Can you tell the difference between \textit{DL-Lite} ontologies?

Roman Kontchakov

\textit{SCSIS, Birkbeck College, London}

http://www.dcs.bbk.ac.uk/~roman

Frank Wolter and Michael Zakharyaschev

University of Liverpool \hspace{1cm} Birkbeck College, London
Developing and maintaining ontologies

● versions:
  comparing **logical consequences** over some common vocabulary \( \Sigma \)
  not not the syntactic form of the axioms (as in diff)

● refinement:
  adding new axioms but **preserving** the relationships
  between terms of a certain part \( \Sigma \) of the vocabulary

● reuse:
  importing an ontology and using its vocabulary \( \Sigma \) as originally defined
  (relationships between terms of \( \Sigma \) should not change)

new types of reasoning problems

related notions: conservative extensions, model conservativity, locality, etc.
**Σ-difference**

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be TBoxes (in some DL $\mathcal{L}$) and $\Sigma$ a signature (concept and role names).  

**Σ-concept difference** $c\text{Diff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ is the set of $\Sigma$-concept inclusions such that $\mathcal{T}_2 \models C \subseteq D$ and $\mathcal{T}_1 \not\models C \subseteq D$.

**Σ-query difference** $q\text{Diff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ is the set of pairs $(\mathcal{A}, q(\overline{x}))$, where $\text{sig}(\mathcal{A}), \text{sig}(q) \subseteq \Sigma$, $(\mathcal{T}_2, \mathcal{A}) \models q(\bar{a})$ and $(\mathcal{T}_1, \mathcal{A}) \not\models q(\bar{a})$, for some $\bar{a}$.

**Strong Σ-query difference** $\text{sqDiff}_\Sigma(\mathcal{T}_1, \mathcal{T}_2)$ is the set of triples $(\mathcal{T}, \mathcal{A}, q(\overline{x}))$, where $\text{sig}(\mathcal{T}, \mathcal{A}), \text{sig}(q) \subseteq \Sigma$, $(\mathcal{T}_2 \cup \mathcal{T}, \mathcal{A}) \models q(\bar{a})$, $(\mathcal{T}_1 \cup \mathcal{T}, \mathcal{A}) \not\models q(\bar{a})$, for some $\bar{a}$.

$\mathcal{T}_1$ and $\mathcal{T}_2$ are **Σ-concept inseparable** iff $c\text{Diff}(\mathcal{T}_1, \mathcal{T}_2) = \emptyset$ and $c\text{Diff}(\mathcal{T}_2, \mathcal{T}_1) = \emptyset$.

$\mathcal{T}_1$ and $\mathcal{T}_2$ are **Σ-query inseparable** iff $q\text{Diff}(\mathcal{T}_1, \mathcal{T}_2) = \emptyset$ and $q\text{Diff}(\mathcal{T}_2, \mathcal{T}_1) = \emptyset$.

---

- ExpTime for $\mathcal{EL}$, 2ExpTime for $\mathcal{ALCQI}$, undecidable for $\mathcal{ALCQIO}$
- tractable for acyclic $\mathcal{EL}$ (e.g., SNOMED)
**DL-Lite: Description Logic for Databases**

A fragment of a conceptual schema:

**Translating into DL:**

1. \( \exists \text{manages}. \top \sqsubseteq \text{ProjectManager} \)
2. \( \exists \text{manages}^{-}. \top \sqsubseteq \text{Project} \)
3. \( \text{Project} \sqsubseteq \exists \text{manages}^{-}. \top \)
4. \( \geq 3 \text{ manages}^{-}. \top \sqsubseteq \bot \)
5. \( \text{Research} \sqcap \text{Visiting} \sqsubseteq \bot \)
6. \( \text{Academic} \sqsubseteq \text{ProjectManager} \)
7. \( \text{ProjectManager} \sqsubseteq \text{Academic} \sqcup \text{Visiting} \)

\[
\begin{align*}
DL-Lite_{\text{horn}} & : B_1 \sqcap \cdots \sqcap B_k \sqsubseteq B \\
DL-Lite_{\text{bool}} & : C_1 \sqsubseteq C_2 
\end{align*}
\]

\[
\begin{align*}
B & ::= \bot \mid A_i \mid \exists R \mid \geq q R \\
C & ::= B \mid \neg C \mid C_1 \sqcap C_2 \mid C_1 \sqcup C_2 
\end{align*}
\]
Example

Let $T_1$ contain the axioms

- Research $\sqsubseteq \exists \text{worksIn}$,
- Project $\sqsubseteq \exists \text{manages}$,
- $\exists \text{teaches} \sqsubseteq \text{Academic} \sqcup \text{Research}$,
- Research $\cap$ Visiting $\sqsubseteq \bot$,
- $\exists \text{writes} \sqsubseteq \text{Academic} \sqcup \text{Research}$,
- $\exists \text{worksIn}^- \sqsubseteq \text{Project}$,
- $\exists \text{manages} \sqsubseteq \text{Academic} \sqcup \text{Visiting}$,
- Academic $\sqsubseteq \exists \text{teaches} \cap \leq 1 \text{teaches}$,
- $\exists \text{writes} \sqsubseteq \text{Academic} \sqcup \text{Research}$,

$T_2 = T_1 \cup \{ \text{Visiting} \sqsubseteq \geq 2 \text{writes} \}$ and $\Sigma = \{ \text{teaches} \}$

- $T_1$ and $T_2$ are $\Sigma$-concept inseparable (\Sigma-entailment in both directions)
  - $T_2 \models \text{Visiting} \sqsubseteq \text{Academic}$, but nothing new in the signature $\Sigma$

- $T_1$ does not $\Sigma$-query entail $T_2$:
  - $\mathcal{A} = \{ \text{teaches}(a,b), \text{teaches}(a,c) \}$
  - $q = \exists x ((\exists \text{teaches})(x) \land (\leq 1 \text{teaches})(x))$
  - ‘is there anybody who teaches precisely one module?’
  - $(T_1, \mathcal{A}) \not\models q$
  - $(T_2, \mathcal{A}) \models q$
**Σ-inseparability in DL-Lite**

**Theorem**

(1) In $DL$-$Lite_{bool}$:
Strong $Σ$-query insep. $\iff$ $Σ$-query inseparability $\Rightarrow$ $Σ$-concept inseparability

In each case the problem is $\Pi^P_2$-complete

(2) In $DL$-$Lite_{horn}$:
Strong $Σ$-query insep. $\Rightarrow$ $Σ$-query inseparability $\Rightarrow$ $Σ$-concept inseparability

In each case the problem is coNP-complete

(3) In $DL$-$Lite_{bool}$:
$Σ$-query entailment and $Σ$-concept entailment

can be encoded by Quantified Boolean Formulas $\forall \exists \psi$
**Σ-entailment: semantic criteria**

Let $Q$ be a set of numerical parameters and $Σ$ a signature

**ΣQ-concepts $B$:** $A_i ∈ Σ$ and $(≥ q R)$ with $q ∈ Q$ and $R ∈ Σ$

<table>
<thead>
<tr>
<th>$ΣQ$-type $t$ is a set of $ΣQ$-concepts containing</th>
</tr>
</thead>
<tbody>
<tr>
<td>● $B$ or $¬B$ (but not both), for all $B$</td>
</tr>
<tr>
<td>● $≥ q R$ whenever $q &lt; q'$ and $≥ q' R ∈ t$, for all $≥ q R$</td>
</tr>
</tbody>
</table>

For a TBox $T$,

a $ΣQ$-type $t$ is $T$-realisable if $t$ is satisfied in a model of $T$

a set $Ξ$ of $ΣQ$-types is precisely $T$-realisable if

there is a model of $T$ realising precisely the types from $Ξ$

**Theorem.** Let $Q$ denote the set of parameters occurring in $T_1 ∪ T_2$

$T_1$ $Σ$-concept entails $T_2$ iff every $T_1$-realisable $ΣQ$-type is $T_2$-realisable

$T_1$ $Σ$-query entails $T_2$ iff every precisely $T_1$-realisable set $Ξ$ of $ΣQ$-types is precisely $T_2$-realisable
Encoding $\Sigma$-concept entailment in QBF

Let $\mathcal{T}$ be a TBox, $Q$ a set of numerical parameters and $t$ a $\text{sig}(\mathcal{T})Q$-type

\[ \text{‘}$t_0$ is $\mathcal{T}$-realisable with $t_1, \ldots, t_n$ being witnesses’\text{‘} = \Phi_{\mathcal{T}}(b_0, b_1, \ldots, b_n) \]

Propositional formula

$b_j$ is the vector of all propositional variables $B^*$ of the type $t_j$

Then the condition

\text{‘}every $\mathcal{T}_1$-realisable $\Sigma Q$-type $t$ is $\mathcal{T}_2$-realisable’ \text{‘}

is described by the following QBF

\[
\forall b_0^{\Sigma Q} \left[ \exists b_0^{T_2 \setminus \Sigma Q} \exists b_1^{T_1} \ldots \exists b_{n_1}^{T_1} \Phi_{\mathcal{T}_1}(b_0^{\Sigma Q} \cdot b_0^{T_1 \setminus \Sigma Q}, b_1^{T_1}, \ldots, b_{n_1}^{T_1}) \rightarrow \\
\exists b_0^{T_2 \setminus \Sigma Q} \exists b_1^{T_2} \ldots \exists b_{n_2}^{T_2} \Phi_{\mathcal{T}_2}(b_0^{\Sigma Q} \cdot b_0^{T_2 \setminus \Sigma Q}, b_1^{T_2}, \ldots, b_{n_2}^{T_2}) \right]
\]

$(b_0^{\Sigma Q}$ is the $\Sigma Q$-part of $b_0$ and $b_0^{T_i \setminus \Sigma Q}$ contains the rest of the variables)
### Experiments

#### TBox instances
(standard Department Ontology + ICNARC)

<table>
<thead>
<tr>
<th>series</th>
<th>description</th>
<th>no. of instances</th>
<th>axioms</th>
<th>basic concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>( T_1 )</td>
<td>( T_2 )</td>
</tr>
<tr>
<td>NN</td>
<td>( T_1 ) does not ( \Sigma )-concept entail ( T_2 )</td>
<td>420</td>
<td>59–154</td>
<td>74–198</td>
</tr>
<tr>
<td>YN</td>
<td>( T_1 ) ( \Sigma )-concept but not ( \Sigma )-query entails ( T_2 )</td>
<td>252</td>
<td>56–151</td>
<td>77–191</td>
</tr>
<tr>
<td>YY</td>
<td>( T_1 ) ( \Sigma )-query entails ( T_2 )</td>
<td>156</td>
<td>54–88</td>
<td>62–110</td>
</tr>
</tbody>
</table>

#### QBF solvers
- sKizzo 0.8.2 ([http://skizzo.info/](http://skizzo.info/)).
- yQuaffle ([http://www.princeton.edu/~chaff/quaffle.html](http://www.princeton.edu/~chaff/quaffle.html)).
- QuBE 6.4 ([http://www.star.dist.unige.it/](http://www.star.dist.unige.it/)).

### QBF entailment

<table>
<thead>
<tr>
<th>series</th>
<th>( \Sigma )-concept entailment QBF</th>
<th>( \Sigma )-query entailment QBF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>variables</td>
<td>clauses</td>
</tr>
<tr>
<td>NN</td>
<td>1,469–11,752</td>
<td>2,391–18,277</td>
</tr>
<tr>
<td>YN</td>
<td>1,460–11,318</td>
<td>2,352–17,424</td>
</tr>
<tr>
<td>YY</td>
<td>1,526–4,146</td>
<td>2,200–6,079</td>
</tr>
</tbody>
</table>
Experimental results: percentage of solved instances

%  %  %
\[ \Sigma \text{-concept entailment} \]

%  %  %
\[ \Sigma \text{-query entailment} \]

NN  YN  YY

2clsQ  sKizzo  yQuaffle  QuBE

KR 2008  Sydney  18.09.08
Forgetting

studied under different names: forgetting, uniform interpolation, variable elimination...

A DL $\mathcal{L}$ admits forgetting (has uniform interpolation) if,
for every $\mathcal{T}$ in $\mathcal{L}$ and every $\Sigma$, there exists $\mathcal{T}_\Sigma$ in $\mathcal{L}$ with $\text{sig}(\mathcal{T}_\Sigma) \subseteq \Sigma$
such that $\mathcal{T}$ and $\mathcal{T}_\Sigma$ are $\Sigma$-concept inseparable in $\mathcal{L}$

Theorem Both $\text{DL-Lite}^{u}_{\text{bool}}$ and $\text{DL-Lite}^{u}_{\text{horn}}$ have uniform interpolation
and the uniform interpolant can be constructed in exponential time

$\text{DL-Lite}^{u}_{\text{bool}}$:

$$C ::= \ldots \mid \exists C \mid \ldots$$

(universal modality)

e.g., $(\geq 2 \text{ teaches}) \sqsubseteq \exists (\exists \text{ teaches} \cap \leq 1 \text{ teaches})$

$\mathcal{T}_\Sigma$ with $\text{sig}(\mathcal{T}_\Sigma) \subseteq \Sigma$ is a uniform interpolant of $\mathcal{T}$ w.r.t. $\Sigma$ in $\text{DL-Lite}^{u}_{\text{bool}}$ if
$\mathcal{T} \models C \sqsubseteq D$ iff $\mathcal{T}_\Sigma \models C \sqsubseteq D$, for every $C \sqsubseteq D$ in $\text{DL-Lite}^{u}_{\text{bool}}$ with $\text{sig}(C \sqsubseteq D) \subseteq \Sigma$

$\mathcal{T}' \Sigma$-query entails $\mathcal{T}$ iff $\mathcal{T}' \models C \sqsubseteq D$, for each $C \sqsubseteq D \in \mathcal{T}_\Sigma$

Theorem For every $\mathcal{T}$ in $\text{DL-Lite}^{u}_{\text{bool}}$ and every $\Sigma$ one can construct
a uniform interpolant $\mathcal{T}_\Sigma$ of $\mathcal{T}$ w.r.t. $\Sigma$ in $\text{DL-Lite}^{u}_{\text{bool}}$ in time exponential in $\mathcal{T}$
Future work

- investigate different variants of the QBF encoding (non-prenex/non-CNF) and/or different solvers (AQME or even a dedicated solver)
- QBF encoding of $\Sigma$-entailment in $DL-Lite_{horn}$ (coNP instead of $\Pi^p_2$)
- module extraction algorithm (extended QBF encoding)
- approximation of $\Sigma$-difference