INCONSISTENCY-TOLERANT QUERYING OF DESCRIPTION LOGIC KNOWLEDGE BASES

Meghyn Bienvenu

Part of "Logic and Languages" module

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!

 \cdot when $\mathcal K$ unsatisfiable, cert $(q,\mathcal K)$ contains all possible tuples

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!

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

Two approaches to inconsistency handling:

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

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

TBox \mathcal{T}_{univ} :

Prof⊑Fac	Prof ⊑ ∃teaches	$Prof \sqsubseteq \neg Lect$	$Fac \sqsubseteq \neg Course$
Lect ⊑ Fac	Lect $\sqsubseteq \exists$ teaches	$Prof \sqsubseteq \neg Fellow$	
Fellow ⊑ Fac	$\exists teaches^- \sqsubseteq Course$	$Lect\sqsubseteq\negFellow$	

Consider following ABoxes:

\mathcal{A}_1	_	{Prof(a	nna),	Lect(anna)	, Fellow ((alex)	}
-----------------	---	---------	-------	-------	-------	-------------------	--------	---

 $A_2 = \{Prof(anna), Fellow(alex), Lect(alex)\}$

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

Prof(anna)	Lect(anna)	Fac(anna)
Fellow(alex)	Lect(alex)	Fac(alex)

TBox \mathcal{T}_{univ} :

Prof ⊑ Fac	Prof ⊑ ∃teaches	Prof ⊑ ¬Lect	Fac ⊑ ¬Course
Lect ⊑ Fac	Lect $\sqsubseteq \exists$ teaches	$Prof \sqsubseteq \neg Fellow$	
Fellow 드 Fac	\exists teaches $^- \sqsubseteq$ Course	$Lect\sqsubseteq\negFellow$	

ABox \mathcal{A}_{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?

 $\begin{array}{ll} q_1(x) = \operatorname{Fac}(x) & q_2(x) = \exists y \operatorname{teaches}(x, y) \\ q_3(x) = \exists y \operatorname{Fac}(x) \wedge \operatorname{teaches}(x, y) & q_4(x, y) = \operatorname{Fac}(x) \wedge \operatorname{teaches}(x, y) \end{array}$

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

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

- $\cdot\,$ if ${\cal K}$ is satisfiable, should return certain answers
- \cdot 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

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

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

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

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

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

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

- $\cdot\,$ set of query results is jointly consistent with TBox
- \cdot safe to combine query results

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

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

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 $\mathcal{I} \models q(\vec{a})$.

- $\cdot\,$ set of query results is jointly consistent with TBox
- \cdot safe to combine query results

Note: neither property implies the other

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

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

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

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

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

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 A w.r.t. a TBox T
= inclusion-maximal subset of A that is 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 <mark>at least one repair</mark>

 \cdot inconsistent KB \Rightarrow typically multiple repairs

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

Prof⊑Fac	Prof ⊑ ∃teaches	$Prof \sqsubseteq \neg Lect$	Fac ⊑ ¬Course
Lect ⊑ Fac	Lect $\sqsubseteq \exists$ teaches	$Prof \sqsubseteq \neg Fellow$	
Fellow ⊑ Fac	\exists teaches $^- \sqsubseteq$ Course	$Lect\sqsubseteq\negFellow$	

and ABox \mathcal{A}_{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}_{univ} -inconsistent subsets:

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

Question: How many repairs of A_{univ} w.r.t. T_{univ} ?

EXAMPLE: REPAIRS (CONT.)

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

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

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

 $A_{Int} = \{Fellow(alex), teaches(alex, csc486)\}$

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

 \cdot ways to achieve consistency, keeping as much information as possible

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

 $\cdot\,$ ways to achieve consistency, keeping as much information as possible

Plausible answers: hold no matter which repair is chosen

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

 \cdot 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

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



For the query $q_1(x) = Fac(x)$, we have:

For the query $q_1(x) = Fac(x)$, we have:

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

For the query $q_1(x) = Fac(x)$, we have:

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

These are the only answers under AR semantics:

For the query $q_1(x) = Fac(x)$, we have:

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

These are the only answers under AR semantics:

- · $\mathcal{K}_{univ} \not\models_{AR} q_1(julie)$ as $(\mathcal{T}_{univ}, \mathcal{R}_7) \not\models$ Fac(julie)
- · can similarly show $\mathcal{K}_{univ} \not\models_{AR} q_1(csc486)$ and $\mathcal{K}_{univ} \not\models_{AR} q_1(csc343)$

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

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

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

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

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

These are the only answers under AR semantics:

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

Repair: ⊆-maximal subset of the data consistent with the ontology

 \cdot 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

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



Satisfies both CONSISTENT SUPPORT and CONSISTENT RESULTS

Idea: only use the surest assertions to answer queries

 \cdot disregard assertions involved in some contradiction

Idea: only use the surest assertions to answer queries

· disregard assertions involved in some contradiction

IAR semantics: query the intersection of the repairs

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

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

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

 $A_{Int} = \{Fellow(alex), teaches(alex, csc486)\}$

For the query $q_1(x) = Fac(x)$, we have:

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

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

 $A_{Int} = \{Fellow(alex), teaches(alex, csc486)\}$

For the query $q_1(x) = Fac(x)$, we have:

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

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

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

 $A_{Int} = \{Fellow(alex), teaches(alex, csc486)\}$

For the query $q_1(x) = Fac(x)$, we have:

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

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)

Idea: only use the surest assertions to answer queries

· disregard assertions involved in some contradiction

IAR semantics: query the intersection of the repairs

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

Under-approximation of the AR semantics

Satisfies both CONSISTENT SUPPORT and CONSISTENT RESULTS

Idea: return all answers supported by consistent part of data

 \cdot can view them as possible answers, having coherent justification

Idea: return all answers supported by consistent part of data

 \cdot can view them as possible answers, having coherent justification

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

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

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

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

Moving from AR to brave semantics yields an additional answer:

 $\cdot \mathcal{K}_{univ} \models_{brave} q_1(anna)$ AR-answer $\cdot \mathcal{K}_{univ} \models_{brave} q_1(kim)$ AR-answer $\cdot \mathcal{K}_{univ} \models_{brave} q_1(alex)$ AR-answer $\cdot \mathcal{K}_{univ} \models_{brave} q_1(alex)$ AR-answer

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

Moving from AR to brave semantics yields an additional answer:

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

These are the only answers to *q*₁ under brave semantics:

 \cdot csc486 and csc343 cannot be obtained as answers from any repair

Idea: return all answers supported by consistent part of data

 \cdot can view them as possible answers, having coherent justification

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

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

Over-approximation of the **AR semantics**

· ... and every semantics that satisfies CONSISTENT SUPPORT

Does not satisfy CONSISTENT RESULTS

Why?

Idea: some repairs are more likely than others

· exploit knowledge about relative reliability of ABox assertions

Idea: some repairs are more likely than others

• exploit knowledge about relative reliability of ABox assertions

Formally: use preference relation \leq to compare repairs

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

Idea: some repairs are more likely than others

• exploit knowledge about relative reliability of ABox assertions

Formally: use preference relation \leq to compare repairs

- compare w.r.t. cardinality (≤)
- · partition ABox into priority levels $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$
 - · compare level-by-level using set inclusion ($\subseteq_{\mathcal{P}}$)
 - \cdot 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 (\leq -AR, \leq -IAR, \leq -brave)

Results apply to $DL-Lite_{\mathcal{R}}$ and all DL-Lite dialects that satisfy:

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

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

- · every minimal support for $q(\vec{a})$ contains at most |q| assertions
- $\cdot\,$ every minimal $\mathcal{T}\text{-inconsistent}$ subset has cardinality at most two
- CQ answering, instance checking, and KB consistency can be performed by FO query rewriting (so in AC⁰ in data complexity)

- · every minimal support for $q(\vec{a})$ contains at most |q| assertions
- $\cdot\,$ every minimal $\mathcal{T}\text{-inconsistent}$ subset has cardinality at most two
- CQ answering, instance checking, and KB consistency can be performed by FO query rewriting (so in AC⁰ in data complexity)
- · CQ answering is NP-complete for combined complexity

- · every minimal support for $q(\vec{a})$ contains at most |q| assertions
- $\cdot\,$ every minimal $\mathcal{T}\text{-inconsistent}$ subset has cardinality at most two
- CQ answering, instance checking, and KB consistency can be performed by FO query rewriting (so in AC⁰ in data complexity)
- · CQ answering is NP-complete for combined complexity
- · instance checking is NL-complete in combined complexity

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

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 $(\mathcal{T}, \mathcal{A}') \not\models q(\vec{a})$

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 $\varphi = c_1 \land \ldots \land c_m$ over v_1, \ldots, v_k



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

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

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

Reduction from 2+2UNSAT: $\varphi = c_1 \land \ldots \land c_m$ over $v_1, \ldots, v_k, \top, \bot$ each clause has two positive and two negative literals



Can show φ unsatisfiable $\Leftrightarrow \mathcal{T}, \mathcal{A} \models_{\mathsf{AR}} q$

For IAR and brave semantics,

have same low data complexity as classical semantics

Theorem CQ answering under IAR semantics are in AC⁰ in data complexity

Theorem CQ answering under brave semantics are in AC⁰ in data complexity For IAR and brave semantics,

have same low data complexity as classical semantics

Theorem CQ answering under IAR semantics are in AC⁰ in data complexity

Theorem CQ answering under brave semantics are in AC⁰ in data complexity

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

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

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

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

 $q_2'(x) = \text{Prof}(x) \lor \text{Lect}(x) \lor \exists y. \text{teaches}(x,y)$

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

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

 $q'_2(x) = Prof(x) \lor Lect(x) \lor \exists y.teaches(x, y)$

Rewriting of *q*² for IAR semantics:

 $q_2''(x) = \operatorname{Prof}(x) \land (\neg \operatorname{Lect}(x) \land \neg \operatorname{Fellow}(x) \land \neg \operatorname{Course}(x) \land \neg \exists z. \operatorname{teaches}(z, x)) \lor$ $\operatorname{Lect}(x) \land (\neg \operatorname{Prof}(x) \land \neg \operatorname{Fellow}(x) \land \neg \operatorname{Course}(x) \land \neg \exists z. \operatorname{teaches}(z, x)) \lor$ $\exists y.(\operatorname{teaches}(x, y) \land (\neg \operatorname{Prof}(y) \land \neg \operatorname{Lect}(y) \land \neg \operatorname{Fellow}(y)))$

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

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

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

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

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

(Normal) rewriting of $q_2(x) = \exists y \text{ teaches}(x, y) \text{ w.r.t. } \mathcal{T}'_{univ}$: $q'_2(x) = \operatorname{Prof}(x) \lor \operatorname{Lect}(x) \lor \exists y. \operatorname{teaches}(x, y)$ Idea: modify UCQ-rewriting to ensure each disjunct can only match $\mathcal{T}\text{-}\mathsf{consistent}$ subset of ABox

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

(Normal) rewriting of $q_2(x) = \exists y \text{ teaches}(x, y) \text{ w.r.t. } \mathcal{T}'_{\text{univ}}$: $q'_2(x) = \text{Prof}(x) \lor \text{Lect}(x) \lor \exists y \text{.teaches}(x, y)$

Rewriting of *q*² for **brave semantics**:

 $q'_2(x) = Prof(x) \lor Lect(x) \lor (\exists y.teaches(x, y) \land x \neq y)$

to **disallow** using assertions of the form teaches(a, a)

Semantics	Data complexity		Combined complexity	
	CQs	lQs	CQs	IQs
classical	in AC ⁰	in AC ⁰	NP	NL
AR	coNP	coNP	Π_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

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

CQAPri first system for AR query answering in DL-Lite

Implements hybrid approach:

- compute IAR and brave answers
 - gives upper and lower **bounds on AR answers**

polytime

- · 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 $\top, \bot, \neg, \exists r.C$

Semantics	Data complexity		Combined complexity	
	CQs	IQs	CQs	lQs
classical	Р	Р	NP	Р
AR	coNP	coNP	Π_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

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

Expressive DL ALC: constructors $\top, \bot, \neg, \sqcap, \sqcup, \exists r.C, \forall r.C$

Semantics	Data complexity		Combined complexity	
o e mantreo	CQs	IQs	CQs	IQs
classical	coNP	coNP	Exp	Ехр
AR	Π_2^p	Π_2^p	Exp	Exp
IAR	Π_2^p	Π_2^p	Exp	Exp
brave	Σ_2^p	Σ_2^p	Exp	Exp

Observe:

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