Complexity of Reasoning with Expressive Ontology Mappings

Chiara Ghidini, Luciano Serafini and Sergio Tessaris

FBK-irst, Trento, Italy
Free University of Bolzen-Bolzano, Bolzano, Italy

FOIS’08 - 1-3 November 2008
Mapping Heterogeneous Ontologies

- Mapping languages focus mainly on mappings between concepts from different ontologies; Very few address mappings between roles;
  - $s$ : Article less general than $t$ : Publication
  - $s$ : partnerOf more general than $t$ : marriedTo

- Mismatches due to schematic differences exist in different ontologies; A typical example is the representation of an element as a concept in one ontology and as a role in another ontology.
Mapping Heterogeneous Ontologies

- Mapping languages focus mainly on mappings between concepts from different ontologies; Very few address mappings between roles;
  - \( s \) : Article less general than \( t \) : Publication
  - \( s \) : \textit{partnerOf} more general than \( t \) : \textit{marriedTo}

- Mismatches due to \textbf{schematic differences} exist in different ontologies; A typical example is the representation of an element as a \textbf{concept} in one ontology and as a \textbf{role} in another ontology.
Examples from Upper Level Ontologies - I

- **Spatial location**

<table>
<thead>
<tr>
<th>Entity</th>
<th>Identifier</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>GUM:SpatialLocating</td>
<td>Any configuration whose function it is to locate some physical object in space. Instances must have one locatum and one placement.</td>
</tr>
<tr>
<td>Relation</td>
<td>DOLCE-Lite:<em>physical-location</em></td>
<td>Analytical location holding between physical endurants and physical regions.</td>
</tr>
</tbody>
</table>
Examples from Upper Level Ontologies - II

- **Cause**

<table>
<thead>
<tr>
<th>Entity</th>
<th>Identifier</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>GUM:Causal</td>
<td>This concept defines a Configuration, in which one of the participants is the cause of the other. Hence, one of the participants occupies the role &quot;cause&quot;, and the other one &quot;effect&quot;.</td>
</tr>
<tr>
<td>Relation</td>
<td>GIST:CausedByDirect</td>
<td><em>No comment present, but from reading the ontology this is the relation holding between two objects in which one is the direct cause of the other.</em></td>
</tr>
</tbody>
</table>
## Examples from Upper Level Ontologies - III

### Membership

<table>
<thead>
<tr>
<th>Entity</th>
<th>Identifier</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>GIST:Membership</td>
<td>Declared membership in some group or collection (such as the ACM).</td>
</tr>
<tr>
<td>Relation</td>
<td>DOLCE-Lite: <em>TemporaryProperPart</em></td>
<td>Being proper part at time t. It holds for endurants only. This is important to model proper parts that can change or be lost over time without affecting the identity of the whole.</td>
</tr>
</tbody>
</table>
Our contribution

- Extend the mapping language of DDL to represent mappings between concepts and roles from different ontologies;
- Study the effects and computational complexity of the proposed mappings;
- Ongoing work: capture reification modelling pattern.
Our contribution

- Extend the mapping language of DDL to represent mappings between concepts and roles from different ontologies;
- Study the effects and computational complexity of the proposed mappings;
- Ongoing work: capture reification modelling pattern.
DDL formalism - Syntax

- Two (or more) distinct TBoxes
  - Using $\mathcal{ALCQI}_b$ for technical reasons

- Bridge rules ($\mathcal{B}$) connecting terms:
  - directional (source and target ontology)
  - connect source to target without back-flow
  - two flavours: into and onto rules

  $s : \text{Article} \rightarrow t : \text{Publication}$

  $s : \text{partnerOf} \rightarrow t : \text{marriedTo}$

- Distributed T-box (DTB): $\mathcal{I} = \langle \{T_i\}_{i \in I}, \mathcal{B} \rangle$. 
DDL formalism - Syntax

- **Two (or more) distinct TBoxes**
  - Using $\mathcal{ALCQI}_b$ for technical reasons

- **Bridge rules ($\mathcal{B}$) connecting terms:**
  - directional (source and target ontology)
  - connect source to target **without** back-flow
  - two flavours: **into** and **onto** rules

\[
s : \text{Article} \xrightarrow{\mathcal{B}} t : \text{Publication}
\]

\[
s : \text{partnerOf} \xrightarrow{\mathcal{B}} t : \text{marriedTo}
\]

- **Distributed T-box (DTB):** $\mathcal{T} = \langle \{T_i\}_{i \in I}, \mathcal{B} \rangle$. 
DDL formalism - Syntax

- Two (or more) distinct TBoxes
  - Using $\mathcal{ALCQI}_b$ for technical reasons
- Bridge rules ($\mathcal{B}$) connecting terms:
  - directional (source and target ontology)
  - connect source to target without back-flow
  - two flavours: into and onto rules

$$s : \text{Article} \xrightarrow{\text{into}} t : \text{Publication}$$
$$s : \text{partnerOf} \xrightarrow{\text{onto}} t : \text{marriedTo}$$

- Distributed T-box (DTB): $\mathcal{T} = \langle \{T_i\}_{i \in I}, \mathcal{B} \rangle$. 
Homogeneous and Heterogeneous Bridge Rules

<table>
<thead>
<tr>
<th>concept</th>
<th>role</th>
</tr>
</thead>
<tbody>
<tr>
<td>concept</td>
<td>into/onto to into/onto</td>
</tr>
<tr>
<td>role</td>
<td>into/onto to into/onto</td>
</tr>
</tbody>
</table>

- $s$:Article $\sqsubseteq t$:Publication
- $s$:partnerOf $\sqsupseteq t$:marriedTo
- $s$:Marriage $\sqsubseteq t$:partnerOf
- $s$:Marriage $\sqsupseteq t$:marriedTo
local interpretations for TBoxes $\mathcal{I}_i = \langle \Delta_i, .\mathcal{I}_i \rangle$

- bridge rules interpreted by means of domain relations $r_{st} \subseteq \Delta_s \times \Delta_t$

- Distributed interpretation $\mathcal{I}$: $\langle \{\mathcal{I}_i\}_{i \in I}, \{r_{ij}\}_{i \neq j \in I} \rangle$. 
DDL formalism - Semantics (Homogeneous Case)

- local interpretations for TBoxes $\mathcal{I}_i = \langle \Delta_i, \mathcal{I}_i \rangle$
- bridge rules interpreted by means of domain relations $r_{st} \subseteq \Delta_s \times \Delta_t$
- Distributed interpretation $\mathcal{I}$: $\langle \{\mathcal{I}_i \}_{i \in I}, \{r_{ij} \}_{i \neq j \in I} \rangle$. 
local interpretations for TBoxes $\mathcal{I}_i = \langle \Delta_i, \cdot \mathcal{I}_i \rangle$

- bridge rules interpreted by means of domain relations $r_{st} \subseteq \Delta_s \times \Delta_t$

- Distributed interpretation $\mathcal{I}_d$: $\langle \{\mathcal{I}_i\}_{i \in I}, \{r_{ij}\}_{i \neq j \in I} \rangle$. 

Chiara Ghidini (FBK-irst, FUB)  
Reasoning with Expressive Ontology Mapping
Homogeneous Bridge rules: Satisfiability

\[ s: \text{Article} \rightarrow t: \text{Publication} \]

\[ \mathcal{I} \models s: \text{Article} \rightarrow t: \text{Publication}, \text{iff} \ r_{st}(\text{Article}^{\mathcal{I}_s}) \subseteq \text{Publication}^{\mathcal{I}_t} \]

\[ r_{st}(\text{Article}^{\mathcal{I}_s}) = \{ x \mid \exists a \in \text{Article}^{\mathcal{I}_s}, x \in r_{st}(a) \} \]
Homogeneous Bridge rules: Satisfiability

\[ \mathcal{I} \models s: \text{Article} \rightarrow t: \text{Publication}, \text{ iff } r_{st}(\text{Article}^\mathcal{I}_s) \supseteq \text{Publication}^\mathcal{I}_t \]

\[ r_{st}(\text{Article}^\mathcal{I}_s) = \{ x \mid \exists a \in \text{Article}^\mathcal{I}_s, x \in r_{st}(a) \} \]
Homogeneous Bridge rules: Satisfiability

\[ \mathcal{I} \models s : \text{partnerOf} \rightarrow t : \text{marriedTo}, \text{iff } r_{st}(\text{partnerOf}^\mathcal{I}_s) \supseteq \text{marriedTo}^\mathcal{I}_t \]

\[ r_{st}(\text{partnerOf}^\mathcal{I}_s) = \{ (x, y) \mid \exists (a, b) \in \text{partnerOf}^\mathcal{I}_s, x \in r_{st}(a), y \in r_{st}(b) \} \]
Heterogeneous Bridge Rules

- Rules mapping concepts to roles (or the other way around)
  
  \[ s : \text{Marriage} \xrightarrow{\subseteq} t : \text{partnerOf} \]

- both into and onto versions
- Semantics provided by means of two additional domain relations
  - concept-role \( cr_{st} \)
  - role-concept \( rc_{st} \)

- Desiderata
  - “connect” objects to pairs
  - “distinguish” among pairs participating in different relations
    - e.g. \( \text{marriedTo/dancesWith} \)
Heterogeneous Bridge Rules

- Rules mapping concepts to roles (or the other way around)

\[ s : \text{Marriage} \rightarrow t : \text{partnerOf} \]

- both into and onto versions
- Semantics provided by means of two additional domain relations
  - concept-role \( cr_{st} \)
  - role-concept \( rc_{st} \)
- Desiderata
  - “connect” objects to pairs
  - “distinguish” among pairs participating in different relations
    - e.g. \( \text{marriedTo}/\text{dancesWith} \)
Admissible Triples

Definition

The set of admissible triple $[\mathcal{R}]^I$ is $\bigcup_{R \in \mathcal{R}} [R]^I$, where

$$[R]^I = \{ R[d_1, d_2] \mid (d_1, d_2) \in R^I \}$$

- Heterogeneous domain relations:

  $$cr_{st} \subseteq \Delta^I_s \times [\mathcal{R}]^I_t$$
  $$rc_{st} \subseteq [\mathcal{R}]^I_s \times \Delta^I_t$$

- Satisfiability of the heterogeneous rule $s$:Marriage $\models t$:partnerOf

  $$cr_{st}(Marriage^I_s) \subseteq [\text{partnerOf}]^I_t$$
Heterogeneous Bridge Rules: Satisfiability

\[ s : \text{Marriage} \equiv t : \text{partnerOf} \]

\[ \mathcal{I}_s \]

\[ \mathcal{I}_t \]

\[ \Delta_s \]

\[ [R] \]

\[ \mathcal{I} \models s : \text{Marriage} \iff t : \text{partnerOf}, \text{ iff } cr_{st}(\text{Marriage}^{\mathcal{I}_s}) \subseteq [\text{partnerOf}]^{\mathcal{I}_t} \]
Effects of Bridge Rules

- Propagation rules characterise the knowledge propagation
  \[
  \begin{align*}
  & \text{axioms in } s \\
  & \text{bridge rules from } s \text{ to } t \\
  \hline
  & \text{axiom in } t
  \end{align*}
  \]

- Correct and complete set of propagation rules
  - define an operator which computes all and only the new knowledge
Effects of Bridge Rules

- Propagation rules characterise the knowledge propagation
  
  \[
  \text{axioms in } s \\
  \text{bridge rules from } s \text{ to } t \\
  \text{axiom in } t
  \]

- Correct and complete set of propagation rules
  - define an operator which computes all and only the new knowledge
Propagation of hierarchies

**Propagation Rule**

\[
\begin{align*}
    s : X_1 & \sqsubseteq X_2 \\
    s : X_1 & \rightarrow t : Y_1 \\
    s : X_2 & \rightarrow t : Y_2 \\
    t : Y_1 & \sqsubseteq Y_2
\end{align*}
\]

- \(X_i \ (Y_j)\) both concepts or both atomic roles;

**Graphically**
Propagation of hierarchies

Propagation Rule

\[
\begin{align*}
& s : X_1 \sqsubseteq X_2 \\
& s : X_1 \xrightarrow{\exists} t : Y_1 \\
& s : X_2 \sqsubseteq t : Y_2 \\
\hline
& t : Y_1 \sqsubseteq Y_2
\end{align*}
\]

Graphically

- \( X_i \) (\( Y_j \)) both concepts or both atomic roles;
Usage of Propagation rules

- Alignment of Upper Level Ontologies: GUM and DOLCE-Light
- Proposed mappings:
  \[ \text{GUM: SpatialLocating} \equiv \rightarrow \text{DOLCE-Lite: physical-location} \]
  \[ \text{GUM: SpatialTemporalLocating} \equiv \rightarrow \text{DOLCE-Lite: spatio-temporally-present-at} \]

- GUM: SpatialTemporalLocating \subseteq SpatialLocating

- Question: do we really want to infer new facts in DOLCE-Light or revise the mappings?
- Example:
  \[ \text{GUM: SpatialTemporalLocating} \subseteq \rightarrow \text{DOLCE-Lite: spatio-temporally-present-at} \]
Usage of Propagation rules

- Alignment of Upper Level Ontologies: GUM and DOLCE-Light
- Proposed mappings:
  \[ \text{GUM: SpatialLocating} \equiv \rightarrow \text{DOLCE-Lite: physical-location} \]
  \[ \text{GUM: SpatialTemporalLocating} \equiv \rightarrow \text{DOLCE-Lite: spatio-temporally-present-at} \]
  \[ \text{GUM: SpatialTemporalLocating} \subseteq \text{SpatialLocating} \]

- Question: do we really want to infer new facts in DOLCE-Light or revise the mappings?
- Example:
  \[ \text{GUM: SpatialTemporalLocating} \subseteq \rightarrow \text{DOLCE-Lite: spatio-temporally-present-at} \]
Propagation of role domain (range)

Propagation Rule

\[ \exists P. \top \sqsubseteq B \]

\[ s : P \quad \overset{\sqsubseteq}{\rightarrow} \quad t : R \]

\[ s : B \quad \overset{\sqsubseteq}{\rightarrow} \quad t : D \]

\[ t : \exists R. \top \sqsubseteq D \]

Graphically

\[ \begin{array}{ccc}
B & \overset{\sqsubseteq}{\rightarrow} & D \\
\downarrow & & \downarrow \\
P & \overset{\sqsubseteq}{\rightarrow} & R
\end{array} \]
Propagation of role domain (range)

**Propagation Rule**

\[
\begin{align*}
s : \exists P. \top \sqsubseteq B \\
s : P & \quad \models \quad t : R \\
s : B & \quad \sqsubseteq \quad t : D \\
t : \exists R. \top & \sqsubseteq D
\end{align*}
\]

**Graphically**

\[
\begin{align*}
\text{B} & \quad \sqsubseteq \quad \text{D} \\
\text{P} & \quad \models \quad \text{R}
\end{align*}
\]
The general rules

\[
\begin{align*}
  i : A &\subseteq \bigsqcup_{k=1}^{n} B_k \\
  i : A &\rightarrow j : C \\
  i : B_k &\subseteq j : D_k, \text{ for } 1 \leq k \leq n \\
  j : C &\subseteq \bigsqcup_{k=1}^{n} D_k \\
  i : P &\subseteq Q \\
  i : P &\rightarrow j : C \\
  i : Q &\subseteq j : D \\
  j : C &\subseteq D \\
  j : R &\subseteq \bigsqcup_{k=1}^{m} S_k \\
  i : A &\subseteq \bigsqcup_{k=1}^{n} B_k \\
  i : A &\rightarrow j : R \\
  i : B_k &\subseteq j : S_k, \text{ for } 1 \leq k \leq n \\
  j : S_k &\subseteq \bigsqcup_{k=1}^{m} S_k \\
  i : \exists (P \sqcap \neg(\bigsqcup_{h=1}^{l} Q_h)) \cdot (\neg \bigsqcup_{h=1}^{p} A_h) \subseteq (\bigsqcup_{h=1}^{m} B_h) \\
  i : P &\rightarrow j : R \\
  i : Q_h &\subseteq j : S_h, \text{ for } 1 \leq h \leq l \\
  i : A_h &\subseteq j : C_h, \text{ for } 1 \leq h \leq p \\
  i : B_h &\subseteq j : D_h, \text{ for } 1 \leq h \leq m \\
  j : \exists (R \sqcap \neg(\bigsqcup_{h=1}^{l} S_h)) \cdot (\neg \bigsqcup_{h=1}^{p} C_h) \subseteq (\bigsqcup_{h=1}^{m} D_h) \\
\end{align*}
\]
Complexity of Reasoning

- found four schemata of propagation rules which characterize the propagation of knowledge from $T_s$ to $T_t$;
- transformed the four propagation rules in a general operator $\mathcal{B}_{st}(T_s)$ which produces a T-box in $t$;
- shown soundness and completeness of the operator $\mathcal{B}_{st}(.)$:

$$\langle T_s, T_t, \mathcal{B}_{st} \rangle \models_{DDL} t : X \sqsubseteq Y \iff T_t \cup \mathcal{B}_{st}(T_s) \models_{DL} X \sqsubseteq Y$$

- checking the subsumption among concepts in $\langle T_s, T_t, \mathcal{B}_{st} \rangle$ is decidable and the complexity is $\text{ExpTime}$-complete on the number of onto bridge rules.
- checking the subsumption among concepts in the general case is also the same complexity.
Hint on Naive Procedure
Hint on Naive Procedure
Hint on Naive Procedure
Hint on Naive Procedure
Hint on Naive Procedure
Hint on Naive Procedure
Hint on Complexity of Reasoning

\[ s : A \sqsubseteq B_1 \cup B_2 \cup \ldots \cup B_n \]
\[ s : A \rightarrow t : C \]
\[ s : B_1 \rightarrow t : D_1 \]
\[ s : B_2 \rightarrow t : D_2 \]
\[ \ldots \]
\[ s : B_n \rightarrow t : D_n \]
\[ t : C \sqsubseteq D_1 \cup D_2 \cup \ldots \cup D_n \]

- Exponential time to compute \( \mathcal{B}_{st}(\mathcal{T}_s) \), possibly obtaining an exponential blow-up of \( \mathcal{T}_t \cup \mathcal{B}_{st}(\mathcal{T}_s) \);
- Exponential time to check concept subsumption of \( X \sqsubseteq Y \) in \( \mathcal{T}_t \cup \mathcal{B}_{st}(\mathcal{T}_s) \) (because of \( ALCQI_b \)).
- lower bound: complexity of \( ALCQI_b \) (\( EXPTime \))
- upper bound: complexity of the naive decision procedure (\( 2EXPTIME \))
Hint on Complexity of Reasoning

\[ s : A \sqsubseteq B_1 \sqcup B_2 \sqcup \ldots \sqcup B_n \]
\[ s : A \supseteq \exists \rightarrow t : C \]
\[ s : B_1 \sqsubseteq \exists \rightarrow t : D_1 \]
\[ s : B_2 \sqsubseteq \exists \rightarrow t : D_2 \]
\[ \ldots \]
\[ s : B_n \sqsubseteq \exists \rightarrow t : D_n \]
\[ t : C \sqsubseteq D_1 \sqcup D_2 \sqcup \ldots \sqcup D_n \]

- Exponential time to compute \( \mathcal{B}_{st}(\mathcal{T}_s) \), possibly obtaining an exponential blow-up of \( \mathcal{T}_t \cup \mathcal{B}_{st}(\mathcal{T}_s) \);
- Exponential time to check concept subsumption of \( X \sqsubseteq Y \) in \( \mathcal{T}_t \cup \mathcal{B}_{st}(\mathcal{T}_s) \) (because of \( ALCQI^b \)).
- lower bound: complexity of \( ALCQI^b \) (\( \text{ExpTime} \))
- upper bound: complexity of the naive decision procedure (\( 2\text{ExpTime} \))
Bringing Reification into the Picture

- Rules above don’t connect relevant attributes to domain and range of *marriedTo*
- Moreover domain and concept-role relation don’t interact
- Main idea: map attributes into “positional arguments”
Bringing Reification into the Picture

- Rules above don’t connect relevant attributes to domain and range of `marriedTo`
- Moreover domain and concept-role relation don’t interact
- Main idea: map attributes into “positional arguments”
Projection Bridge Rules

- new bridge rules bridging “attributes” and relations

\[ s : \text{wife} \rightarrow t : \text{marriedTo} \ #^1 \]
Conclusions

- We have
  - enriched the Heterogeneous DDL framework with a set of bridge rules to capture mappings between concepts and roles
  - characterised the propagation of new knowledge by means of a correct and complete set of propagation rules
  - studied the computational complexity of the local concept satisfiability problem in a network of $n$ T-boxes

- Ongoing research
  - extending the framework in order to capture reification
  - in T-box and A-box
  - devising restrictions in order to guarantee decidability