Example of Advanced Topics: Inconsistency Handling
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  ▶ outdated information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- outdated information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

\[ \mathcal{T} = \{ \text{AP} \sqsubseteq \text{Prof}, \text{FP} \sqsubseteq \text{Prof}, \text{AP} \sqsubseteq \neg \text{FP} \} \]
\[ \mathcal{A} = \{ \text{AP}(\text{ann}), \text{FP}(\text{ann}), \text{Postdoc(alex)} \} \]

Which assertions would 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 $\langle \mathcal{T}, \mathcal{A} \rangle \models_S q(\bar{a})$ if $\bar{a}$ is an answer to $q$ w.r.t. $\langle \mathcal{T}, \mathcal{A} \rangle$ under semantics $S$.
Consistency Properties

A (consistent) $\mathcal{T}$-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
Consistency Properties

A (consistent) $\mathcal{T}$-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 $\mathcal{I}$ of $\mathcal{T}$ such that $\langle \mathcal{T}, A \rangle \models_S q(\bar{a})$ implies $\mathcal{I} \models q(\bar{a})$

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

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

<table>
<thead>
<tr>
<th>TBox</th>
<th>ABox</th>
<th>Repair $\mathcal{R}_1$</th>
<th>Repair $\mathcal{R}_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{APprof} \sqsubseteq \text{Prof}$</td>
<td>$\text{APprof}(ann)$ $\text{FProf}(ann)$ $\text{Postdoc}(\text{alex})$</td>
<td>$\text{APprof}(ann)$ $\text{Postdoc}(\text{alex})$</td>
<td>$\text{FProf}(ann)$ $\text{Postdoc}(\text{alex})$</td>
</tr>
</tbody>
</table>
**Plausible Answers: AR Semantics**

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

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

<table>
<thead>
<tr>
<th>TBox</th>
<th>( \mathcal{T} )</th>
<th>ABox</th>
<th>( \mathcal{A} )</th>
<th>Repair ( \mathcal{R}_1 )</th>
<th>Repair ( \mathcal{R}_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>AProf \sqsubseteq Prof</td>
<td></td>
<td>AProf(\textit{ann})</td>
<td></td>
<td>AProf(\textit{ann})</td>
<td>FProf(\textit{ann})</td>
</tr>
<tr>
<td>FProf \sqsubseteq Prof</td>
<td></td>
<td>FProf(\textit{ann})</td>
<td></td>
<td>Postdoc(\textit{alex})</td>
<td>Postdoc(\textit{alex})</td>
</tr>
<tr>
<td>AProf \sqsubseteq \neg FProf</td>
<td></td>
<td>Postdoc(\textit{alex})</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Example of Advanced Topics: Inconsistency Handling**
Surest Answers: IAR Semantics

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

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

TBox $\mathcal{T}$
- AProf $\sqsubseteq$ Prof
- FProf $\sqsubseteq$ Prof
- AProf $\sqsubseteq \neg$FProf

ABox $\mathcal{A}$
- AProf($ann$)
- FProf($ann$)
- Postdoc($alex$)

Repair $\mathcal{R}_1$
- AProf($ann$)
- Postdoc($alex$)

Repair $\mathcal{R}_2$
- FProf($ann$)
- Postdoc($alex$)

Consequences($\mathcal{R}_1$)
Consequences($\mathcal{R}^\cap$)
Consequences($\mathcal{R}_2$)

Example of Advanced Topics: Inconsistency Handling
Possible Answers: Brave Semantics

Brave answers: hold in some repair

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

- **TBox**
  - AProf ⊆ Prof
  - FProf ⊆ Prof
  - AProf ⊆ ¬FProf

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

- **Repair**
  - R₁
    - AProf(ann)
    - Postdoc(alex)
  - R₂
    - FProf(ann)
    - Postdoc(alex)

---

Example of Advanced Topics: Inconsistency Handling
AR, IAR and Brave Semantics

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

How do the three semantics compare?
AR, IAR and Brave Semantics

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

- AR : consistent support and consistent results property
- IAR : consistent support and consistent results property
- brave : consistent support only

How do the three semantics compare?

\[ \langle T, A \rangle \models_{IAR} q \iff \langle T, A \rangle \models_{AR} q \iff \langle T, A \rangle \models_{brave} q \]
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
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-sup} q(\bar{a})$ 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
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-sup} q(\bar{a}) \Rightarrow \langle \mathcal{T}, \mathcal{A} \rangle \models_{k+1-sup} q(\bar{a})$
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q(\bar{a}) \iff \exists k \geq 1, \langle \mathcal{T}, \mathcal{A} \rangle \models_{k-sup} q(\bar{a})$

- **$k$-defeater semantics**
  - fine-grained over-approximation of AR
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k-def} q(\bar{a})$ 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 $q(\bar{a})$
  - $0$-defeater = brave
  - $\langle \mathcal{T}, \mathcal{A} \rangle \models_{k+1-def} q(\bar{a}) \Rightarrow \langle \mathcal{T}, \mathcal{A} \rangle \models_{k-def} q(\bar{a})$
  - for every KB, there exists $k$ such that $\langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q(\bar{a}) \iff \langle \mathcal{T}, \mathcal{A} \rangle \models_{k-def} q(\bar{a})$
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 (closure of $\mathcal{R} = \text{set of assertions entailed from } \langle \mathcal{T}, \mathcal{R} \rangle$)
  - same as AR for queries without quantifier

**Example of Advanced Topics: Inconsistency Handling**
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 $R = \text{set of assertions entailed from } \langle T, R \rangle$)
  - 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 $T$-support in $A$
  - closed ABox repairs: maximally complete standard ABox repairs with facts from the closure of $A$
  - apply AR (CAR) or IAR (ICAR) using closed ABox repairs
  - do not satisfy consistent support!

$$T = \{ A \sqsubseteq B, C \sqsubseteq D, A \sqsubseteq \neg C \}, \quad A = \{ A(a), C(a) \},$$

$$q = B(x) \land D(x)$$

Example of Advanced Topics: Inconsistency Handling
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
Some Complexity Results for DL-Lite

Results apply to all languages that satisfy
- minimal $\mathcal{T}$-supports for $q(\bar{a})$ contain at most $|q|$ assertions
- minimal $\mathcal{T}$-inconsistent subsets have cardinality at most 2
- CQ answering and satisfiability can be performed by FO rewriting (so in AC0 in data complexity)
Some Complexity Results for 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 $\langle \mathcal{T}, \mathcal{R} \rangle \not\models q(\bar{a})$

- $\langle \mathcal{T}, \mathcal{R} \rangle \not\models q(\bar{a})$ in AC0
- repair checking in PTIME: check $\mathcal{T}$-consistency of $\mathcal{R} +$ check $\mathcal{T}$-inconsistency of every $\mathcal{R} \cup \{\alpha\}$ for $\alpha \in \mathcal{A} \setminus \mathcal{R}$

Example of Advanced Topics: Inconsistency Handling
Some Complexity Results for 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
\[ \langle \mathcal{T}, \mathcal{R} \rangle \nRightarrow q(\bar{a}) \]

- \[ \langle \mathcal{T}, \mathcal{R} \rangle \nRightarrow q(\bar{a}) \] in **AC0**
- **repair checking in PTIME**: check \( \mathcal{T} \)-consistency of \( \mathcal{R} \) + check \( \mathcal{T} \)-inconsistency of every \( \mathcal{R} \cup \{\alpha\} \) for \( \alpha \in \mathcal{A} \setminus \mathcal{R} \)

Lower bound: reduction from **UNSAT**: \( \varphi = c_1 \land \cdots \land c_m \)
conjunction of clauses over variables \( x_1, \ldots, x_k \), build a KB \( \langle \mathcal{T}, \mathcal{A} \rangle \)
and query \( q(\bar{a}) \) such that \( \varphi \) is unsatisfiable iff \( \langle \mathcal{T}, \mathcal{A} \rangle \models_{AR} q(\bar{a}) \)
Some Complexity Results for DL-Lite

CQ entailment under IAR and brave semantics is in \text{PTime} in data complexity.

→ Compute the conflicts of the KB and the sets of assertions that match of the query (bounded size), then check whether there is one such that none of its assertions is involved in a conflict / which is $\mathcal{T}$-consistent.
Some Complexity Results for DL-Lite

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

→ Compute the conflicts of the KB and the sets of assertions that match of the query (bounded size), then check whether there is one such that none of its assertions is involved in a conflict / which is $\mathcal{T}$-consistent

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

→ Can use FO-rewriting to compute IAR and brave answers
   Idea: modify UCQ-rewriting to ensure ABox assertions matching CQs are not involved in any contradictions / form a $\mathcal{T}$-consistent subset of the ABox

Example of Advanced Topics: Inconsistency Handling
## 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 IQs</td>
<td>CQs IQs</td>
</tr>
<tr>
<td>classical</td>
<td>in AC0 in AC0</td>
<td>NP-co in PTIME</td>
</tr>
<tr>
<td>AR</td>
<td>coNP-co coNP-co</td>
<td>( \Pi_2^P )-co coNP-co</td>
</tr>
<tr>
<td>IAR</td>
<td>in AC0 in AC0</td>
<td>NP-co in PTIME</td>
</tr>
<tr>
<td>brave</td>
<td>in AC0 in AC0</td>
<td>NP-co in PTIME</td>
</tr>
<tr>
<td>( k )-support</td>
<td>in AC0 in AC0</td>
<td>NP-co in PTIME</td>
</tr>
<tr>
<td>( k )-defeater</td>
<td>in AC0 in AC0</td>
<td>NP-co in PTIME</td>
</tr>
<tr>
<td>ICR</td>
<td>coNP-co coNP-co</td>
<td>( \Delta_2^P[O\log(n)] )-co coNP-co</td>
</tr>
<tr>
<td>CAR</td>
<td>coNP-co in AC0</td>
<td>( \Pi_2^P )-co in PTIME</td>
</tr>
<tr>
<td>ICAR</td>
<td>in AC0 in AC0</td>
<td>NP-co in PTIME</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

---

*Example of Advanced Topics: Inconsistency Handling*
A Practical Approach to Query Answering under AR Semantics

- Exploit tractable approximations:
  IAR $\Rightarrow$ AR and not brave $\Rightarrow$ not AR

- for remaining cases (brave and not IAR): reduce AR entailment to SAT and use a SAT solver
Research Problems on Inconsistency Handling

- Inconsistency-tolerant semantics
- Practical algorithms, implementations
- Beyond DL-Lite: unbounded size of query supports and conflicts
- Explanations of query results
- Improving data quality, helping user to resolve inconsistencies