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- outdated information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- outdated information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

\[ \mathcal{T} = \{ \text{APProf} \sqsubseteq \text{Prof}, \text{FProf} \sqsubseteq \text{Prof}, \text{APProf} \sqsubseteq -\text{FProf} \} \]

\[ \mathcal{A} = \{ \text{APProf}(ann), \text{FProf}(ann), \text{Postdoc}(alex) \} \]

Which assertions would it be reasonable to infer?
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 $\langle T, A \rangle \models_S q(\vec{a})$ if $\vec{a}$ is an answer to $q$ w.r.t. $\langle T, A \rangle$ under semantics $S$. 
Consistency Properties

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

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

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

- consistent explanation/justification for the query result

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

- set of query results is jointly consistent with the ontology
- safe to combine query results
A $T$-support of $q(\vec{a})$ is a subset $C \subseteq A$ such that

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

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

- consistent explanation/justification for the query result

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

- 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
  $\langle T, A \rangle \models_{S'} q(\bar{a})$ implies $\langle T, A \rangle \models_S q(\bar{a})$

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

<table>
<thead>
<tr>
<th>TBox</th>
<th></th>
<th>ABox</th>
<th></th>
<th>Repair</th>
<th></th>
<th>Repair</th>
</tr>
</thead>
<tbody>
<tr>
<td>AProf ⊑ Prof</td>
<td></td>
<td>AProf(ann)</td>
<td></td>
<td>AProf(ann)</td>
<td></td>
<td>FProf(ann)</td>
</tr>
<tr>
<td>FProf ⊑ Prof</td>
<td></td>
<td>FProf(ann)</td>
<td></td>
<td>Postdoc(alex)</td>
<td></td>
<td>Postdoc(alex)</td>
</tr>
<tr>
<td>AProf ⊑ ¬FProf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Plausible Answers: AR Semantics

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

\[ \langle T, A \rangle \models_{AR} q(\vec{a}) \iff \langle T, R \rangle \models q(\vec{a}) \text{ for every repair } R \]

\[
\begin{array}{|l|l|}
\hline
\text{TBox} & \mathcal{T} \\
\hline
\text{AProf} \sqsubseteq \text{Prof} & \\
\text{FProf} \sqsubseteq \text{Prof} & \\
\text{AProf} \sqsubseteq \neg \text{FProf} & \\
\hline
\end{array}
\]

\[
\begin{array}{|l|l|}
\hline
\text{ABox} & \mathcal{A} \\
\hline
\text{AProf}(\text{ann}) & \\
\text{FProf}(\text{ann}) & \\
\text{Postdoc}(\text{alex}) & \\
\hline
\end{array}
\]

\[
\begin{array}{|l|l|}
\hline
\text{Repair} & \mathcal{R}_1 \\
\hline
\text{AProf}(\text{ann}) & \\
\text{Postdoc}(\text{alex}) & \\
\hline
\end{array}
\]

\[
\begin{array}{|l|l|}
\hline
\text{Repair} & \mathcal{R}_2 \\
\hline
\text{FProf}(\text{ann}) & \\
\text{Postdoc}(\text{alex}) & \\
\hline
\end{array}
\]

\[
\begin{array}{|l|l|}
\hline
\text{Consequences}(\mathcal{R}_1) & \\
\hline
\end{array}
\]

\[
\begin{array}{|l|l|}
\hline
\text{Consequences}(\mathcal{R}_2) & \\
\hline
\end{array}
\]
Surest Answers: IAR Semantics

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

$$\langle \mathcal{T}, \mathcal{A} \rangle \models_{IAR} q(\bar{a}) \iff \langle \mathcal{T}, \mathcal{R}^\cap \rangle \models q(\bar{a})$$ with $$\mathcal{R}^\cap$$ repairs intersection
Possible Answers: Brave Semantics

Brave answers: hold in some repair

\[ \langle T, A \rangle \models_{\text{brave}} q(\vec{a}) \iff \langle T, R \rangle \models q(\vec{a}) \text{ for some } R \]

- **TBox**
  - T
  - AProf $\sqsubseteq$ Prof
  - FProf $\sqsubseteq$ Prof
  - AProf $\sqsubseteq \neg$FProf

- **ABox**
  - A
  - AProf(ann)
  - FProf(ann)
  - Postdoc(alex)

- **Repair**
  - R
  - Repair $R_1$
    - AProf(ann)
    - Postdoc(alex)
  - Repair $R_2$
    - FProf(ann)
    - Postdoc(alex)

- **Consequences**
  - Consequences($R_1$)
  - Consequences($R_2$)
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
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-supp} q(\bar{a})$ iff there exist $C_1, \ldots, C_k$ $\mathcal{T}$-supports of $q(\bar{a})$ such that every repair contains at least one of the $C_i$
  - 1-support = IAR
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-supp} q(\bar{a}) \Rightarrow \langle \mathcal{T}, \mathcal{A} \rangle \models_{k+1-supp} q(\bar{a})$
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q(\bar{a}) \iff \exists k \geq 1, \langle \mathcal{T}, \mathcal{A} \rangle \models_{k-supp} q(\bar{a})$
Some Other Inconsistency-Tolerant Semantics

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

- **k-defeater semantics**
  - fine-grained over-approximation of AR
  - \( \langle T, A \rangle \models_{k-def} q(\bar{a}) \) iff there does not exist a \( T \)-consistent \( S \subseteq A \) such that \( |S| \leq k \) and \( \langle T, S \cup C \rangle \models \bot \) for every minimal \( T \)-support \( C \) of \( q(\bar{a}) \)
  - 0-defeater = brave
  - \( \langle T, A \rangle \models_{k+1-def} q(\bar{a}) \Rightarrow \langle T, A \rangle \models_{k-def} q(\bar{a}) \)
  - for every KB, there exists \( k \) such that \( \langle T, A \rangle \models_{AR} q(\bar{a}) \Leftrightarrow \langle T, A \rangle \models_{k-def} q(\bar{a}) \)
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 \((\text{closure of } \mathcal{R} = \text{set of assertions entailed from } \langle \mathcal{T}, \mathcal{R} \rangle)\)
  - 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 $\mathcal{R} = \text{set of assertions entailed from } \langle \mathcal{T}, \mathcal{R} \rangle$)
  - 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!

$$\mathcal{T} = \{ A \sqsubseteq B, C \sqsubseteq D, A \sqsubseteq \neg C \}, \quad \mathcal{A} = \{ A(a), C(a) \}, \quad q = B(x) \land D(x)$$
Exercise: AR, IAR, brave, $k$-supp, $k$-def, ICR?

$T = \{ \text{AP} \subseteq \text{Prof}, \text{F} \subseteq \text{Prof}, \text{Prof} \subseteq \text{PhD}, \text{Postdoc} \subseteq \text{PhD}, \text{PhD} \subseteq \text{Person}, \exists \text{Teach} \subseteq \text{Person}, \exists \text{Teach}^{-} \subseteq \text{Course}, \text{Prof} \subseteq \exists \text{WorkFor}, \text{Student} \subseteq \exists \text{MemberOf}, \text{WorkFor} \subseteq \text{MemberOf}, \text{AP} \subseteq \neg \text{F}, \text{Prof} \subseteq \neg \text{Postdoc}, \text{Student} \subseteq \neg \text{Prof}, \text{Person} \subseteq \neg \text{Course}, \exists \text{MemberOf}^{-} \subseteq \neg \text{Postdoc} \}$
Exercise: AR, IAR, brave, $k$-supp, $k$-def, ICR ?

$T = \{ \text{AProf} \sqsubseteq \text{Prof}, \text{FProf} \sqsubseteq \text{Prof}, \text{Prof} \sqsubseteq \text{PhD}, \text{Postdoc} \sqsubseteq \text{PhD}, \text{PhD} \sqsubseteq \text{Person}, \exists \text{Teach} \sqsubseteq \text{Person}, \exists \text{Teach}^- \sqsubseteq \text{Course}, \text{Prof} \sqsubseteq \exists \text{WorkFor}, \text{Student} \sqsubseteq \exists \text{MemberOf}, \text{WorkFor} \sqsubseteq \exists \text{MemberOf}, \text{AProf} \sqsubseteq \neg \text{FProf}, \text{Prof} \sqsubseteq \neg \text{Postdoc}, \text{Student} \sqsubseteq \neg \text{Prof}, \text{Person} \sqsubseteq \neg \text{Course}, \exists \text{MemberOf}^- \sqsubseteq \neg \text{Postdoc} \}$

$A_a = \{ \text{AProf}(\text{ann}), \text{FProf}(\text{ann}), \text{Prof}(\text{ann}), \text{Teach}(\text{ann}, c_a), \text{Teach}(\text{ann}, \text{ann}) \}$

$q(x) = \exists yz \text{PhD}(x) \land \text{MemberOf}(x, y) \land \text{Teach}(x, z)$
Exercise: AR, IAR, brave, $k$-supp, $k$-def, ICR?

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

$\mathcal{A}_b = \{\text{AProf}(bob), \text{FProf}(bob), \text{Postdoc}(bob), \text{MemberOf}(bob, dpt), \text{Teach}(bob, c_b)\}$

$q(x) = \exists y z \text{PhD}(x) \land \text{MemberOf}(x, y) \land \text{Teach}(x, z)$
Exercise: AR, IAR, brave, $k$-supp, $k$-def, ICR?

$\mathcal{T} = \{ \text{AP} \sqsubseteq \text{Prof}, \text{FP} \sqsubseteq \text{Prof}, \text{Prof} \sqsubseteq \text{PhD}, \text{Postdoc} \sqsubseteq \text{PhD}, \text{PhD} \sqsubseteq \text{Person}, \exists \text{Teach} \sqsubseteq \text{Person}, \exists \text{Teach}^- \sqsubseteq \text{Course}, \text{Prof} \sqsubseteq \exists \text{WorkFor}, \text{Student} \sqsubseteq \exists \text{MemberOf}, \text{WorkFor} \sqsubseteq \exists \text{MemberOf}, \text{AP} \sqsubseteq \neg \text{FP}, \text{Prof} \sqsubseteq \neg \text{Postdoc}, \text{Student} \sqsubseteq \neg \text{Prof}, \text{Person} \sqsubseteq \neg \text{Course}, \exists \text{MemberOf}^- \sqsubseteq \neg \text{Postdoc} \}$

$\mathcal{A}_c = \{ \text{AP}(\text{carl}), \text{Teach}(\text{carl}, c_{c1}), \text{Teach}(\text{carl}, c_{c2}), \text{Teach}(c_{c1}, c_{c2}), \text{Teach}(c_{c2}, c_{c1}) \}$

$q(x) = \exists y z \text{PhD}(x) \land \text{MemberOf}(x, y) \land \text{Teach}(x, z)$
Exercise: AR, IAR, brave, \(k\)-supp, \(k\)-def, ICR ?

\[ \mathcal{T} = \{ \text{APref} \sqsubseteq \text{Prof}, \text{FProf} \sqsubseteq \text{Prof}, \text{Prof} \sqsubseteq \text{PhD}, \text{Postdoc} \sqsubseteq \text{PhD}, \]
\[ \text{PhD} \sqsubseteq \text{Person}, \exists \text{Teach} \sqsubseteq \text{Person}, \exists \text{Teach}^{-} \sqsubseteq \text{Course}, \]
\[ \text{Prof} \sqsubseteq \exists \text{WorkFor}, \text{Student} \sqsubseteq \exists \text{MemberOf}, \text{WorkFor} \sqsubseteq \text{MemberOf}, \]
\[ \text{APref} \sqsubseteq \neg \text{FProf}, \text{Prof} \sqsubseteq \neg \text{Postdoc}, \text{Student} \sqsubseteq \neg \text{Prof}, \]
\[ \text{Person} \sqsubseteq \neg \text{Course}, \exists \text{MemberOf}^{-} \sqsubseteq \neg \text{Postdoc} \} \]

\[ \mathcal{A}_d = \{ \text{APref}(\text{dan}), \text{Teach}(\text{dan}, c_{d1}), \text{Teach}(\text{dan}, c_{d2}), \]
\[ \text{APref}(c_{d1}), \text{APref}(c_{d2}) \} \]

\[ q(x) = \exists y z \text{PhD}(x) \land \text{MemberOf}(x, y) \land \text{Teach}(x, z) \]
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 ⊆ \( \text{LogSpace} \subseteq \text{PTime} \))
  - via query rewriting
- Based on the DL-Lite\(_{\mathcal{R}}\) language of the DL-Lite family

DL-Lite\(_{\mathcal{R}}\): 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\(_{\mathcal{R}}\):

- \( \text{DL-Lite}_{\mathcal{R}} = \text{DL-Lite}_{\mathcal{R}} + \) 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 $\text{AC0} \subseteq \text{PTIME}$ in data complexity)
Complexity of AR in DL-Lite

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

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

- $\langle \mathcal{T}, \mathcal{R} \rangle \not\models q(\bar{a})$ in AC0
- repair checking in \textbf{PTIME}?
Complexity of AR in DL-Lite

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

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

$\langle \mathcal{T}, \mathcal{R} \rangle \not|= q(\bar{a})$ in AC0

 repair checking in PTIME?

Lower bound: by reduction from propositional unsatisfiability

$\langle \mathcal{T}, \mathcal{A} \rangle$ and query $q(\bar{a})$ such that $\varphi$ is unsatisfiable iff $\langle \mathcal{T}, \mathcal{A} \rangle |=_{AR} q(\bar{a})$

$\mathcal{T} = \{ \exists P^- \sqsubseteq \neg \exists N^-, \exists P \sqsubseteq \neg \exists U^-, \exists N \sqsubseteq \neg \exists U^-, \exists U \sqsubseteq \mathcal{A} \}$

$\mathcal{A} = \{ P(c_j, x_i) \mid x_i \in C_j \} \cup \{ N(c_j, x_i) \mid \neg x_i \in C_j \} \cup \{ U(a, c_j) \mid 1 \leq j \leq m \}$

$q = A(a)$
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 $\text{PTime}$ in data complexity

Any idea of $\text{PTime}$ algorithms?

Actually, CQ entailment under IAR and brave semantics is in $\text{AC0}$ 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

\[ T = \{ \text{AP} \sqsubseteq \text{Prof}, \text{F} \sqsubseteq \text{Prof}, \text{Prof} \sqsubseteq \text{PhD}, \text{Postdoc} \sqsubseteq \text{PhD}, \text{PhD} \sqsubseteq \text{Person}, \exists \text{Teach} \sqsubseteq \text{Person}, \exists \text{Teach}^- \sqsubseteq \text{Course}, \text{Prof} \sqsubseteq \exists \text{WorkFor}, \text{Student} \sqsubseteq \exists \text{MemberOf}, \text{WorkFor} \sqsubseteq \text{MemberOf}, \text{AP} \sqsubseteq \neg \text{F}, \text{Prof} \sqsubseteq \neg \text{Postdoc}, \text{Student} \sqsubseteq \neg \text{Prof}, \text{Person} \sqsubseteq \neg \text{Course}, \exists \text{MemberOf}^- \sqsubseteq \neg \text{Postdoc} \} \]

\[ q_1(x) = \text{PhD}(x) \]

\[ q_2(x) = \exists y \text{MemberOf}(x, y) \]

\[ q_3(x) = \exists y \text{Prof}(x) \land \text{Teach}(x, y) \]
FO Rewriting for Brave Semantics

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

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

$q_1(x) = \exists y \text{PhD}(x) \land \text{MemberOf}(x, y)$

$q_2(x) = \exists y \text{Prof}(x) \land \text{Teach}(x, y)$
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

<table>
<thead>
<tr>
<th></th>
<th>Data Complexity</th>
<th>Combined Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQs</td>
<td>IQs</td>
</tr>
<tr>
<td>classical</td>
<td>in AC0</td>
<td>in AC0</td>
</tr>
<tr>
<td>AR</td>
<td>coNP-co</td>
<td>coNP-co</td>
</tr>
<tr>
<td>IAR</td>
<td>in AC0</td>
<td>in AC0</td>
</tr>
<tr>
<td>brave</td>
<td>in AC0</td>
<td>in AC0</td>
</tr>
<tr>
<td>k-support</td>
<td>in AC0</td>
<td>in AC0</td>
</tr>
<tr>
<td>k-defeater</td>
<td>in AC0</td>
<td>in AC0</td>
</tr>
<tr>
<td>ICR</td>
<td>coNP-co</td>
<td>coNP-co</td>
</tr>
<tr>
<td>CAR</td>
<td>coNP-co</td>
<td>in AC0</td>
</tr>
<tr>
<td>ICAR</td>
<td>in AC0</td>
<td>in AC0</td>
</tr>
</tbody>
</table>

**Note on AC0 cases:**

- FO-rewritings, but rewritings may be huge and not efficiently evaluated over databases
- alternative \( \text{PTime} \) 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 $\varphi$ is unsatisfiable

$$\varphi = \bigwedge_{C \in \mathcal{T}\text{-supp}} \bigvee_{\alpha \in C, \{\alpha, \beta\} \in \text{conflicts}} x_\beta \land \bigwedge_{\{\alpha, \beta\} \in \text{conflicts}} \neg x_\alpha \lor \neg x_\beta$$
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

▶ \( \langle T, A \rangle \models_{AR} q \): minimal set \( \{C_1, \ldots, C_k\} \) of minimal \( T \)-supports for \( q \) such that every repair contains at least one of the \( C_i \)

▶ \( \langle T, A \rangle \not\models_{AR} q \): minimal \( B \subseteq A \) such that \( B \) is \( T \)-consistent and for every \( T \)-support \( C \) of \( q \), \( B \cup C \) is not \( T \)-consistent

▶ IAR semantics

▶ \( \langle T, A \rangle \models_{IAR} q \): minimal \( T \)-support included in every repair

▶ \( \langle T, A \rangle \not\models_{IAR} q \): minimal \( B \subseteq A \) such that for every \( T \)-support \( C \) of \( q \), there exists \( B' \subseteq B \) such that \( B' \) is \( T \)-consistent and \( B' \cup C \) is not \( T \)-consistent

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