Description Logics and Reasoning on Data
Inconsistency Handling

C. Bourgaux, M. Thomazo

Outline

Introduction

Inconsistency-tolerant semantics

Complexity issues

A practical approach for AR semantics

Some research problems

References

Handling Inconsistent Data

In real world data often contains errors

- human errors
- automatic extraction
- outdated information

Likely to be inconsistent with the ontology (today: focus on the case where the ontology is assumed reliable)

Standard semantics: everything is entailed from an inconsistent knowledge base !

Handling Inconsistent Data

In real world data often contains errors

- human errors
- automatic extraction
- outdated information

Likely to be inconsistent with the ontology (today: focus on the case where the ontology is assumed reliable)

Standard semantics: everything is entailed from an inconsistent knowledge base !

It is not always possible to resolve the inconsistencies (lack of information, time, permission...)

Alternative semantics: meaningful answers to queries despite inconsistencies

Example

$$\mathcal{T} = \{\mathsf{AProf} \sqsubseteq \mathsf{Prof},\ \mathsf{FProf} \sqsubseteq \mathsf{Prof},\ \mathsf{AProf} \sqsubseteq \neg\mathsf{FProf}\}$$
$$\mathcal{A} = \{\mathsf{AProf}(ann),\ \mathsf{FProf}(ann),\ \mathsf{Postdoc}(alex)\}$$

Which assertions would it be reasonable to infer ?

Inconsistency-Tolerant Semantics

Many inconsistency-tolerant semantics have been proposed

A semantics S associates a set of answers to every KB and query

- if the KB is satisfiable, should return certain answers
- for unsatisfiable KBs, give different answers than classical semantics

Write $\langle \mathcal{T}, \mathcal{A} \rangle \models_S q(\vec{a})$ if $\vec{a}$ is an answer to q w.r.t. $\langle \mathcal{T}, \mathcal{A} \rangle$ under semantics S

Consistency Properties

A $\mathcal{T}$-support of $q(\vec{a})$ is a subset $C \subseteq \mathcal{A}$ such that

- $\langle \mathcal{T}, C \rangle$ is satisfiable
- $\langle \mathcal{T}, C \rangle \models q(\vec{a})$

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

- consistent explanation/justification for the query result

Consistency Properties

A $\mathcal{T}$-support of $q(\vec{a})$ is a subset $C \subseteq \mathcal{A}$ such that

- $\langle \mathcal{T}, C \rangle$ is satisfiable
- $\langle \mathcal{T}, C \rangle \models q(\vec{a})$

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

- consistent explanation/justification for the query result

Semantics S satisfies the consistent results property if for every KB $\langle \mathcal{T}, \mathcal{A} \rangle$, there exists a model $\mathcal{I}$ of $\mathcal{T}$ such that $\langle \mathcal{T}, \mathcal{A} \rangle \models_S q(\vec{a})$ implies $\mathcal{I} \models q(\vec{a})$

- set of query results is jointly consistent with the ontology
- safe to combine query results

Comparing Semantics

Given two semantics S and S'

- S' is an under-approximation (or sound approximation) of S if $\langle \mathcal{T}, \mathcal{A} \rangle \models_{S'} q(\vec{a})$ implies $\langle \mathcal{T}, \mathcal{A} \rangle \models_{S} q(\vec{a})$
- S' is an over-approximation (or complete approximation) of S if $\langle \mathcal{T}, \mathcal{A} \rangle \models_{S} q(\vec{a})$ implies $\langle \mathcal{T}, \mathcal{A} \rangle \models_{S'} q(\vec{a})$

Repairs

Many semantics are based upon the notion of repair:
inclusion-maximal subset of the data consistent with the ontology

Possible worlds, different ways of achieving consistency while retaining as much of the original data as possible

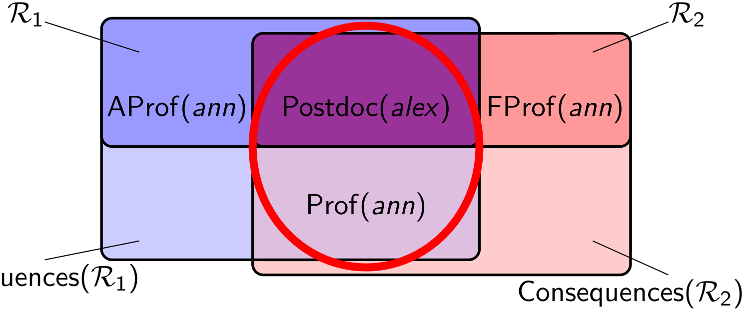
TBox $\mathcal{T}$	ABox $\mathcal{A}$	Repair $\mathcal{R}_1$	Repair $\mathcal{R}_2$
$\mathsf{AProf} \sqsubseteq \mathsf{Prof}$	$\mathsf{AProf}(ann)$	$\mathsf{AProf}(ann)$	
$\mathsf{FProf} \sqsubseteq \mathsf{Prof}$	$\mathsf{FProf}(ann)$		$\mathsf{FProf}(ann)$
$\mathsf{AProf} \sqsubseteq \neg\mathsf{FProf}$	$\mathsf{Postdoc}(alex)$	$\mathsf{Postdoc}(alex)$	$\mathsf{Postdoc}(alex)$

Plausible Answers: AR Semantics

AR (ABox Repair) answers: hold no matter which repair is chosen

$$\langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q(\vec{a}) \Leftrightarrow \langle \mathcal{T}, \mathcal{R} \rangle \models q(\vec{a}) \text{ for every repair } \mathcal{R}$$

TBox $\mathcal{T}$	ABox $\mathcal{A}$	Repair $\mathcal{R}_1$	Repair $\mathcal{R}_2$
$\mathsf{AProf} \sqsubseteq \mathsf{Prof}$	AProf(*ann*)	AProf(*ann*)	
$\mathsf{FProf} \sqsubseteq \mathsf{Prof}$	FProf(*ann*)		FProf(*ann*)
$\mathsf{AProf} \sqsubseteq \neg\mathsf{FProf}$	Postdoc(*alex*)	Postdoc(*alex*)	Postdoc(*alex*)

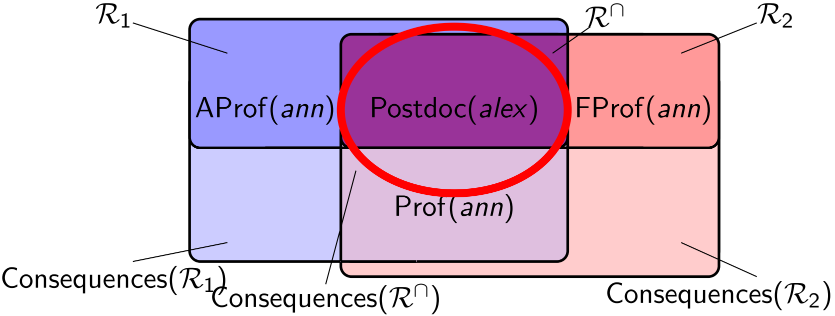


Surest Answers: IAR Semantics

IAR (Intersection AR) answers: hold in the repairs intersection

$$\langle \mathcal{T}, \mathcal{A} \rangle \models_{IAR} q(\vec{a}) \Leftrightarrow \langle \mathcal{T}, \mathcal{R}^{\cap} \rangle \models q(\vec{a}) \text{ with } \mathcal{R}^{\cap} \text{ repairs intersection}$$

TBox $\mathcal{T}$	ABox $\mathcal{A}$	Repair $\mathcal{R}_1$	Repair $\mathcal{R}_2$
AProf $\sqsubseteq$ Prof	AProf(*ann*)	AProf(*ann*)	
FProf $\sqsubseteq$ Prof	FProf(*ann*)		FProf(*ann*)
AProf $\sqsubseteq$ ¬FProf	Postdoc(*alex*)	Postdoc(*alex*)	Postdoc(*alex*)

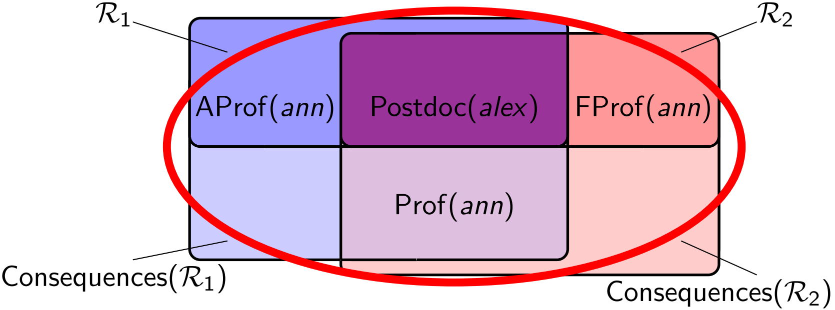


Possible Answers: Brave Semantics

Brave answers: hold in some repair

$$\langle \mathcal{T}, \mathcal{A} \rangle \models_{brave} q(\vec{a}) \Leftrightarrow \langle \mathcal{T}, \mathcal{R} \rangle \models q(\vec{a}) \text{ for some } \mathcal{R}$$

TBox $\mathcal{T}$	ABox $\mathcal{A}$	Repair $\mathcal{R}_1$	Repair $\mathcal{R}_2$
AProf $\sqsubseteq$ Prof	AProf(*ann*)	AProf(*ann*)	
FProf $\sqsubseteq$ Prof	FProf(*ann*)		FProf(*ann*)
AProf $\sqsubseteq$ $\neg$FProf	Postdoc(*alex*)	Postdoc(*alex*)	Postdoc(*alex*)



AR, IAR and Brave Semantics

Which consistency properties are satisfied by AR, IAR, brave ?

- consistent support property ?
- consistent results property ?

How do the three semantics compare ?

- under/over-approximation

AR, IAR and Brave Semantics

- AR is the most well-known and accepted semantics
  - cautious reasoning used in many area (belief revision...)
  - consistent query answering in databases
- but AR is usually intractable ($\textsc{coNP}$-complete in data complexity for DL-Lite and $\mathcal{EL}$)
- IAR and brave are under- and over-approximations of AR
  - IAR most cautious: disregard all facts involved in some contradiction
  - brave least cautious: all answers supported by some consistent set of facts
- IAR and brave are tractable for DL-Lite

Some Other Inconsistency-Tolerant Semantics

- *k*-support semantics
  - fine-grained under-approximation of AR
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-supp} q(\vec{a})$ iff there exist $C_1, \ldots, C_k$ $\mathcal{T}$-supports of $q(\vec{a})$ such that every repair contains at least one of the C_i
  - 1-support = IAR
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-supp} q(\vec{a}) \Rightarrow \langle \mathcal{T}, \mathcal{A} \rangle \models_{k+1-supp} q(\vec{a})$
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q(\vec{a}) \Leftrightarrow \exists k \geq 1, \langle \mathcal{T}, \mathcal{A} \rangle \models_{k-supp} q(\vec{a})$

Some Other Inconsistency-Tolerant Semantics

- k-support semantics
  - fine-grained under-approximation of AR
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-supp} q(\vec{a})$ iff there exist $C_1, \ldots, C_k$ $\mathcal{T}$-supports of $q(\vec{a})$ such that every repair contains at least one of the C_i
  - 1-support $=$ IAR
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-supp} q(\vec{a}) \Rightarrow \langle \mathcal{T}, \mathcal{A} \rangle \models_{k+1-supp} q(\vec{a})$
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q(\vec{a}) \Leftrightarrow \exists k \geq 1, \langle \mathcal{T}, \mathcal{A} \rangle \models_{k-supp} q(\vec{a})$
- k-defeater semantics
  - fine-grained over-approximation of AR
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-def} q(\vec{a})$ iff there does not exist a $\mathcal{T}$-consistent $S \subseteq \mathcal{A}$ such that $|S| \leq k$ and $\langle \mathcal{T}, S \cup C \rangle \models \bot$ for every minimal $\mathcal{T}$-support C of $q(\vec{a})$
  - 0-defeater $=$ brave
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k+1-def} q(\vec{a}) \Rightarrow \langle \mathcal{T}, \mathcal{A} \rangle \models_{k-def} q(\vec{a})$
  - for every KB, there exists k such that $\langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q(\vec{a}) \Leftrightarrow \langle \mathcal{T}, \mathcal{A} \rangle \models_{k-def} q(\vec{a})$

Some Other Inconsistency-Tolerant Semantics

- ICR (Intersection Closed Repairs) semantics
  - under-approximation of AR and over-approximation of IAR
  - intersects the closures of the repairs (closure of $\mathcal{R}$ = set of assertions entailed from $\langle \mathcal{T}, \mathcal{R} \rangle$)
  - same as AR for queries without quantifier

Some Other Inconsistency-Tolerant Semantics

- ICR (Intersection Closed Repairs) semantics
  - under-approximation of AR and over-approximation of IAR
  - intersects the closures of the repairs (closure of $\mathcal{R}$ = set of assertions entailed from $\langle \mathcal{T}, \mathcal{R} \rangle$)
  - same as AR for queries without quantifier
- CAR and ICAR semantics
  - define semantics that are (almost) syntax-independent
  - apply closure operator on original ABox
  - need alternative notion of closure for inconsistent KB: set of assertions with a $\mathcal{T}$-support in $\mathcal{A}$
  - closed ABox repairs: maximally complete standard ABox repairs with facts from the closure of $\mathcal{A}$
  - apply AR (CAR) or IAR (ICAR) using closed ABox repairs
  - do not satisfy consistent support !
    $\mathcal{T} = \{A \sqsubseteq B, C \sqsubseteq D, A \sqsubseteq \neg C\}$, $\mathcal{A} = \{A(a), C(a)\}$,
    $q = B(x) \wedge D(x)$

Exercise: AR, IAR, brave, k-supp, k-def, ICR ?

$$
\begin{aligned}
\mathcal{T} = \{ & \mathsf{AProf} \sqsubseteq \mathsf{Prof},\ \mathsf{FProf} \sqsubseteq \mathsf{Prof},\ \mathsf{Prof} \sqsubseteq \mathsf{PhD},\ \mathsf{Postdoc} \sqsubseteq \mathsf{PhD}, \\
& \mathsf{PhD} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach}^- \sqsubseteq \mathsf{Course}, \\
& \mathsf{Prof} \sqsubseteq \exists \mathsf{WorkFor}, \mathsf{Student} \sqsubseteq \exists \mathsf{MemberOf}, \mathsf{WorkFor} \sqsubseteq \mathsf{MemberOf}, \\
& \mathsf{AProf} \sqsubseteq \neg \mathsf{FProf},\ \mathsf{Prof} \sqsubseteq \neg \mathsf{Postdoc},\ \mathsf{Student} \sqsubseteq \neg \mathsf{Prof}, \\
& \mathsf{Person} \sqsubseteq \neg \mathsf{Course},\ \exists \mathsf{MemberOf}^- \sqsubseteq \neg \mathsf{Postdoc} \}
\end{aligned}
$$

Exercise: AR, IAR, brave, k-supp, k-def, ICR ?

$$\begin{aligned}\mathcal{T} = \{&\mathsf{AProf} \sqsubseteq \mathsf{Prof},\ \mathsf{FProf} \sqsubseteq \mathsf{Prof},\ \mathsf{Prof} \sqsubseteq \mathsf{PhD},\ \mathsf{Postdoc} \sqsubseteq \mathsf{PhD},\\ &\mathsf{PhD} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach}^- \sqsubseteq \mathsf{Course},\\ &\mathsf{Prof} \sqsubseteq \exists \mathsf{WorkFor}, \mathsf{Student} \sqsubseteq \exists \mathsf{MemberOf}, \mathsf{WorkFor} \sqsubseteq \mathsf{MemberOf},\\ &\mathsf{AProf} \sqsubseteq \neg \mathsf{FProf},\ \mathsf{Prof} \sqsubseteq \neg \mathsf{Postdoc},\ \mathsf{Student} \sqsubseteq \neg \mathsf{Prof},\\ &\mathsf{Person} \sqsubseteq \neg \mathsf{Course},\ \exists \mathsf{MemberOf}^- \sqsubseteq \neg \mathsf{Postdoc}\}\end{aligned}$$

$$\begin{aligned}\mathcal{A}_a = \{&\mathsf{AProf}(ann), \mathsf{FProf}(ann), \mathsf{Prof}(ann),\\ &\mathsf{Teach}(ann, c_a), \mathsf{Teach}(ann, ann)\}\end{aligned}$$

$$q(x) = \exists yz \mathsf{PhD}(x) \wedge \mathsf{MemberOf}(x, y) \wedge \mathsf{Teach}(x, z)$$

Exercise: AR, IAR, brave, k-supp, k-def, ICR ?

$$
\begin{aligned}
\mathcal{T} = \{ & \mathsf{AProf} \sqsubseteq \mathsf{Prof},\ \mathsf{FProf} \sqsubseteq \mathsf{Prof},\ \mathsf{Prof} \sqsubseteq \mathsf{PhD},\ \mathsf{Postdoc} \sqsubseteq \mathsf{PhD},\\
& \mathsf{PhD} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach}^- \sqsubseteq \mathsf{Course},\\
& \mathsf{Prof} \sqsubseteq \exists \mathsf{WorkFor}, \mathsf{Student} \sqsubseteq \exists \mathsf{MemberOf}, \mathsf{WorkFor} \sqsubseteq \mathsf{MemberOf},\\
& \mathsf{AProf} \sqsubseteq \neg \mathsf{FProf},\ \mathsf{Prof} \sqsubseteq \neg \mathsf{Postdoc},\ \mathsf{Student} \sqsubseteq \neg \mathsf{Prof},\\
& \mathsf{Person} \sqsubseteq \neg \mathsf{Course},\ \exists \mathsf{MemberOf}^- \sqsubseteq \neg \mathsf{Postdoc} \}
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{A}_b = \{ & \mathsf{AProf}(bob), \mathsf{FProf}(bob), \mathsf{Postdoc}(bob),\\
& \mathsf{MemberOf}(bob, dpt), \mathsf{Teach}(bob, c_b) \}
\end{aligned}
$$

$$
q(x) = \exists yz \mathsf{PhD}(x) \wedge \mathsf{MemberOf}(x, y) \wedge \mathsf{Teach}(x, z)
$$

Exercise: AR, IAR, brave, k-supp, k-def, ICR ?

$$
\begin{aligned}
\mathcal{T} = \{ & \mathsf{AProf} \sqsubseteq \mathsf{Prof},\ \mathsf{FProf} \sqsubseteq \mathsf{Prof},\ \mathsf{Prof} \sqsubseteq \mathsf{PhD},\ \mathsf{Postdoc} \sqsubseteq \mathsf{PhD},\\
& \mathsf{PhD} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach}^- \sqsubseteq \mathsf{Course},\\
& \mathsf{Prof} \sqsubseteq \exists \mathsf{WorkFor}, \mathsf{Student} \sqsubseteq \exists \mathsf{MemberOf}, \mathsf{WorkFor} \sqsubseteq \mathsf{MemberOf},\\
& \mathsf{AProf} \sqsubseteq \neg \mathsf{FProf},\ \mathsf{Prof} \sqsubseteq \neg \mathsf{Postdoc},\ \mathsf{Student} \sqsubseteq \neg \mathsf{Prof},\\
& \mathsf{Person} \sqsubseteq \neg \mathsf{Course},\ \exists \mathsf{MemberOf}^- \sqsubseteq \neg \mathsf{Postdoc} \}
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{A}_c = \{ & \mathsf{AProf}(carl), \mathsf{Teach}(carl, c_{c1}), \mathsf{Teach}(carl, c_{c2}),\\
& \mathsf{Teach}(c_{c1}, c_{c2}), \mathsf{Teach}(c_{c2}, c_{c1}) \}
\end{aligned}
$$

$$
q(x) = \exists yz \mathsf{PhD}(x) \wedge \mathsf{MemberOf}(x, y) \wedge \mathsf{Teach}(x, z)
$$

Exercise: AR, IAR, brave, k-supp, k-def, ICR ?

$$\begin{aligned}\mathcal{T} = \{&\mathsf{AProf} \sqsubseteq \mathsf{Prof},\ \mathsf{FProf} \sqsubseteq \mathsf{Prof},\ \mathsf{Prof} \sqsubseteq \mathsf{PhD},\ \mathsf{Postdoc} \sqsubseteq \mathsf{PhD},\\ &\mathsf{PhD} \sqsubseteq \mathsf{Person},\ \exists\mathsf{Teach} \sqsubseteq \mathsf{Person},\ \exists\mathsf{Teach}^- \sqsubseteq \mathsf{Course},\\ &\mathsf{Prof} \sqsubseteq \exists\mathsf{WorkFor}, \mathsf{Student} \sqsubseteq \exists\mathsf{MemberOf}, \mathsf{WorkFor} \sqsubseteq \mathsf{MemberOf},\\ &\mathsf{AProf} \sqsubseteq \neg\mathsf{FProf},\ \mathsf{Prof} \sqsubseteq \neg\mathsf{Postdoc},\ \mathsf{Student} \sqsubseteq \neg\mathsf{Prof},\\ &\mathsf{Person} \sqsubseteq \neg\mathsf{Course},\ \exists\mathsf{MemberOf}^- \sqsubseteq \neg\mathsf{Postdoc}\}\end{aligned}$$

$$\begin{aligned}\mathcal{A}_d = \{&\mathsf{AProf}(dan), \mathsf{Teach}(dan, c_{d1}), \mathsf{Teach}(dan, c_{d2}),\\ &\mathsf{AProf}(c_{d1}), \mathsf{AProf}(c_{d2})\}\end{aligned}$$

$$q(x) = \exists yz \mathsf{PhD}(x) \wedge \mathsf{MemberOf}(x, y) \wedge \mathsf{Teach}(x, z)$$

Some Complexity Results for DL-Lite

The DL-Lite family and OWL 2 QL

- OWL 2 QL : OWL 2 profile for efficient query answering
- Target large datasets: CQ answering is in AC0 in data complexity (AC0 $\subseteq$ LogSpace $\subseteq$ PTime)
  - via query rewriting
- Based on the DL-Lite$_{\mathcal{R}}$ language of the DL-Lite family

DL-Lite$_{\text{core}}$: concept inclusions of the form $B \sqsubseteq C$ where

$$C := B \mid \neg B, \quad B := A \mid \exists S, \quad S := R \mid R^{-}$$

with A an atomic concept and R an atomic role

We focus on DL-Lite$_{\mathcal{R}}$:

- DL-Lite$_{\mathcal{R}}$ = DL-Lite$_{\text{core}}$ + role inclusions $S \sqsubseteq Q$ with $Q := S \mid \neg S$

Some Complexity Results for DL-Lite

Complexity results will apply to all languages that satisfy

- minimal $\mathcal{T}$-supports for $q(\vec{a})$ contain at most $|q|$ assertions
- minimal $\mathcal{T}$-inconsistent subsets have bounded cardinality
  - in DL-Lite: bounded by 2
- CQ answering and satisfiability can be performed by FO rewriting (so in AC0 $\subseteq$ PTIME in data complexity)

Complexity of AR in DL-Lite

CQ entailment under AR semantics is $\textsc{coNP}$-complete in data complexity

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

- $\langle \mathcal{T}, \mathcal{R} \rangle \not\models q(\vec{a})$ in AC0
- repair checking in $\textsc{PTime}$?

Complexity of AR in DL-Lite

CQ entailment under AR semantics is coNP-complete in data complexity

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

- $\langle \mathcal{T}, \mathcal{R} \rangle \not\models q(\vec{a})$ in AC0
- repair checking in PTime?

Lower bound: by reduction from propositional unsatisfiability

- UNSAT: Given $\varphi = C_1 \wedge \cdots \wedge C_m$ conjunction of clauses over propositional variables $x_1, \ldots, x_k$, decide whether φ is unsatisfiable
- build a KB $\langle \mathcal{T}, \mathcal{A} \rangle$ and query $q(\vec{a})$ such that φ is unsatisfiable iff $\langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q(\vec{a})$
  - $\mathcal{T} = \{\exists P^- \sqsubseteq \neg\exists N^-,\ \exists P \sqsubseteq \neg\exists U^-,\ \exists N \sqsubseteq \neg\exists U^-, \exists U \sqsubseteq A\}$
  - $\mathcal{A} = \{P(c_j, x_i) \mid x_i \in C_j\} \cup \{N(c_j, x_i) \mid \neg x_i \in C_j\} \cup \{U(a, c_j) \mid 1 \leq j \leq m\}$
  - $q = A(a)$

Complexity of IAR and Brave in DL-Lite

CQ entailment under IAR and brave semantics is in PTIME in data complexity

Any idea of PTIME algorithms ?

Complexity of IAR and Brave in DL-Lite

CQ entailment under IAR and brave semantics is in PTIME in data complexity

Any idea of PTIME algorithms ?

Actually, CQ entailment under IAR and brave semantics is in AC0 in data complexity

Can use FO-rewriting to compute IAR and brave answers

FO Rewriting for IAR Semantics

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

$$\begin{aligned}\mathcal{T} = \{&\mathsf{AProf} \sqsubseteq \mathsf{Prof},\ \mathsf{FProf} \sqsubseteq \mathsf{Prof},\ \mathsf{Prof} \sqsubseteq \mathsf{PhD},\ \mathsf{Postdoc} \sqsubseteq \mathsf{PhD},\\ &\mathsf{PhD} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach} \sqsubseteq \mathsf{Person},\ \exists \mathsf{Teach}^- \sqsubseteq \mathsf{Course},\\ &\mathsf{Prof} \sqsubseteq \exists \mathsf{WorkFor}, \mathsf{Student} \sqsubseteq \exists \mathsf{MemberOf}, \mathsf{WorkFor} \sqsubseteq \mathsf{MemberOf},\\ &\mathsf{AProf} \sqsubseteq \neg \mathsf{FProf},\ \mathsf{Prof} \sqsubseteq \neg \mathsf{Postdoc},\ \mathsf{Student} \sqsubseteq \neg \mathsf{Prof},\\ &\mathsf{Person} \sqsubseteq \neg \mathsf{Course},\ \exists \mathsf{MemberOf}^- \sqsubseteq \neg \mathsf{Postdoc}\}\end{aligned}$$

$$q_1(x) = \mathsf{PhD}(x)$$

$$q_2(x) = \exists y \mathsf{MemberOf}(x, y)$$

$$q_3(x) = \exists y \mathsf{Prof}(x) \wedge \mathsf{Teach}(x, y)$$

FO Rewriting for Brave Semantics

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

$$\begin{aligned}\mathcal{T} = \{&\mathsf{AProf} \sqsubseteq \mathsf{Prof},\ \mathsf{FProf} \sqsubseteq \mathsf{Prof},\ \mathsf{Prof} \sqsubseteq \mathsf{PhD},\ \mathsf{Postdoc} \sqsubseteq \mathsf{PhD},\\ &\mathsf{PhD} \sqsubseteq \mathsf{Person},\ \exists\mathsf{Teach} \sqsubseteq \mathsf{Person},\ \exists\mathsf{Teach}^- \sqsubseteq \mathsf{Course},\\ &\mathsf{Prof} \sqsubseteq \exists\mathsf{WorkFor}, \mathsf{Student} \sqsubseteq \exists\mathsf{MemberOf}, \mathsf{WorkFor} \sqsubseteq \mathsf{MemberOf},\\ &\mathsf{AProf} \sqsubseteq \neg\mathsf{FProf},\ \mathsf{Prof} \sqsubseteq \neg\mathsf{Postdoc},\ \mathsf{Student} \sqsubseteq \neg\mathsf{Prof},\\ &\mathsf{Person} \sqsubseteq \neg\mathsf{Course},\ \exists\mathsf{MemberOf}^- \sqsubseteq \neg\mathsf{Postdoc}\}\end{aligned}$$

$$q_1(x) = \exists y \mathsf{PhD}(x) \wedge \mathsf{MemberOf}(x, y)$$

$$q_2(x) = \exists y \mathsf{Prof}(x) \wedge \mathsf{Teach}(x, y)$$

More FO Rewritings

k-support and k-defeater semantics are also FO-rewritable. Any idea for the general shape of the rewritings ?

Complexity Picture for DL-Lite

	Data Complexity		Combined Complexity	
	CQs	**IQs**	**CQs**	**IQs**
classical	in AC0	in AC0	NP-co	in PTIME
AR	coNP-co	coNP-co	Π_2^P-co	coNP-co
IAR	in AC0	in AC0	NP-co	in PTIME
brave	in AC0	in AC0	NP-co	in PTIME
k-**support**	in AC0	in AC0	NP-co	in PTIME
k-**defeater**	in AC0	in AC0	NP-co	in PTIME
ICR	coNP-co	coNP-co	$\Delta_2^P[O log(n)]$-co	coNP-co
CAR	coNP-co	in AC0	Π_2^P-co	in PTIME
ICAR	in AC0	in AC0	NP-co	in PTIME

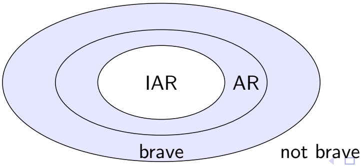
Note on AC0 cases:

- FO-rewritings, but rewritings may be huge and not efficiently evaluated over databases
- alternative PTIME algorithms based on supports and conflicts may be more efficient in practice

A Practical Approach for AR Semantics

- Precompute the conflicts : minimal subsets of the ABox inconsistent with the TBox (of size at most 2 in DL-Lite)
- Compute the minimal $\mathcal{T}$-supports of the query
- Exploit tractable approximations:
  - IAR $\Rightarrow$ AR and not brave $\Rightarrow$ not AR
  - decide IAR/not brave using the $\mathcal{T}$-supports and conflicts
- For remaining cases (brave and not IAR): reduce AR entailment to SAT and use a SAT solver
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models q$ iff φ is unsatisfiable

$$\varphi = \Big(\bigwedge_{C \in \mathcal{T}\text{-supp}} \; \bigvee_{\substack{\alpha \in C, \\ \{\alpha,\beta\} \in \mathit{conflicts}}} x_\beta \Big) \wedge \Big(\bigwedge_{\{\alpha,\beta\} \in \mathit{conflicts}} \neg x_\alpha \vee \neg x_\beta \Big)$$



Examples of Research Problems

- Alternative semantics or repairs
  - taking into account qualitative/quantitative information on data quality: priority, probabilities...
  - case where the TBox may not be correct: general repairs that modify the TBox, soft constraints...
- Practical algorithms, implementations, experimental studies
  - languages with unbounded size of query supports and conflicts
  - impact of the data structure
- Explanations of query results
- Improving data quality, helping user to resolve inconsistencies
- Extending the framework: temporal data, fuzzy data...

Examples of Research Problems

Semantics based upon preferred repairs

Idea: some repairs are more likely than others

Defined preferred repairs based on

- cardinality
- priority levels
- weights
- ...

AR/IAR/brave/... semantics based upon most preferred repairs

Using preferred repairs generally (but not always) increases the computational complexity

Examples of Research Problems

Explanations

Idea: explain the user why a query is entailed (or not) under a given semantics

- AR semantics
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q$: minimal set $\{C_1, \dots, C_k\}$ of minimal $\mathcal{T}$-supports for q such that every repair contains at least one of the C_i
  - $\langle \mathcal{T}, \mathcal{A} \rangle \not\models_{AR} q$: minimal $\mathcal{B} \subseteq \mathcal{A}$ such that $\mathcal{B}$ is $\mathcal{T}$-consistent and for every $\mathcal{T}$-support C of q, $\mathcal{B} \cup C$ is not $\mathcal{T}$-consistent
- IAR semantics
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{IAR} q$: minimal $\mathcal{T}$-support included in every repair
  - $\langle \mathcal{T}, \mathcal{A} \rangle \not\models_{IAR} q$: minimal $\mathcal{B} \subseteq \mathcal{A}$ such that for every $\mathcal{T}$-support C of q, there exists $\mathcal{B}' \subseteq \mathcal{B}$ such that $\mathcal{B}'$ is $\mathcal{T}$-consistent and $\mathcal{B}' \cup C$ is not $\mathcal{T}$-consistent

Basic explanations that should be completed (with some TBox axioms/reasoning steps/conflicting assertions...)

References

- Bienvenu, Bourgaux (RW 2016): Inconsistency-Tolerant Querying of Description Logic Knowledge Bases (https://hal.inria.fr/hal-01633000/file/BieBou-RW16.pdf).