Armstrong ABoxes for $\mathcal{ALC}$ TBoxes

Henriette Harmse

ICTAC 2018
Armstrong ABoxes are inspired by Armstrong relations of relational database theory.

Armstrong ABoxes are the Description Logic counterpart of Armstrong relations.

A DL ontology or knowledge base consists of a TBox and an ABox.

An Armstrong ABox is an ABox that for a specific class of constraints, satisfies all constraints that hold, and violates all constraints that do not necessarily hold.
Armstrong ABoxes are formalized relative to particular classes of constraints, with each class of constraints resulting in a different Armstrong ABox formalization.

For arbitrary classes of constraints Armstrong ABoxes are undecidable.
Theoretical Basis

- Armstrong ABoxes for ALC TBoxes
- Ontology Completion
- Partial Contexts
- Attribute Exploration
- Formal Concept Analysis
- Description Logics
- Closed Sets in Lectic Order
- Armstrong ABoxes for ALC TBoxes

Henriette Harmse

Armstrong ABoxes for ALC TBoxes
Description Logics - Axioms, \( \mathcal{ALC} \) Concept Constructors

- The syntactic building blocks for a DL are based on the disjoint sets \( N_C \) (concept names), \( N_R \) (role names) and \( N_I \) (individual names).
- TBox axioms: \( C \sqsubseteq D, \ C \equiv D \).
- ABox assertions: \( C(x), \ r(x, y) \).
- \( \mathcal{ALC} \) concept descriptions (referred to as concepts) are constructed using the following concept constructors

\[
C := \top | A | \neg C | C \sqcap D | C \sqcup D | \exists r.C
\]

where \( A \) is an atomic concept, \( C \) and \( D \) are (possibly complex) concepts and \( r \) is a role.
Description Logics - Semantics of $\mathcal{ALC}$

For $\mathcal{ALC}$ for a given interpretation $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$, the interpretation function $\cdot^\mathcal{I}$ is extended to interpret complex concepts in the following way:

<table>
<thead>
<tr>
<th>Name</th>
<th>Constructor</th>
<th>Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>$\top^\mathcal{I}$</td>
<td>$\Delta^\mathcal{I}$</td>
</tr>
<tr>
<td>Bottom</td>
<td>$\bot^\mathcal{I}$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>Negation</td>
<td>$(\neg C)^\mathcal{I}$</td>
<td>$\Delta^\mathcal{I} \setminus C^\mathcal{I}$</td>
</tr>
<tr>
<td>Conjunction</td>
<td>$(C_1 \cap C_2)^\mathcal{I}$</td>
<td>$C_1^\mathcal{I} \cap C_2^\mathcal{I}$</td>
</tr>
<tr>
<td>Disjunction</td>
<td>$(C_1 \sqcup C_2)^\mathcal{I}$</td>
<td>$C_1^\mathcal{I} \cup C_2^\mathcal{I}$</td>
</tr>
<tr>
<td>Existential restriction</td>
<td>$(\exists r . C)^\mathcal{I}$</td>
<td>${ x \in \Delta^\mathcal{I}</td>
</tr>
</tbody>
</table>
Description Logics - Satisfaction

- $\mathcal{I} \models \alpha$ indicates that an interpretation $\mathcal{I}$ satisfies an axiom $\alpha$.
- Satisfaction of $\alpha$ is defined as

\[
\begin{align*}
\mathcal{I} \models C \subseteq D & \iff C^\mathcal{I} \subseteq D^\mathcal{I}, \\
\mathcal{I} \models C(x) & \iff x^\mathcal{I} \in C^\mathcal{I}, \text{ and} \\
\mathcal{I} \models r(x, y) & \iff (x^\mathcal{I}, y^\mathcal{I}) \in r^\mathcal{I}.
\end{align*}
\]
Description Logics - Models, Entailment

- $\mathcal{I}$ is a **model** of a TBox $\mathcal{T}$ (ABox $\mathcal{A}$) if it satisfies all its axioms (assertions).
- If $\mathcal{I}$ is a model of $\mathcal{T}$ and $\mathcal{A}$, it is called a model of the ontology $(\mathcal{T}, \mathcal{A})$ and $(\mathcal{T}, \mathcal{A})$ is said to be **consistent** if such a model exists.
- An axiom $\alpha$ is said to be **entailed** by an ontology $\mathcal{O}$, written as $\mathcal{O} \models \alpha$, if every model of $\mathcal{O}$ is also a model of $\alpha$.
- For a set of axioms $\Sigma = \{\sigma_0, \ldots, \sigma_n\}$, we abbreviate $\mathcal{O} \models \sigma_0, \ldots, \mathcal{O} \models \sigma_n$ with $\mathcal{O} \models \Sigma$.
- If $\mathcal{O}$ is empty, we abbreviate $\mathcal{O} \models \alpha$ as $\models \alpha$. 
Why is All This Necessary?

- Closed Sets in Lectic Order
- Formal Concept Analysis

> Attribute Exploration

- Partial Contexts

> Description Logics

- Ontology Completion

> Armstrong ABoxes for ALC TBoxes
Formal Concept Analysis - Basics

Formal context: \( K := (G, M, I) \) where

- \( G \) (objects),
- \( M \) (attributes), and
- \( I \subseteq G \times M \) (relation).

Implications between attributes can be used to analyse \( K \), i.e. \( \{m_2\} \rightarrow \{m_1, m_3\} \)
\( L \rightarrow R \) holds in \( K \) if every object that has all the attributes in \( L \) also has all the attributes in \( R \).

\( X \subseteq M \) respects an implication \( L \rightarrow R \) if \( L \not\subseteq X \) or \( R \subseteq X \).

\( X \subseteq M \) respects a set \( \mathcal{L} \) of implications if \( X \) respects every implication in \( \mathcal{L} \).

\( L \rightarrow R \) follows from \( \mathcal{L} \) if every \( X \subseteq M \) that respects all implications in \( \mathcal{L} \), also respects \( L \rightarrow R \).

\( \text{Mod}(\mathcal{L}) := \{X \subseteq M \mid X \text{ respects } \mathcal{L}\} \) is a closure system on \( M \), for which a closure operator \( \mathcal{L} : 2^M \rightarrow 2^M \) can be defined.
Attribute Exploration used to complete a subcontext $K'$. 

Iterates through implications from $\{\emptyset\} \rightarrow M$ to $M \rightarrow \{\emptyset\}$. 

$L$ is an implication base of $K$ if 

- every implication from $L$ holds in $K$, 
- every implication that holds in $K$ follows from $L$, and 
- no implication in $L$ follows from other implications in $L$. 

Minimal implication base, in particular a Duquenne Guigues base.
Formal Concept Analysis - Traversing Implications

- Steps to traverse Duquenne Guigues base.
  1. Start with $L = \{\emptyset\}$, find largest $R$ such that $L \rightarrow R$ does not have a counterexample in the $K'$.
  2. Find next $L$ to consider using \texttt{NextClosure} and the implication closure operator.

- Lectic order: Fixes some linear order on $M$ and defines lectic order such that for $A, B \subseteq M$ we can answer whether $A < B$?

- \texttt{NextClosure}: Given $A \subseteq M$ and some closure operator $\varphi$ it determines the next closed set $B$ in the lectic order such that $A < B$. 
From FCA to Ontology completion

**Table: Partial Context**

<table>
<thead>
<tr>
<th>$\mathcal{K}$</th>
<th>$m_1$</th>
<th>$m_2$</th>
<th>$m_3$</th>
<th>$m_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$o_1$</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>−</td>
</tr>
<tr>
<td>$o_2$</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>$o_3$</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>?</td>
</tr>
</tbody>
</table>

**Table: A Partial Context for an ontology ($\mathcal{T}, \mathcal{A}$)**

<table>
<thead>
<tr>
<th>$\mathcal{K}_{\mathcal{T}, \mathcal{A}}(M)$</th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$C_3$</th>
<th>$C_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>−</td>
</tr>
<tr>
<td>$x_2$</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>$x_3$</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>?</td>
</tr>
</tbody>
</table>
What are we doing again?

Diagram:

- Closed Sets in Lectic Order
- Formal Concept Analysis
- Attribute Exploration
- Partial Contexts
- Description Logics
- Ontology Completion
- Armstrong ABoxes for ALC TBoxes
Ontology Completion vs Armstrong ABoxes

- **Ontology Completion** assumes we start with an ontology \((\mathcal{T}_0, \mathcal{A}_0)\) and we want to determine an ontology \((\mathcal{T}, \mathcal{A})\) that is representative of the application domain.

- **Armstrong ABoxes** assumes we start with an ontology \((\mathcal{T}, \emptyset)\) that is representative of the application. We want to determine an ontology \((\mathcal{T}, \mathcal{A}_\bigcup)\) where \(\mathcal{A}_\bigcup\) represents perfect synthetic test/example data.
Armstrong ABoxes - Preliminaries

- For convenience the notation $\bigcap C_i$ and $\bigcap D_j$ will respectively be used as shorthand for $C_{i_0} \cap \ldots \cap C_{i_n}$ and $D_{j_0} \cap \ldots \cap D_{j_m}$.
- Armstrong ABoxes are restricted an interesting set $M$ of concepts.
- $M$ is said to be **permissible** if it is finite and no concept in $M$ is equivalent to $\top$.
- We define $M \rightarrow$ to be the set of GCIs representing the finite set of all the implications $L \rightarrow R$ over $M$. 
Armstrong ABoxes - Violating Exemplar

Let $\mathcal{T}$ be a consistent $\mathcal{ALC}$ TBox and let

$$\sigma' := \bigcap C_i \subseteq \bigcap D_j$$

for which $\mathcal{T} \nvDash \sigma'$ holds. An ABox $\mathcal{A}'$ is a **violating exemplar** of the entailment $\mathcal{T} \models \sigma'$ if

$$\left\{ (\bigcap C_i)(x), (\neg \bigcap D_j)(x) \right\} \subseteq \mathcal{A}'$$

holds for some named individual $x$ that does not appear in any other assertions of $\mathcal{A}'$. This is denoted by $\mathcal{A}' \not\vDash \sigma'$. 
Let $\mathcal{T}$ be a consistent $\mathcal{ALC}$ TBox, and let

$$\sigma := \bigcap C_i \subseteq \bigcap D_j$$

for which $\mathcal{T} \models \sigma$ and $\not\models \sigma$ holds. An ABox $\mathcal{A}$ is a **satisfying exemplar** of the entailment $\mathcal{T} \models \sigma$ if

$$\{ (\bigcap C_i)(x), (\bigcap D_j)(x) \} \subseteq \mathcal{A}$$

holds for some named individual $x$ that does not appear in any other assertions of $\mathcal{A}$. This is denoted by $\mathcal{A} \models \sigma$. 
Let $\mathcal{T}$ be a consistent $\mathcal{ALC}$ TBox and let $M$ be permissible. Let

$$\Sigma' := \{ \sigma' \mid \mathcal{T} \nvdash \sigma' \text{ and } \sigma' \in M \rightarrow \}.$$ 

$\Sigma'$ is called the **candidate axiom set** of $\mathcal{T}$ over $M$. Assume $\Sigma' = \{ \sigma'_0, \ldots, \sigma'_n \}$. An ABox $\mathcal{A}'$ is a **violating exemplar** of $\mathcal{T} \models \Sigma'$ if $\mathcal{A}' \nvdash \sigma'_0, \ldots, \mathcal{A}' \nvdash \sigma'_n$ holds. An ABox $\mathcal{A}'$ is a **minimal violating exemplar** of $\mathcal{T} \models \Sigma'$ iff there is no ABox $\mathcal{A}'_0 \subset \mathcal{A}'$ that is violating exemplar of $\mathcal{T} \models \Sigma'$.
Let $\mathcal{T}$ be a consistent $\mathcal{ALC}$ TBox and let $M$ be permissible. Let

$$\Sigma := \{ \sigma \mid \mathcal{T} \models \sigma, \ \not\models \sigma \text{ and } \sigma \in M^{-}\}.$$ 

$\Sigma$ is called the **entailment set** of $\mathcal{T}$ over $M$. Assume $\Sigma = \{\sigma_0, \ldots, \sigma_n\}$. An ABox $A$ is a **satisfying exemplar** of $\mathcal{T} \models \Sigma$ if $A \models \sigma_0, \ldots, A \models \sigma_n$ holds, which is denoted by $A \models \Sigma$. An ABox $A$ is a **minimal satisfying exemplar** of $\mathcal{T} \models \Sigma$ iff there is no ABox $A_0 \subset A$ that is satisfying exemplar of $\mathcal{T} \models \Sigma$. 
Let $\mathcal{T}$ be a consistent $\mathcal{ALC}$ TBox with $\Sigma$ and $\Sigma'$ respectively the entailment- and candidate axiom sets of $\mathcal{T}$. $\mathcal{A}_\Box$ is said to be an **Armstrong ABox** for $\mathcal{T}$ if and only if:

1. for every $\sigma \in \Sigma$, $\mathcal{A}_\Box \models \sigma$ holds,
2. for every $\sigma' \in \Sigma'$, $\mathcal{A}_\Box \not\models \sigma'$ holds and
3. there is no proper subset of $\mathcal{A}_\Box$ such that properties (1) and (2) hold.

$\mathcal{O}_\Box = \mathcal{T} \cup \mathcal{A}_\Box$ is called an **Armstrong ontology**.
An Example

\[ \mathcal{T}_0 = \{ \text{Course } \sqsubseteq \neg \text{Person}, \]
\[ \text{Teacher } \equiv \text{Person } \sqcap \exists \text{teaches.Course}, \]
\[ \exists \text{teaches.} \top \sqsubseteq \text{Person}, \]
\[ \text{Student } \equiv \text{Person } \sqcap \exists \text{attends.Course}, \]
\[ \exists \text{attends.} \top \sqsubseteq \text{Person} \} \]
**Table:** Armstrong ABox for $\mathcal{T}_0$ and $M = \{\text{Person, Student} \sqcup \text{Teacher}\}$

<table>
<thead>
<tr>
<th>Entailment</th>
<th>Satisfying exemplar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student $\sqcup$ Teacher $\sqsubseteq$ Person</td>
<td>${(\text{Student} \sqcup \text{Teacher})(x_2), \text{Person}(x_2)}$</td>
</tr>
<tr>
<td>Non-Entailment</td>
<td>Violating exemplar</td>
</tr>
<tr>
<td>Person $\sqsupseteq (\text{Student} \sqcup \text{Teacher})$ $\sqsubseteq$ Person $\sqcap (\text{Student} \sqcup \text{Teacher})$</td>
<td>${(\text{Person} \sqcap (\text{Student} \sqcup \text{Teacher}))(x_1), \neg (\text{Person} \sqcap (\text{Student} \sqcup \text{Teacher}))(x_1)}$</td>
</tr>
</tbody>
</table>
An Example (cont.)

\[ \mathcal{T}_1 = \mathcal{T}_0 \cup \{\text{Person} \equiv \text{Student} \sqcup \text{Teacher}\} \]

Table: Armstrong ABox for \( \mathcal{T}_1 \) and \( M = \{\text{Person}, \text{Student} \sqcup \text{Teacher}\} \)

<table>
<thead>
<tr>
<th>Entailment</th>
<th>Satisfying exemplar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person ( \sqcap ) (Student ( \sqcup ) Teacher) ( \sqsubseteq ) Person ( \sqcap ) (Student ( \sqcup ) Teacher)</td>
<td>{((\text{Person} \ sqcap (\text{Student} \ sqcup \text{Teacher}))(x_1), \ (\text{Person} \ sqcap (\text{Student} \ sqcup \text{Teacher}))(x_1))}</td>
</tr>
<tr>
<td>Student ( \sqcup ) Teacher ( \sqsubseteq ) Person</td>
<td>{((\text{Student} \ sqcup \text{Teacher})(x_2), \ \text{Person}(x_2))}</td>
</tr>
<tr>
<td>Person ( \sqsubseteq ) Student ( \sqcup ) Teacher</td>
<td>{\text{Person}(x_3), \ (\text{Student} \ sqcup \text{Teacher})(x_3)}</td>
</tr>
</tbody>
</table>
Potential Benefits

- Can help stakeholders to understand the meaning of entailments and non-entailments, which may be helpful to identify missing axioms.

- Where stakeholders are reasonably sure $\mathcal{T}$ is representative of their application domain, Armstrong ABoxes can give quick feedback when compared to ontology completion.

- Does not replace ontology completion. Rather, Armstrong ABoxes complements ontology completion.
Overview
Theoretical Basis
Description Logics
FCA, Partial Contexts, Ontology Completion
Armstrong ABox Formal Definition
An Example
Questions?