TECHNISCHE UNIVERSITÄT DRESDEN

Institute of Theoretical Computer Science Chair of Automata Theory

AXIOMATIZATION OF GENERAL CONCEPT INCLUSIONS FROM STREAMS OF INTERPRETATIONS WITH OPTIONAL ERROR TOLERANCE

Francesco Kriegel

FCA4AI @ ECAI, The Hague, 30 August 2016

Introduction: Two Problems

Global Problem

$$\begin{array}{ccccccccc} \mathcal{I}_0 & & \mathcal{I}_1 & & \mathcal{I}_2 & & \mathcal{I}_3 & & \cdots \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\ \mathcal{T}_0 & \rightarrow & \mathcal{T}_1 & \rightarrow & \mathcal{T}_2 & \rightarrow & \mathcal{T}_3 & \rightarrow & \cdots \end{array}$$

Input: Sequence of interpretations $\mathcal{I}_n$ with a time delay between consecutive interpretations

Output: Sequence of TBoxes $\mathcal{T}_n$ such that $\mathcal{T}_n \models C \sqsubseteq D$ if, and only if, $\mathcal{I}_1, \ldots, \mathcal{I}_n \models C \sqsubseteq D$

Introduction: Two Problems

Global Problem

$$\begin{array}{ccccccccc}
\mathcal{I}_0 & & \mathcal{I}_1 & & \mathcal{I}_2 & & \mathcal{I}_3 & & \cdots \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
\mathcal{T}_0 & \rightarrow & \mathcal{T}_1 & \rightarrow & \mathcal{T}_2 & \rightarrow & \mathcal{T}_3 & \rightarrow & \cdots
\end{array}$$

Input: Sequence of interpretations $\mathcal{I}_n$ with a time delay between consecutive interpretations

Output: Sequence of TBoxes $\mathcal{T}_n$ such that $\mathcal{T}_n \models C \sqsubseteq D$ if, and only if, $\mathcal{I}_1, \ldots, \mathcal{I}_n \models C \sqsubseteq D$

Local Problem

Input: A TBox $\mathcal{T}$ and an interpretation $\mathcal{I}$
Output: A TBox $\mathcal{U}$ such that $\mathcal{U} \models C \sqsubseteq D$ if, and only if, $\mathcal{T} \models C \sqsubseteq D$ as well as $\mathcal{I} \models C \sqsubseteq D$

The Description Logic $\mathcal{EL}^{\perp}$

Syntax

- *Signature* (N_C, N_R) where N_C is a set of *concept names* and N_R is a set of *role names*

The Description Logic $\mathcal{EL}^{\bot}$

Syntax

- *Signature* (N_C, N_R) where N_C is a set of *concept names* and N_R is a set of *role names*
- $\mathcal{EL}^{\bot}$*-concept descriptions* are inductively built by the rule

$$C ::= \bot \mid \top \mid A \mid C \sqcap C \mid \exists r.C$$

The Description Logic $\mathcal{EL}^{\bot}$

Syntax

- *Signature* (N_C, N_R) where N_C is a set of *concept names* and N_R is a set of *role names*
- $\mathcal{EL}^{\bot}$*-concept descriptions* are inductively built by the rule

$$C ::= \bot \mid \top \mid A \mid C \sqcap C \mid \exists r.C$$

- $\mathcal{EL}^{\bot}(N_C, N_R)$ is the set of all $\mathcal{EL}^{\bot}$-concept descriptions over (N_C, N_R)

The Description Logic $\mathcal{EL}^{\perp}$

Syntax

- *Signature* (N_C, N_R) where N_C is a set of *concept names* and N_R is a set of *role names*
- $\mathcal{EL}^{\perp}$*-concept descriptions* are inductively built by the rule

$$C ::= \perp \mid \top \mid A \mid C \sqcap C \mid \exists r.C$$

- $\mathcal{EL}^{\perp}(N_C, N_R)$ is the set of all $\mathcal{EL}^{\perp}$-concept descriptions over (N_C, N_R)
- *General concept inclusion (GCI)* $C \sqsubseteq D$

The Description Logic $\mathcal{EL}^\perp$

Syntax

- *Signature* (N_C, N_R) where N_C is a set of *concept names* and N_R is a set of *role names*
- $\mathcal{EL}^\perp$*-concept descriptions* are inductively built by the rule

$$C ::= \perp \mid \top \mid A \mid C \sqcap C \mid \exists r.C$$

- $\mathcal{EL}^\perp(N_C, N_R)$ is the set of all $\mathcal{EL}^\perp$-concept descriptions over (N_C, N_R)
- *General concept inclusion (GCI)* $C \sqsubseteq D$
- *Terminological Box (TBox)* $\mathcal{T}$ is a (finite) set of GCIs

The Description Logic $\mathcal{EL}^{\bot}$

Semantics

- *Interpretation* $\mathcal{I} \coloneqq (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ where $\Delta^{\mathcal{I}} \neq \emptyset$ is the *domain* and $\cdot^{\mathcal{I}}$ is a mapping such that $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ for all concept names $A \in N_C$ and $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ for all role names $r \in N_R$

The Description Logic $\mathcal{EL}^\bot$

Semantics

- *Interpretation* $\mathcal{I} := (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ where $\Delta^\mathcal{I} \neq \emptyset$ is the *domain* and $\cdot^\mathcal{I}$ is a mapping such that $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$ for all concept names $A \in N_C$ and $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$ for all role names $r \in N_R$
- Canonical extension to all $\mathcal{EL}^\bot$-concept descriptions:

$$\begin{aligned}
\bot^\mathcal{I} &:= \emptyset \\
\top^\mathcal{I} &:= \Delta^\mathcal{I} \\
(C \sqcap D)^\mathcal{I} &:= C^\mathcal{I} \cap D^\mathcal{I} \\
(\exists r.C)^\mathcal{I} &:= \{\, d \in \Delta^\mathcal{I} \mid \exists e \in \Delta^\mathcal{I} : (d,e) \in r^\mathcal{I} \text{ and } e \in C^\mathcal{I} \,\}
\end{aligned}$$

The Description Logic $\mathcal{EL}^\perp$

Semantics

- *Interpretation* $\mathcal{I} := (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ where $\Delta^\mathcal{I} \neq \emptyset$ is the *domain* and $\cdot^\mathcal{I}$ is a mapping such that $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$ for all concept names $A \in N_C$ and $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$ for all role names $r \in N_R$
- Canonical extension to all $\mathcal{EL}^\perp$-concept descriptions:

$$
\begin{aligned}
\perp^\mathcal{I} &:= \emptyset \\
\top^\mathcal{I} &:= \Delta^\mathcal{I} \\
(C \sqcap D)^\mathcal{I} &:= C^\mathcal{I} \cap D^\mathcal{I} \\
(\exists r. C)^\mathcal{I} &:= \{ d \in \Delta^\mathcal{I} \mid \exists e \in \Delta^\mathcal{I} : (d, e) \in r^\mathcal{I} \text{ and } e \in C^\mathcal{I} \}
\end{aligned}
$$

- $\mathcal{I} \models C \sqsubseteq D$ if $C^\mathcal{I} \subseteq D^\mathcal{I}$

The Description Logic $\mathcal{EL}^{\bot}$

Semantics

- *Interpretation* $\mathcal{I} := (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ where $\Delta^{\mathcal{I}} \neq \emptyset$ is the *domain* and $\cdot^{\mathcal{I}}$ is a mapping such that $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ for all concept names $A \in N_C$ and $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ for all role names $r \in N_R$
- Canonical extension to all $\mathcal{EL}^{\bot}$-concept descriptions:

$$\begin{aligned} \bot^{\mathcal{I}} &:= \emptyset \\ \top^{\mathcal{I}} &:= \Delta^{\mathcal{I}} \\ (C \sqcap D)^{\mathcal{I}} &:= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\ (\exists r.C)^{\mathcal{I}} &:= \{\, d \in \Delta^{\mathcal{I}} \mid \exists e \in \Delta^{\mathcal{I}} \colon (d,e) \in r^{\mathcal{I}} \text{ and } e \in C^{\mathcal{I}} \,\} \end{aligned}$$

- $\mathcal{I} \models C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
- $\mathcal{I} \models \mathcal{T}$ if $\mathcal{I} \models C \sqsubseteq D$ for all GCIs $C \sqsubseteq D \in \mathcal{T}$

The Description Logic $\mathcal{EL}^{\bot}$

Semantics

- *Interpretation* $\mathcal{I} := (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ where $\Delta^{\mathcal{I}} \neq \emptyset$ is the *domain* and $\cdot^{\mathcal{I}}$ is a mapping such that $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ for all concept names $A \in N_C$ and $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ for all role names $r \in N_R$
- Canonical extension to all $\mathcal{EL}^{\bot}$-concept descriptions:

$$\begin{aligned} \bot^{\mathcal{I}} &:= \emptyset \\ \top^{\mathcal{I}} &:= \Delta^{\mathcal{I}} \\ (C \sqcap D)^{\mathcal{I}} &:= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\ (\exists r.C)^{\mathcal{I}} &:= \{\, d \in \Delta^{\mathcal{I}} \mid \exists e \in \Delta^{\mathcal{I}} \colon (d,e) \in r^{\mathcal{I}} \text{ and } e \in C^{\mathcal{I}} \,\} \end{aligned}$$

- $\mathcal{I} \models C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
- $\mathcal{I} \models \mathcal{T}$ if $\mathcal{I} \models C \sqsubseteq D$ for all GCIs $C \sqsubseteq D \in \mathcal{T}$
- $\mathcal{T} \models C \sqsubseteq D$ if $\mathcal{I} \models C \sqsubseteq D$ for all models $\mathcal{I} \models \mathcal{T}$

The Description Logic $\mathcal{EL}^{\bot}$

Semantics

- *Interpretation* $\mathcal{I} := (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ where $\Delta^{\mathcal{I}} \neq \emptyset$ is the *domain* and $\cdot^{\mathcal{I}}$ is a mapping such that $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ for all concept names $A \in N_C$ and $r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ for all role names $r \in N_R$
- Canonical extension to all $\mathcal{EL}^{\bot}$-concept descriptions:

$$
\begin{aligned}
\bot^{\mathcal{I}} &:= \emptyset \\
\top^{\mathcal{I}} &:= \Delta^{\mathcal{I}} \\
(C \sqcap D)^{\mathcal{I}} &:= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\
(\exists r. C)^{\mathcal{I}} &:= \{\, d \in \Delta^{\mathcal{I}} \mid \exists e \in \Delta^{\mathcal{I}} \colon (d,e) \in r^{\mathcal{I}} \text{ and } e \in C^{\mathcal{I}} \,\}
\end{aligned}
$$

- $\mathcal{I} \models C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
- $\mathcal{I} \models \mathcal{T}$ if $\mathcal{I} \models C \sqsubseteq D$ for all GCIs $C \sqsubseteq D \in \mathcal{T}$
- $\mathcal{T} \models C \sqsubseteq D$ if $\mathcal{I} \models C \sqsubseteq D$ for all models $\mathcal{I} \models \mathcal{T}$
- $\mathcal{T} \models \mathcal{U}$ if $\mathcal{I} \models \mathcal{U}$ for all models $\mathcal{I} \models \mathcal{T}$

Least Common Subsumer

The *least common subsumer (lcs)* $C \vee D$ of $\mathcal{EL}^{\bot}$-concept descriptions C and D is defined by the following conditions:

- $\emptyset \models C \sqsubseteq C \vee D$
- $\emptyset \models D \sqsubseteq C \vee D$
- For all $\mathcal{EL}^{\bot}$-concept descriptions E, $\emptyset \models \{C \sqsubseteq E, D \sqsubseteq E\}$ implies $\emptyset \models C \vee D \sqsubseteq E$.

The Description Logic $\mathcal{EL}^\bot$

Least Common Subsumer

The *least common subsumer (lcs)* $C \vee D$ of $\mathcal{EL}^\bot$-concept descriptions C and D is defined by the following conditions:

- $\emptyset \models C \sqsubseteq C \vee D$
- $\emptyset \models D \sqsubseteq C \vee D$
- For all $\mathcal{EL}^\bot$-concept descriptions E, $\emptyset \models \{C \sqsubseteq E, D \sqsubseteq E\}$ implies $\emptyset \models C \vee D \sqsubseteq E$.

The Lattice of $\mathcal{EL}^\bot$-Concept Descriptions

$\boldsymbol{\mathcal{EL}}^\bot(N_C, N_R) := (\mathcal{EL}^\bot(N_C, N_R), \sqsubseteq)/\equiv$ is a lattice where

- $\sqsubseteq$ denotes subsumption w.r.t. the empty TBox $\emptyset$
- $\equiv$ is the induced equivalence relation of $\sqsubseteq$
- Conjunction $\sqcap$ is the *infimum*
- Lcs $\vee$ is the *supremum*
- $\bot$ is the smallest, and $\top$ the greatest element

An Existing Static Approach

- Baader and Distel have found a method for the construction of a base of GCIs for (finite) interpretations.
- They utilize so-called *model-based most specific concept descriptions* and a transformation to *Formal Concept Analysis*.

An Existing Static Approach

- Baader and Distel have found a method for the construction of a base of GCIs for (finite) interpretations.
- They utilize so-called *model-based most specific concept descriptions* and a transformation to *Formal Concept Analysis*.

Most Specific Concept Description

Given: Interpretation $\mathcal{I}$, subset $X \subseteq \Delta^{\mathcal{I}}$
The *model-based most specific concept description (mmsc)* $X^{\mathcal{I}}$ of X with respect to $\mathcal{I}$ is defined by the following conditions:

- $X \subseteq X^{\mathcal{I}\mathcal{I}}$
- For all $\mathcal{EL}^{\perp}$-concept descriptions C, $X \subseteq C^{\mathcal{I}}$ implies $\emptyset \models X^{\mathcal{I}} \sqsubseteq C$.

An Existing Static Approach

- Baader and Distel have found a method for the construction of a base of GCIs for (finite) interpretations.
- They utilize so-called *model-based most specific concept descriptions* and a transformation to *Formal Concept Analysis*.

Most Specific Concept Description

Given: Interpretation $\mathcal{I}$, subset $X \subseteq \Delta^{\mathcal{I}}$
The *model-based most specific concept description (mmsc)* $X^{\mathcal{I}}$ of X with respect to $\mathcal{I}$ is defined by the following conditions:

- $X \subseteq X^{\mathcal{I}\mathcal{I}}$
- For all $\mathcal{EL}^{\perp}$-concept descriptions C, $X \subseteq C^{\mathcal{I}}$ implies $\emptyset \models X^{\mathcal{I}} \sqsubseteq C$.

- Existence of mmscs either via gfp-semantics or restriction of role depth.
- A slight reformulation of their results yields the following statements.

An Existing Static Approach

- Baader and Distel have found a method for the construction of a base of GCIs for (finite) interpretations.
- They utilize so-called *model-based most specific concept descriptions* and a transformation to *Formal Concept Analysis*.

Most Specific Concept Description

Given: Interpretation $\mathcal{I}$, subset $X \subseteq \Delta^{\mathcal{I}}$
The *model-based most specific concept description (mmsc)* $X^{\mathcal{I}}$ of X with respect to $\mathcal{I}$ is defined by the following conditions:

- $X \subseteq X^{\mathcal{I}\mathcal{I}}$
- For all $\mathcal{EL}^{\bot}$-concept descriptions C, $X \subseteq C^{\mathcal{I}}$ implies $\emptyset \models X^{\mathcal{I}} \sqsubseteq C$.

- Existence of mmscs either via gfp-semantics or restriction of role depth.
- A slight reformulation of their results yields the following statements.
- The mapping $C \mapsto C^{\mathcal{I}\mathcal{I}}$ is a closure operator in the dual of $\boldsymbol{\mathcal{EL}}^{\bot}(N_C, N_R)$.
- A base of $C \mapsto C^{\mathcal{I}\mathcal{I}}$ is a base of GCIs for $\mathcal{I}$.

Most Specific Consequences and a new Closure Operator

Most Specific Consequence

Given: $\mathcal{EL}^{\bot}$-concept description C, $\mathcal{EL}^{\bot}$-TBox $\mathcal{T}$
The *most specific consequence* $C^{\mathcal{T}}$ of C with respect to $\mathcal{T}$ is defined by the following conditions:

$\mathcal{T} \models C \sqsubseteq C^{\mathcal{T}}$.

For all $\mathcal{EL}^{\bot}$-concept descriptions D, $\mathcal{T} \models C \sqsubseteq D$ implies $\emptyset \models C^{\mathcal{T}} \sqsubseteq D$.

Most Specific Consequences and a new Closure Operator

Most Specific Consequence

Given: $\mathcal{EL}^\bot$-concept description C, $\mathcal{EL}^\bot$-TBox $\mathcal{T}$
The *most specific consequence* $C^\mathcal{T}$ of C with respect to $\mathcal{T}$ is defined by the following conditions:

$\mathcal{T} \models C \sqsubseteq C^\mathcal{T}$.

For all $\mathcal{EL}^\bot$-concept descriptions D, $\mathcal{T} \models C \sqsubseteq D$ implies $\emptyset \models C^\mathcal{T} \sqsubseteq D$.

- Existence of most specific consequences either via gfp-semantics or restriction of role depth.
- The mapping $C \mapsto C^\mathcal{T}$ is a closure operator in the dual of $\boldsymbol{\mathcal{EL}}^\bot(N_C, N_R)$.
- A GCI is entailed by $\mathcal{T}$ if, and only if, it is valid in $C \mapsto C^\mathcal{T}$.

Most Specific Consequences and a new Closure Operator

Most Specific Consequence

Given: $\mathcal{EL}^{\bot}$-concept description C, $\mathcal{EL}^{\bot}$-TBox $\mathcal{T}$
The *most specific consequence* $C^{\mathcal{T}}$ of C with respect to $\mathcal{T}$ is defined by the following conditions:

$\mathcal{T} \models C \sqsubseteq C^{\mathcal{T}}$.

For all $\mathcal{EL}^{\bot}$-concept descriptions D, $\mathcal{T} \models C \sqsubseteq D$ implies $\emptyset \models C^{\mathcal{T}} \sqsubseteq D$.

- Existence of most specific consequences either via gfp-semantics or restriction of role depth.
- The mapping $C \mapsto C^{\mathcal{T}}$ is a closure operator in the dual of $\boldsymbol{\mathcal{EL}}^{\bot}(N_C, N_R)$.
- A GCI is entailed by $\mathcal{T}$ if, and only if, it is valid in $C \mapsto C^{\mathcal{T}}$.

Observation (Solution for the Local Problem)

To construct a base of GCIs that are both valid in $\mathcal{I}$ and entailed by $\mathcal{T}$, it suffices to compute a base of implications of the infimum of $C \mapsto C^{\mathcal{I}\mathcal{I}}$ and $C \mapsto C^{\mathcal{T}}$.

Error-Tolerance

- Dualizing the previous approach, we get an error-tolerant way for axiomatizing GCIs from an interpretation.

Error-Tolerance

– Dualizing the previous approach, we get an error-tolerant way for axiomatizing GCIs from an interpretation.

Observation

A base of the supremum of $C \mapsto C^{\mathcal{I}\mathcal{I}}$ and $C \mapsto C^{\mathcal{T}}$ yields a base for only GCIs that are valid for the individuals in $\mathcal{I}$ which respect the GCIs in $\mathcal{T}$, i.e., satisfy some constraints (e.g. hand-crafted by an expert).

Error-Tolerance

- Dualizing the previous approach, we get an error-tolerant way for axiomatizing GCIs from an interpretation.

Observation

A base of the supremum of $C \mapsto C^{\mathcal{I}\mathcal{I}}$ and $C \mapsto C^{\mathcal{T}}$ yields a base for only GCIs that are valid for the individuals in $\mathcal{I}$ which respect the GCIs in $\mathcal{T}$, i.e., satisfy some constraints (e.g. hand-crafted by an expert).

- It is also possible to combine both approaches (infima/suprema) to construct bases of GCIs from streams of possibly erroneous interpretations.

Thanks for your attention!

Thanks for your attention!

Questions?