Knowledge Representation for the Semantic Web

Winter Quarter 2012

Slides 6 – 02/02/2012

Pascal Hitzler
Kno.e.sis Center
Wright State University, Dayton, OH
http://www.knoesis.org/pascal/
Textbook (required)

Pascal Hitzler, Markus Krötzsch, Sebastian Rudolph

Foundations of Semantic Web Technologies
Chapman & Hall/CRC, 2010

Choice Magazine Outstanding Academic Title 2010 (one out of seven in Information & Computer Science)

http://www.semantic-web-book.org
A Semantic Puzzle


\[
\begin{align*}
\text{Person} & \sqsubseteq \neg \text{Movie} \\
\text{RRated} & \sqsubseteq \text{CatMovie} \\
\text{CatMovie} & \sqsubseteq \text{Movie} \\
\text{RRated} & \equiv (\exists \text{hasScript. ThrillerScript}) \sqsubseteq (\forall \text{hasViolenceLevel. High}) \sqsubseteq \text{Domain(hasViolenceLevel, Movie)}
\end{align*}
\]

Fig. 1. A justification for Person \(\sqsubseteq \bot\)
Today: Model-theoretic Semantics
Today’s Session: DL Semantics

1. Model-theoretic Semantics of SROIQ(D)
2. Class Project
3. Class Presentations
Model-theoretic Semantics

• Recall:

How does one make a model-theoretic semantics?

What – which mathematical entity – actually captures the “meaning”?

How would we get at this here?
Model-theoretic Semantics

- There are two semantics for OWL.

1. Description Logic Semantics
   also: Direct Semantics; FOL Semantics
   Can be obtained by translation to FOL.
   Some global restrictions apply! (see next slide)

2. RDF-based Semantics (requires RDF/XML syntax: done later)
   No syntax restrictions apply.
   Extends the direct semantics with RDFS-reasoning features.

In the following, we will deal with the direct semantics only.
Direct Semantics

To obtain decidability, syntactic restrictions apply.

• Type separation / punning

• No cycles in property chains.  
  (See global restrictions mentioned earlier.)

• No transitive properties in cardinality restrictions.  
  (See global restrictions mentioned earlier.)
Decidability

• A problem is *decidable* if there exists an always terminating algorithm which determines, whether or not a solution exists.

• A problem is *semi-decidable* if there exists an algorithm which, in case a solution exists, finds this out in finite time.

• A problem is *undecidable* if it is not decidable.

• Note there exist problems which are semi-decidable and undecidable.
Decidability of DLs

• A description logic is decidable if “entailment of axioms” is decidable.

• Most description logics are decidable. Decidability is one of the design criteria for “good” description logics.
Direct Semantics

- model-theoretic semantics
- starts with interpretations
- an interpretation $\mathcal{I}$ maps

  individual names, class names and property names...

...into a domain
Interpretation Example

If we consider, for example, the knowledge base consisting of the axioms

\[
\text{Professor} \sqsubseteq \text{FacultyMember} \\
\text{Professor}(\text{rudiStuder}) \\
\text{hasAffiliation}(\text{rudiStuder}, \text{aifb})
\]

then we could set

\[
\Delta = \{a, b, \text{Ian}\} \\
I_I(\text{rudiStuder}) = \text{Ian} \\
I_I(\text{aifb}) = b \\
I_C(\text{Professor}) = \{a\} \\
I_C(\text{FacultyMember}) = \{a, b\} \\
I_R(\text{hasAffiliation}) = \{(a, b), (b, \text{Ian})\}
\]

Intuitively, these settings are nonsense, but they nevertheless determine a valid interpretation.
OWL Direct Semantics

- mapping is extended to complex class expressions:
  - $T^I = \Delta^I$
  - $\bot^I = \emptyset$
  - $(C \cap D)^I = C^I \cap D^I$
  - $(C \cup D)^I = C^I \cup D^I$
  - $(\neg C)^I = \Delta^I \setminus C^I$
  - $(\forall R.C)^I = \{ x | \text{for all } (x,y) \in R^I \text{ we have } y \in C^I \}$
  - $(\exists R.C)^I = \{ x | \text{there is } (x,y) \in R^I \text{ with } y \in C^I \}$
  - $(\geq n R.C)^I = \{ x | \#\{ y | (x,y) \in R^I \text{ and } y \in C^I \} \geq n \}$
  - $(\leq n R.C)^I = \{ x | \#\{ y | (x,y) \in R^I \text{ and } y \in C^I \} \leq n \}$
- ...and to role expressions:
  - $U^I = \Delta^I \times \Delta^I$
  - $(R^-)^I = \{ (y,x) | (x,y) \in R^I \}$
- ...and to axioms:
  - $C(a)$ holds, if $a^I \in C^I$
  - $R(a,b)$ holds, if $(a^I,b^I) \in R^I$
  - $C \subseteq D$ holds, if $C^I \subseteq D^I$
  - $R \subseteq S$ holds, if $R^I \subseteq S^I$
  - Disjoint$(R,S)$ holds if $R^I \cap S^I = \emptyset$
  - $S_1 \circ S_2 \circ ... \circ S_n \subseteq R$ holds if $S_1^I \circ S_2^I \circ ... \circ S_n^I \subseteq R^I$
OWL Direct Semantics

• what’s below gives us a notion of *model*:

An interpretation is a model of a set of axioms if all the axioms hold (are evaluated to true) in the interpretation.

• Notion of *logical consequence* obtained as usual.

• ...and to axioms:

  – $C(a)$ holds, if $a^I \in C^I$
  – $C \subseteq D$ holds, if $C^I \subseteq D^I$
  – Disjoint($R,S$) holds if $R^I \cap S^I = \emptyset$
  – $S_1 \circ S_2 \circ \ldots \circ S_n \subseteq R$ holds if $S_1^I \circ S_2^I \circ \ldots \circ S_n^I \subseteq R^I$

• what’s below gives us a notion of *model*:

An interpretation is a model of a set of axioms if all the axioms hold (are evaluated to true) in the interpretation.

• Notion of *logical consequence* obtained as usual.
Not a model!

If we consider, for example, the knowledge base consisting of the axioms

\[
\begin{align*}
\text{Professor} & \sqsubseteq \text{FacultyMember} \\
\text{Professor}(\text{rudiStuder}) \\
\text{hasAffiliation}(\text{rudiStuder}, \text{aifb})
\end{align*}
\]

then we could set

\[
\Delta = \{a, b, \text{Ian}\}
\]

\[
\begin{align*}
\mathcal{I}(\text{rudiStuder}) &= \text{Ian} \\
\mathcal{I}(\text{aifb}) &= b \\
\mathcal{I}(\text{Professor}) &= \{a\} \\
\mathcal{I}(\text{FacultyMember}) &= \{a, b\} \\
\mathcal{I}(\text{hasAffiliation}) &= \{(a, b), (b, \text{Ian})\}
\end{align*}
\]

Intuitively, these settings are nonsense, but they nevertheless determine a valid interpretation.
A model

\[
\begin{align*}
\text{Professor} & \sqsubseteq \text{FacultyMember} \\
\text{Professor}(\text{rudiStuder}) & \\
\text{hasAffiliation}(\text{rudiStuder}, \text{aifb})
\end{align*}
\]

\[
\Delta = \{a, r, s\}
\]

\[
\begin{align*}
I_I(\text{rudiStuder}) &= r \\
I_I(\text{aifb}) &= a \\
I_C(\text{Professor}) &= \{r\} \\
I_C(\text{FacultyMember}) &= \{r, s\} \\
I_R(\text{hasAffiliation}) &= \{(r, a)\}
\end{align*}
\]
Professor ⊆ FacultyMember
Professor(rudiStuder)
hasAffiliation(rudiStuder, aifb)

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ</td>
<td>{a, r, s}</td>
<td>{1, 2}</td>
<td>{♣}</td>
</tr>
<tr>
<td>I (rudiStuder)</td>
<td>(r)</td>
<td>(1)</td>
<td>♣</td>
</tr>
<tr>
<td>I (aifb)</td>
<td>(a)</td>
<td>(2)</td>
<td>♣</td>
</tr>
<tr>
<td>I C(Professor)</td>
<td>{(r)}</td>
<td>{1}</td>
<td>{♣}</td>
</tr>
<tr>
<td>I C(FacultyMember)</td>
<td>{a, r, s}</td>
<td>{1, 2}</td>
<td>{♣}</td>
</tr>
<tr>
<td>I R(hasAffiliation)</td>
<td>{(r, a)}</td>
<td>{(1, 1), (1, 2)}</td>
<td>{(♣, ♣)}</td>
</tr>
</tbody>
</table>

Is FacultyMember(aifb) a logical consequence?
Returning to our running example knowledge base, let us show formally that \textit{FacultyMember(aifb)} is not a logical consequence. This can be done by giving a model \( M \) of the knowledge base where \( aifb^M \not\in FacultyMember^M \). The following determines such a model.

\[ \Delta = \{ a, r \} \]

\[ I_I(rudiStuder) = r \]
\[ I_I(aifb) = a \]
\[ I_C(Professor) = \{ r \} \]
\[ I_C(FacultyMember) = \{ r \} \]
\[ I_R(hasAffiliation) = \{ (r, a) \} \]
but often OWL 2 DL is said to be a fragment of first-order predicate logic (FOL) [with equality]...

yes, there is a translation of OWL 2 DL into FOL

\[
\pi(C \sqsubseteq D) = (\forall x)(\pi_x(C) \rightarrow \pi_x(D)) \\
\pi_x(A) = A(x) \\
\pi_x(\neg C) = \neg \pi_x(C) \\
\pi_x(C \sqcap D) = \pi_x(C) \land \pi_x(D) \\
\pi_x(C \sqcup D) = \pi_x(C) \lor \pi_x(D) \\
\pi_x(\forall R.C) = (\forall x_1)(R(x_1, x) \rightarrow \pi_x_1(C)) \\
\pi_x(\exists R.C) = (\exists x_1)(R(x_1, x) \land \pi_x_1(C)) \\
\pi_x(\geq n S.C) = (\exists x_1) \ldots (\exists x_n) \left( \bigwedge_{i \neq j} (x_i \neq x_j) \land \bigwedge_i (S(x, x_i) \land \pi_x_1(C)) \right) \\
\pi_x(\leq n S.C) = \neg (\exists x_1) \ldots (\exists x_{n+1}) \left( \bigwedge_{i \neq j} (x_i \neq x_j) \land \bigwedge_i (S(x, x_i) \land \pi_x_1(C)) \right) \\
\pi_x(\{a\}) = (x = a) \\
\pi_x(\exists S.\text{Self}) = S(x, x) \\
\pi_x(\exists R_1 \ldots R_n) = (\exists x_1) \ldots (\exists x_n) \left( \bigwedge_{i=1}^{n-2} \pi_{x_1, x_{i+1}}(R_i) \land \pi_{x_{n-1}, y}(R_n) \right) \\
\pi(\text{Ref}(R)) = (\forall x)\pi_{x, x}(R) \\
\pi(\text{Asy}(R)) = (\forall x)(\forall y)(\pi_{x, y}(R) \rightarrow \neg \pi_{y, x}(R)) \\
\pi(\text{Dis}(R_1, R_2)) = (\exists y)(\exists y)(\pi_{x, y}(R_1) \land \pi_{x, y}(R_2))
\]
Inconsistency and Satisfiability

- A set of axioms (knowledge base) is satisfiable (or consistent) if it has a model.
- It is unsatisfiable (inconsistent) if it does not have a model.

- Inconsistency is often caused by modeling errors.

- Unicorn(beauty)
  Unicorn ⊆ Fictitious
  Unicorn ⊆ Animal
  Animal ⊆ ¬Fictitious
Inconsistency and Satisfiability

- A knowledge base is incoherent if a named class is equivalent to $\bot$.
- It usually also points to a modeling error.

\[
\text{Unicorn} \sqsubseteq \text{Fictitious} \\
\text{Unicorn} \sqsubseteq \text{Animal} \\
\text{Fictitious} \sqcap \text{Animal} \sqsubseteq \bot
\]
A Semantic Puzzle


\[
\begin{align*}
\text{Person} & \subseteq \neg \text{Movie} \\
\text{RRated} & \subseteq \text{CatMovie} \\
\text{CatMovie} & \subseteq \text{Movie} \\
\text{RRated} & \equiv (\exists \text{hasScript. ThrillerScript}) \sqsubseteq (\forall \text{hasViolenceLevel. High}) \sqsubseteq \text{Domain}(\text{hasViolenceLevel, Movie})
\end{align*}
\]

**Fig. 1.** A justification for Person $\subseteq \bot$
Rationale behind OWL

- Open World Assumption
- Favourable trade-off between expressivity and scalability
- Integrates with RDFS
- Purely declarative semantics

Features:
- Fragment of first-order predicate logic (FOL)
- Decidable
- Known complexity classes (N2ExpTime for OWL 2 DL)
- Reasonably efficient for real KBs
Today’s Session: DL Semantics

1. Model-theoretic Semantics of SROIQ(D)
2. Class Project
3. Class Presentations
Class Project

• Use the classes and properties from your ontology (if necessary, add some new ones).
• Use them as class names and role names, and write down (in DL notation) a number of SROIQ axioms which make sense in the context of your project ontology.
• If you find it appropriate, feel free to completely rewrite your ontology.
• Make sure you use each of the following constructs at least once:
  – \( \cap, \cup, \neg, \exists, \forall \)
  – a nominal
  – an inverse property
  – a qualified cardinality constraint
  – three of the property characteristics
Class Project

• Send me by Sunday 12th of February 9pm:
  – Current version of your ontology in Turtle syntax (those parts not expressed using DL axioms).
  – The DL axioms (comprising the rest of your ontology).
    • Either on paper, handwritten (e.g. via Tonya Davis for me)
    • Or as a pdf (e.g. generated from LaTex).
  – Each DL axiom accompanied with a natural language sentence which captures its meaning.
Today’s Session: DL Semantics

1. Model-theoretic Semantics of SROIQ(D)
2. Class Project
3. Class Presentations
Class presentations – topics

- **SPARQL 1.1 entailment regimes:**
  http://www.w3.org/2009/sparql/docs/entailment/xmlspec.xml

- Aidan Hogan, Andreas Harth, Axel Polleres: **SAOR: Authoritative Reasoning for the Web.** ASWC 2008: 76-90

- Jacopo Urbani, Spyros Kotoulas, Jason Maassen, Frank van Harmelen, Henri E. Bal: **OWL Reasoning with WebPIE: Calculating the Closure of 100 Billion Triples.** ESWC (1) 2010: 213-227

- Yuan Ren, Jeff Z. Pan, Yuting Zhao: **Soundness Preserving Approximation for TBox Reasoning.** AAAI 2010
Class presentations – topics


Class presentations – topics


- Jiao Tao, Evren Sirin, Jie Bao, Deborah L. McGuinness: Integrity Constraints in OWL. In: Maria Fox, David Poole (Eds.): Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, July 11-15, 2010. AAAI Press 2010

Class presentations – topics


Presentation format

- 30 minutes, including 5 minutes questions (timing will be strict)

- Content selection is up to you. Presentation must contain the key results from the paper you are presenting. It can also contain material related to these key results which are not in the paper itself (but which you may find more interesting than the rest of the paper).

- Let me know by 19th of February at the latest (by email) which paper you want to present. It’s first come, first serve.
- Presentations will be in the last two weeks of classes.
Presentation evaluation dimensions

Example criteria – may not all be weighted equally:

- Quality of slides
- Quality and effectiveness of explanations
- Quality of presentation style (use of verbal and body language, use of media, flexibility in case of interaction with audience, time management)
- Correctness of content
- Grade of reaching the audience and getting the content across
- How “interesting” the material is presented
- Competence in answering questions
Tuesday 10\textsuperscript{th} of January: RDF Schema
Thursday 12\textsuperscript{th} of January: RDF and RDFS Semantics
Tuesday 17\textsuperscript{th} of January: RDF and RDFS Semantics
Thursday 19\textsuperscript{th} of January: exercise session 1
Tuesday 24\textsuperscript{th} of January: OWL part 1 – Description Logics
Thursday 2\textsuperscript{nd} of February: OWL pt 2 – model-theoretic Semantics
Tuesday 7\textsuperscript{th} of February: Partonomies
Thursday 9\textsuperscript{th} of February: SPARQL
Tuesday 14\textsuperscript{th} of February: OWL part 3 – web syntax
Thursday 16\textsuperscript{th} of February: exercise session 2