OWL 2 – Theory and Practice

Bernardo Cuenca Grau
University of Oxford
UK

Pascal Hitzler
Kno.e.sis Center
Wright State University
Dayton, OH, USA

Birte Glimm
University of Oxford
UK

Hector Perez-Urbina
Clark & Parsia, LLC
Today’s Schedule

09:00 - 10:00 OWL introduction (Pascal)

10:00 - 10:30 coffee break

10:30 - 12:30 OWL introduction (Pascal)

12:30 - 14:00 lunch

14:00 - 16:00 hands-on session (Birte)

16:00 - 16:30 coffee break

16:30 - 18:30 applications (Bernardo)
Textbook

Pascal Hitzler, Markus Krötzsch, Sebastian Rudolph

Foundations of Semantic Web Technologies
Chapman & Hall/CRC, 2009

Grab a flyer!

http://www.semantic-web-book.org
Pascal Hitzler, Markus Krötzsch, Sebastian Rudolph

语义Web技术基础
Tsinghua University Press (清华大学出版社), 2011, to appear

Translators:
Yong Yu, Haofeng Wang, Guilin Qi (俞勇，王昊奋，漆桂林)

http://www.semantic-web-book.org
Available from

Part 1

OWL 2 – Syntax, Semantics, Reasoning
Main References Part 1


OWL 2 Document Overview: http://www.w3.org/TR/owl2-overview/

OWL – Overview

- Web Ontology Language
  - W3C Recommendation for the Semantic Web, 2004
  - OWL 2 (revised W3C Recommendation), 2009

- Semantic Web KR language based on description logics (DLs)
  - OWL DL is essentially DL SROIQ(D)
  - KR for web resources, using URIs.
  - Using web-enabled syntaxes, e.g. based on XML or RDF.

We present

- DL syntax (used in research – not part of the W3C recommendation)
- (some) RDF Turtle syntax
Contents

• OWL – Basic Ideas
• OWL as the Description Logic SROIQ(D)
• Different Perspectives on OWL
• OWL Semantics
• OWL Profiles
• Proof Theory
• Tools
Contents

• OWL – Basic Ideas
• OWL as the Description Logic SROIQ(D)
• Different Perspectives on OWL
• OWL Semantics
• OWL Profiles
• Proof Theory
• Tools
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
• **individuals (written as URIs)**
  – also: constants (FOL), resources (RDF)
  – [http://example.org/sebastianRudolph](http://example.org/sebastianRudolph)
  – we write these lowercase and abbreviated, e.g. "sebastianRudolph"

• **classes (also written as URIs!)**
  – also: concepts, unary predicates (FOL)
  – we write these uppercase, e.g. "Father"

• **properties (also written as URIs!)**
  – also: roles (DL), binary predicates (FOL)
  – we write these lowercase, e.g. "hasDaughter"
DL syntax | FOL syntax
--- | ---
| | ABox statements
| | TBox statements
- Person(mary) | Person(mary)
- Woman $\subseteq$ Person
  - Person $\equiv$ HumanBeing
- hasWife(john,mary) | $\forall x \ (\text{Woman}(x) \rightarrow \text{Person}(x))$
- hasWife $\subseteq$ hasSpouse
  - hasSpouse $\equiv$ marriedWith
- $\forall x \ \forall y \ (\text{hasWife}(x,y) \rightarrow \text{hasSpouse}(x,y))$
<table>
<thead>
<tr>
<th>DL syntax</th>
<th>FOL syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Person(mary)</td>
<td>• :mary rdf:type :Person .</td>
</tr>
<tr>
<td>• Woman ⊆ Person</td>
<td>• :Woman rdfs:subClassOf :Person .</td>
</tr>
<tr>
<td>– Person ⊆ HumanBeing</td>
<td>• :Woman rdfs:subClassOf :Person .</td>
</tr>
<tr>
<td>• hasWife(john,mary)</td>
<td>• :john :hasWife :mary .</td>
</tr>
<tr>
<td>• hasWife ⊆ hasSpouse</td>
<td>• :hasWife rdfs:subPropertyOf :hasSpouse .</td>
</tr>
<tr>
<td>– hasSpouse ⊆ marriedWith</td>
<td>• :hasWife rdfs:subPropertyOf :hasSpouse .</td>
</tr>
</tbody>
</table>
Special classes and properties

• **owl:Thing** (RDF syntax)
  – DL-syntax: $\top$
  – contains everything
• **owl:Nothing** (RDF syntax)
  – DL-syntax: $\bot$
  – empty class
• **owl:topProperty** (RDF syntax)
  – DL-syntax: $U$
  – every pair is in $U$
• **owl:bottomProperty** (RDF syntax)
  – empty property
Class constructors

• conjunction
  – Mother ≡ Woman ∩ Parent
  – :Mother owl:equivalentClass _:x .
    _:x rdf:type owl:Class .
    _:x owl:intersectionOf ( :Woman :Parent ) .

• disjunction
  – Parent ≡ Mother ∪ Father
  – :Parent owl:equivalentClass _:x .
    _:x rdf:type owl:Class .
    _:x owl:unionOf ( :Mother :Father ) .

• negation
  – ChildlessPerson ≡ Person ∩ ¬Parent
  – :ChildlessPerson owl:equivalentClass _:x .
    _:x rdf:type owl:Class .
    _:x owl:intersectionOf ( :Person _:y ) .
    _:y owl:complementOf :Parent .

∀x (Mother(x) ↔ Woman(x) ∧ Parent(x))

∀x (Parent(x) ↔ Mother(x) ∨ Father(x))

∀x (ChildlessPerson(x) ↔ Person(x) ∧ ¬Parent(x))
Class constructors

- **existential quantification**
  - only to be used with a role – also called a *property restriction*
  - \( \text{Parent} \equiv \exists \text{hasChild\.Person} \)
  - \( \text{:_x rdf:type owl:Restriction} . \)
  - \( \text{:_x owl:onProperty :hasChild} . \)
  - \( \text{:_x owl:someValuesFrom :Person} . \)

- **universal quantification**
  - only to be used with a role – also called a *property restriction*
  - \( \text{Person \cap Happy} \equiv \forall \text{hasChild\.Happy} \)
  - \( \text{:_x rdf:type owl:Class} . \)
  - \( \text{:_x owl:intersectionOf ( :Person :Happy )} . \)
  - \( \text{:_x owl:equivalentClass :_y} . \)
  - \( \text{:_y rdf:type owl:Restriction} . \)
  - \( \text{:_y owl:onProperty :hasChild} . \)
  - \( \text{:_y owl:allValuesFrom :Happy} . \)

- **Class constructors can be nested arbitrarily**
Contents

• OWL – Basic Ideas
• OWL as the Description Logic SROIQ(D)
• Different Perspectives on OWL
• OWL Semantics
• OWL Profiles
• Proof Theory
• Tools
The description logic ALC

- **ABox expressions:**
  - Individual assignments: Father(john)
  - Property assignments: hasWife(john, mary)

- **TBox expressions:**
  - Subclass relationships: \( \sqsubseteq \)
  - Conjunction: \( \sqcap \)
  - Disjunction: \( \sqcup \)
  - Negation: \( \neg \)
  - Also: \( \top, \bot \)
  - Property restrictions: \( \forall \), \( \exists \)

Complexity: ExpTime
Understanding SROIQ(D)

ALC + role chains = SR

- hasParent o hasBrother ⊆ hasUncle

\[ \forall x \forall y (\exists z ((\text{hasParent}(x,z) \land \text{hasBrother}(z,y)) \rightarrow \text{hasUncle}(x,y))) \]

- includes top property and bottom property

- includes S = ALC + transitivity
  - hasAncestor o hasAncestor ⊆ hasAncestor

- includes SH = S + role hierarchies
  - hasFather ⊆ hasParent
Understanding SROIQ(D)

- **O** – nominals (closed classes)
  - MyBirthdayGuests ≡ \{bill, john, mary\}
  - Note the difference to
    MyBirthdayGuests(bill)
    MyBirthdayGuests(john)
    MyBirthdayGuests(mary)

- Individual equality and inequality (no unique name assumption!)
  - bill = john
    - \{bill\} ≡ \{john\}
  - bill ≠ john
    - \{bill\} ∩ \{john\} ≡ ⊥
Understanding SROIQ(D)

- **I** – inverse roles
  - hasParent $\equiv$ hasChild$^-$
  - Orphan $\equiv$ $\forall$ hasChild$^-$.Dead

- **Q** – qualified cardinality restrictions
  - $\leq 4$ hasChild.Parent(john)
  - HappyFather $\equiv$ $\geq 2$ hasChild.Female
  - Car $\square = 4$ hasTyre.⊤

- **Complexity**
  - SHIQ, SHOQ, SHIO: ExpTime.
  - SHOIQ: NExpTime
  - SROIQ: N2ExpTime
Understanding SROIQ(D)

Properties can be declared to be

- Transitive \ hasAncestor
- Symmetric \ hasSpouse
- Asymmetric \ hasChild
- Reflexive \ hasRelative
- Irreflexive \ parentOf
- Functional \ hasHusband
- InverseFunctional \ hasHusband

called *property characteristics*
Understanding SROIQ(D)

(D) – datatypes

• so far, we have only seen properties with individuals in second argument, called *object properties* or *abstract roles* (DL)

• properties with datatype literals in second argument are called *data properties* or *concrete roles* (DL)

• allowed are many XML Schema datatypes, including `xsd:integer`, `xsd:string`, `xsd:float`, `xsd:boolean`, `xsd:anyURI`, `xsd:dateTime`

and also e.g. `owl:real`
Understanding SROIQ(D)

(D) – datatypes

- hasAge(john, "51"^^xsd:integer)

- additional use of constraining facets (from XML Schema)
  - e.g. Teenager \equiv \text{Person} \sqcap \exists \text{hasAge.} (\text{xsd:integer}: \geq 12 \text{ and } \leq 19)
  
  note: this is not standard DL notation!
Understanding SROIQ(D)

further expressive features

• Self
  – PersonCommittingSuicide $\equiv \exists \text{kills}.\text{Self}$

• Keys (not really in SROIQ(D), but in OWL)
  – set of (object or data) properties whose values uniquely identify an object

• disjoint properties
  – Disjoint(hasParent,hasChild)

• explicit anonymous individuals
  – as in RDF: can be used instead of named individuals
SROIQ(D) constructors – overview

- ABox assignments of individuals to classes or properties
- ALC: \( \subseteq, \equiv \) for classes
  \( \cap, \cup, \neg, \exists, \forall \)
  \( \top, \bot \)
- SR: + property chains, property characteristics, role hierarchies \( \subseteq \)
- SRO: + nominals \( \{o\} \)
- SROI: + inverse properties
- SROIQ: + qualified cardinality constraints
- SROIQ(D): + datatypes (including facets)
- + top and bottom roles (for objects and datatypes)
- + disjoint properties
- + Self
- + Keys (not in SROIQ(D), but in OWL)
This applies to the non-DL syntaxes (e.g. RDF syntax).

- **disjoint classes**
  - Apple ∩ Pear ⊆ ⊥

- **disjoint union**
  - Parent ⊑ Mother ⊔ Father
    Mother ∩ Father ⊆ ⊥

- **negative property assignments** (also for datatypes)
  - ¬hasAge(jack, "53"^^xsd:integer)
Contents

• OWL – Basic Ideas
• OWL As the Description Logic SROIQ(D)
• **Different Perspectives on OWL**
• OWL Semantics
• OWL Profiles
• Proof Theory
• Tools
OWL – Extralogical Features

• OWL ontologies have URIs and can be referenced by others via
  – import statements
• Namespace declarations
• Entity declarations (must be done)
• Versioning information etc.

• Annotations
  – Entities and axioms (statements) can be endowed with annotations, e.g. using rdfs:comment.
  – OWL syntax provides *annotation properties* for this purpose.
The modal logic perspective

• Description logics can be understood from a modal logic perspective.

• Each pair of $\forall R$ and $\exists R$ statements give rise to a pair of modalities.

• Essentially, some description logics are multi-modal logics.

• See e.g. Baader et al., The Description Logic Handbook, Cambridge University Press, 2007.
The RDFS perspective

- :mary rdf:type :Person.
- :Mother rdfs:subClassOf :Woman.
- :john :hasWife :Mary.
- :hasWife rdfs:subPropertyOf :hasSpouse
- :hasWife rdfs:range :Woman.
- :hasWife rdfs:domain :Man.
- Person(mary)
- Mother ⊆ Woman
- hasWife(john,mary)
- hasWife ⊆ hasSpouse

RDFS also allows to

- make statements about statements → only possible through annotations in OWL
- mix class names, individual names, property names (they are all URIs) → punning in OWL

RDFS semantics is weaker

\[ T \subseteq \forall \text{hasWife.Woman} \]
\[ T \subseteq \forall \text{hasWife^-}.\text{Man} \] or
\[ \exists \text{hasWife}. T \subseteq \text{Man} \]
Punning

- Description logics impose *type separation*, i.e. names of individuals, classes, and properties must be disjoint.

- In OWL 2 Full, type separation does not apply.

- In OWL 2 DL, type separation is relaxed, but a class \( X \) and an individual \( X \) are interpreted semantically as if they were different.

- Father(john)
  SocialRole(Father)

- See further below on the two different semantics for OWL.
Contents

• OWL – Basic Ideas
• OWL As the Description Logic SROIQ(D)
• Different Perspectives on OWL
• OWL Semantics
• OWL Profiles
• Proof Theory
• Tools
There are two semantics for OWL.

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

2. RDF-based Semantics
   No syntax restrictions apply.
   Extends the direct semantics with RDFS-reasoning features.

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

To obtain decidability, syntactic restrictions apply.

- Type separation / punning
- No cycles in property chains.
- No transitive properties in cardinality restrictions.
OWL Direct Semantics: Restrictions

- arbitrary property chain axioms lead to undecidability
- restriction: set of property chain axioms has to be *regular*
  - there must be a strict linear order $\prec$ on the properties
  - every property chain axiom has to have one of the following forms:
    
    \[
    \begin{align*}
    R \circ R & \subseteq R \\
    S^- & \subseteq R \\
    S_1 \circ S_2 \circ \ldots \circ S_n & \subseteq R \\
    R \circ S_1 \circ S_2 \circ \ldots \circ S_n & \subseteq R \\
    S_1 \circ S_2 \circ \ldots \circ S_n \circ R & \subseteq R
    \end{align*}
    \]
  - thereby, $S_i \prec R$ for all $i = 1, 2, \ldots, n$.

- Example 1: $R \circ S \subseteq R$  
  $S \circ S \subseteq S$  
  $R \circ S \circ R \subseteq T$  
  $\Rightarrow$ regular with order $S \prec R \prec T$

- Example 2: $R \circ T \circ S \subseteq T$  
  $\Rightarrow$ not regular because form not admissible

- Example 3: $R \circ S \subseteq S$  
  $S \circ R \subseteq R$  
  $\Rightarrow$ not regular because no adequate order exists
combining property chain axioms and cardinality constraints may lead to undecidability

restriction: use only *simple* properties in cardinality expressions (i.e. those which cannot be – directly or indirectly – inferred from property chains)

technically:

- for any property chain axiom $S_1 \circ S_2 \circ \ldots \circ S_n \sqsubseteq R$ with $n > 1$, $R$ is non-simple
- for any subproperty axiom $S \sqsubseteq R$ with $S$ non-simple, $R$ is non-simple
- all other properties are simple

Example: $Q \circ P \sqsubseteq R$  $R \circ P \sqsubseteq R$  $R \sqsubseteq S$  $P \sqsubseteq R$  $Q \sqsubseteq S$

non-simple: $R, S$  simple: $P, Q$
OWL 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

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

then we could set

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

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

- Mapping is extended to complex class expressions:
  - $T^I = \Delta^I$
  - $\bot^I = \emptyset$
  - $(C \cap D)^I = C^I \cap D^I$
  - $(C \cup D)^I = C^I \cup D^I$
  - $(\neg C)^I = \Delta^I \setminus C^I$
  - $(\forall R.C)^I = \{ x \mid \text{for all } (x,y) \in R^I \text{ we have } y \in C^I \}$
  - $(\exists R.C)^I = \{ x \mid \text{there is } (x,y) \in R^I \text{ with } y \in C^I \}$
  - $(\geq n R.C)^I = \{ x \mid \#\{ y \mid (x,y) \in R^I \text{ and } y \in C^I \} \geq n \}$
  - $(\leq n R.C)^I = \{ x \mid \#\{ y \mid (x,y) \in R^I \text{ and } y \in C^I \} \leq n \}$

- ...And to role expressions:
  - $U^I = \Delta^I \times \Delta^I$
  - $(R^-)^I = \{ (y,x) \mid (x,y) \in R^I \}$

- ...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 $R^I \cap S^I = \emptyset$
  - $S_1 \circ S_2 \circ \ldots \circ S_n \subseteq R$ holds if $S_1^I \circ S_2^I \circ \ldots \circ S_n^I \subseteq R^I$
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 via models (below).

- ...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 $R^I \cap S^I = \emptyset$  
  - $S_1 \circ S_2 \circ \ldots \circ S_n \subseteq R$ holds if $S_1^I \circ S_2^I \circ \ldots \circ S_n^I \subseteq R^I$
A model for an OWL KB is such a mapping \( I \) which satisfies all axioms in the KB.

An axiom \( \alpha \) is a logical consequence of a KB if every model of the KB is also a model of \( \alpha \).

The logical consequences of a KB are all those things which are necessarily the case in all “realities” in which the KB is the case.
Notion of logical consequence
Not a model!

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.
A model

Professor \sqsubseteq FacultyMember
Professor(rudiStuder)
hasAffiliation(rudiStuder, aifb)

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

\[ I_I(rudiStuder) = r \]
\[ I_I(aifb) = a \]
\[ I_C(Professor) = \{r\} \]
\[ I_C(FacultyMember) = \{r, s\} \]
\[ I_R(hasAffiliation) = \{(r, a)\} \]
## Models

Professor \sqsupseteq FacultyMember

Professor(rudiStuder)

hasAffiliation(rudiStuder, aifb)

<table>
<thead>
<tr>
<th>Δ</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>I_I(rudiStuder)</td>
<td>{a, r, s}</td>
<td>{1, 2}</td>
<td>{\spadesuit}</td>
</tr>
<tr>
<td>I_I(aifb)</td>
<td>r</td>
<td>1</td>
<td>\spadesuit</td>
</tr>
<tr>
<td>I_C(Professor)</td>
<td>a</td>
<td>2</td>
<td>\spadesuit</td>
</tr>
<tr>
<td>I_C(FacultyMember)</td>
<td>{r}</td>
<td>{1}</td>
<td>{\spadesuit}</td>
</tr>
<tr>
<td>I_R(hasAffiliation)</td>
<td>{(r, a)}</td>
<td>{(1, 1), (1, 2)}</td>
<td>{(\spadesuit, \spadesuit)}</td>
</tr>
</tbody>
</table>

Is FacultyMember(aifb) a logical consequence?
Counterexample

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

$$\Delta = \{a, r\}$$

$$I_I(\text{rudiStuder}) = r$$

$$I_I(aifb) = a$$

$$I_C(\text{Professor}) = \{r\}$$

$$I_C(\text{FacultyMember}) = \{r\}$$

$$I_R(\text{hasAffiliation}) = \{(r, a)\}$$
**Logical Consequence**

Professor $\sqsubseteq$ FacultyMember

Professor(rudiStuder)

hasAffiliation(rudiStuder, aifb)

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>${a, r, s}$</td>
<td>${1, 2}$</td>
<td></td>
<td>${\spadesuit}$</td>
</tr>
<tr>
<td>$I_I$rudiStuder</td>
<td>$r$</td>
<td>1</td>
<td>$\spadesuit$</td>
</tr>
<tr>
<td>$I_I$aifb</td>
<td>$a$</td>
<td>2</td>
<td>$\spadesuit$</td>
</tr>
<tr>
<td>$I_C$Professor</td>
<td>${r}$</td>
<td>${1}$</td>
<td>${\spadesuit}$</td>
</tr>
<tr>
<td>$I_C$FacultyMember</td>
<td>${a, r, s}$</td>
<td>${1, 2}$</td>
<td>${\spadesuit}$</td>
</tr>
<tr>
<td>$I_R$hasAffiliation</td>
<td>${(r, a)}$</td>
<td>${(1, 1), (1, 2)}$</td>
<td>${(\spadesuit, \spadesuit)}$</td>
</tr>
</tbody>
</table>

Is FacultyMember(rudiStuder) a logical consequence?
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

...which (interpreted under FOL semantics) coincides with the definition just given.
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.

Unicorn \( \subseteq \) Fictitious

Unicorn \( \subseteq \) Animal

Fictitious \( \sqsubset \) Animal \( \sqsubseteq \) \( \bot \)
A Semantic Puzzle


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

**Fig. 1.** A justification for Person $\sqsubseteq \bot$
What Semantics Is Good For

• Opinions Differ. Here’s my take.

• Semantic Web requires a shareable, declarative and **computable** semantics.
• I.e., the semantics must be a formal entity which is clearly defined and automatically computable.

• Ontology languages provide this by means of their formal semantics.
• Semantic Web Semantics is given by a relation – the **logical consequence** relation.

• Note: This is considerably more than saying that the semantics of an ontology is the set of its logical consequences!
In other words

We capture the meaning of information

not by specifying its meaning (which is impossible)
but by specifying

how information interacts with other information.

We describe the meaning indirectly through its effects.
Simple Logical Reasoning

If I ask for soccer team members, I also want to get the goalkeepers listed ...

If I ask for cities, I also want all capitals listed ...

inheritance reasoning
Less Simple Reasoning

What was again the name of that Russian researcher who worked on resolution-based calculi for EL?

Are lobsters spiders?

What is "Käuzchen" in English?

Answering requires merging of knowledge from many websites and using background knowledge.
SNOMED CT

- SNOMED CT: commercial ontology, medical domain
ca. 300,000 axioms

- InjuryOfFinger \equiv \text{Injury} \sqcap \exists \text{site.Finger}_S
  InjuryOfHand \equiv \text{Injury} \sqcap \exists \text{site.Hand}_S
  \text{Finger}_S \sqsubseteq \text{Hand}_P
  \text{Hand}_P \sqsubseteq \text{Hand}_S \sqcap \exists \text{part.Hand}_E

- Reasoning has been used e.g. for
  - classification (computing the hidden taxonomy)
    e.g., InjuryOfFinger \sqsubseteq InjuryOfHand
  - bug finding
Contents

• OWL – Basic Ideas
• OWL As the Description Logic SROIQ(D)
• Different Perspectives on OWL
• OWL Semantics
• OWL Profiles
• Proof Theory
• Tools
OWL Profiles

- OWL Full – using the RDFS-based semantics
- OWL DL – using the FOL semantics

The OWL 2 documents describe further profiles, which are of polynomial complexity:

- OWL EL (EL++)
- OWL QL (DL Lite_R)
- OWL RL (DLP)
OWL 2 EL

- **allowed:**
  - subclass axioms with intersection, existential quantification, top, bottom
    - closed classes must have only one member
  - property chain axioms, range restrictions (under certain conditions)
- **disallowed:**
  - negation, disjunction, arbitrary universal quantification, role inverses

\[
\neg \exists U \subseteq \text{NEW}
\]

- **Examples:**
  - Human \(\sqsubseteq \exists\text{hasParent.P}erson\)
  - \(\exists\text{married.}\top \sqcap \text{CatholicPriest} \sqsubseteq \bot;\)
  - \(\text{hasParent} \circ \text{hasParent} \sqsubseteq \text{hasGrandparent}\)
Motivated by the question: what fraction of OWL 2 DL can be expressed \textit{naively} by rules (with equality)?

Examples:
- $\exists\text{parentOf.} \exists\text{parentOf.} \top \sqsubseteq \text{Grandfather}$
  
  \text{rule version: } \text{parentOf}(x,y) \ \text{parentOf}(y,z) \rightarrow \text{Grandfather}(x)$
- $\text{Orphan} \sqsubseteq \forall \text{hasParent.} \text{Dead}$
  \text{rule version: } \text{Orphan}(x) \ \text{hasParent}(x,y) \rightarrow \text{Dead}(y)$
- $\text{Monogamous} \sqsubseteq \leq 1 \text{married.} \text{Alive}$
  \text{rule version: } \text{Monogamous}(x) \ \text{married}(x,y) \ \text{Alive}(y) \ \text{married}(x,z) \ \text{Alive}(z) \rightarrow y = z$
- $\text{childOf} \circ \text{childOf} \sqsubseteq \text{grandchildOf}$
  \text{rule version: } \text{childOf}(x,y) \ \text{childOf}(y,z) \rightarrow \text{grandchildOf}(x,z)$
- $\text{Disj(childOf, parentOf)}$
  \text{rule version: } \text{childOf}(x,y) \ \text{parentOf}(x,y) \rightarrow \text{Disj}$
• Syntactic characterization:
  – essentially, all axiom types are allowed
  – disallow certain constructors on lhs and rhs of subclass statements
  – cardinality restrictions: only on rhs and only $\leq 1$ and $\leq 0$ allowed
  – closed classes: only with one member

• Reasoner conformance requires only soundness.
• Motivated by the question: what fraction of OWL 2 DL can be captured by standard database technology?

• Formally: query answering LOGSPACE w.r.t. data (via translation into SQL)

• Allowed:
  – subproperties, domain, range
  – subclass statements with
    • left hand side: class name or expression of type $\exists r. \top$
    • right hand side: intersection of class names, expressions of type $\exists r.C$ and negations of lhs expressions
    • no closed classes!

• Example:
  $\exists_{married. \top} \subseteq \neg_{Free} \land \exists_{has.\text{Sorrow}}$
Contents

• OWL – Basic Ideas
• OWL As the Description Logic SROIQ(D)
• Different Perspectives on OWL
• OWL Semantics
• OWL Profiles
• Proof Theory
• Tools
A is a logical consequence of K written $K \models A$

if and only if

every model of K is a model of A.

- To show an entailment, we need to check all models?
- But that‘s infinitely many!!!
A Reasoning Problem

We need algorithms which do not apply the model-based definition of logical consequence in a naive manner.

These algorithms should be syntax-based. (Computers can only do syntax manipulations.)

Luckily, such algorithms exist!

However, their correctness (soundness and completeness) needs to be proven formally. Which is often a non-trivial problem requiring substantial mathematical build-up.

We won’t do the proofs here.
Proof Theory

We will show the Tableaux Method – implemented, e.g., in Pellet and Racer.

Alternatives:

• Transformation to disjunctive datalog using basic superposition done for SHIQ
• Naive mapping to Datalog for OWL RL
• Mapping to SQL for OWL QL
• Special-purpose algorithms for OWL EL e.g. transformation to Datalog
Proof theory Via Tableaux

• Adaptation of FOL tableaux algorithms.

• Problem: OWL is decidable, but FOL tableaux algorithms do not guarantee termination.

• Solution: blocking.
Contents

- Important inference problems
- Tableaux algorithm for ALC
- Tableaux algorithm for SHIQ
Important Inference Problems

- Global consistency of a knowledge base. \( KB \models \text{false?} \)
  - Is the knowledge base meaningful?
- Class consistency \( C \equiv \bot? \)
  - Is \( C \) necessarily empty?
- Class inclusion (Subsumption) \( C \sqsubseteq D? \)
  - Structuring knowledge bases
- Class equivalence \( C \equiv D? \)
  - Are two classes in fact the same class?
- Class disjointness \( C \cap D = \bot? \)
  - Do they have common members?
- Class membership \( C(a)? \)
  - Is \( a \) contained in \( C \)?
- Instance Retrieval „find all \( x \) with \( C(x) \)“
  - Find all (known!) individuals belonging to a given class.
Reduction to Unsatisfiability

- Global consistency of a knowledge base.
  - Failure to find a model.
- Class consistency
  - $KB \cup \{C(a)\}$ unsatisfiable
- Class inclusion (Subsumption)
  - $KB \cup \{C \cap \neg D(a)\}$ unsatisfiable (a new)
- Class equivalence
  - $C \sqsubseteq D$ and $D \sqsubseteq C$
- Class disjointness
  - $KB \cup \{(C \cap D)(a)\}$ unsatisfiable (a new)
- Class membership
  - $KB \cup \{\neg C(a)\}$ unsatisfiable
- Instance Retrieval „find all x with C(x)“
  - Check class membership for all individuals.
Reduction to Satisfiability

• We will present so-called tableaux algorithms.

• They attempt to construct a model of the knowledge base in a „general, abstract“ manner.
  – If the construction fails, then (provably) there is no model – i.e. the knowledge base is unsatisfiable.
  – If the construction works, then it is satisfiable.

→ Hence the reduction of all inference problems to the checking of unsatisfiability of the knowledge base!
Contents

• Important inference problems
• Tableaux algorithm for ALC
• Tableaux algorithm for SHIQ
ALC tableaux: contents

- Transformation to negation normal form
- Naive tableaux algorithm
- Tableaux algorithm with blocking
Transform. to negation normal form

Given a knowledge base $K$.

- Replace $C \equiv D$ by $C \subseteq D$ and $D \subseteq C$.
- Replace $C \subseteq D$ by $\neg C \sqcup D$.
- Apply the equations on the next slide exhaustively.

Resulting knowledge base: $\text{NNF}(K)$

*Negation normal form* of $K$.

Negation occurs only directly in front of atomic classes.
\[
\begin{align*}
\text{NNF}(C) &= C \quad \text{if } C \text{ is a class name} \\
\text{NNF}(\neg C) &= \neg C \quad \text{if } C \text{ is a class name} \\
\text{NNF}(\neg \neg C) &= \text{NNF}(C) \\
\text{NNF}(C \sqcup D) &= \text{NNF}(C) \sqcup \text{NNF}(D) \\
\text{NNF}(C \sqcap D) &= \text{NNF}(C) \sqcap \text{NNF}(D) \\
\text{NNF}(\neg (C \sqcup D)) &= \text{NNF}(\neg C) \sqcap \text{NNF}(\neg D) \\
\text{NNF}(\neg (C \sqcap D)) &= \text{NNF}(\neg C) \sqcup \text{NNF}(\neg D) \\
\text{NNF}(\forall R.C) &= \forall R.\text{NNF}(C) \\
\text{NNF}(\exists R.C) &= \exists R.\text{NNF}(C) \\
\text{NNF}(\neg \forall R.C) &= \exists R.\text{NNF}(\neg C) \\
\text{NNF}(\neg \exists R.C) &= \forall R.\text{NNF}(\neg C)
\end{align*}
\]

K and NNF(K) have the same models (are \textit{logically equivalent}).
Example

\[ P \subseteq (E \cap U) \cup \neg(\neg E \cup D). \]

In negation normal form:

\[ \neg P \cup (E \cap U) \cup (E \cap \neg D). \]
ALC tableaux: contents

- Transformation to negation normal form
- Naive tableaux algorithm
- Tableaux algorithm with blocking
Reduction to (un)satisfiability.

Idea:
• Given knowledge base K
• Attempt construction of a tree (called Tableau), which represents a model of K.
  (It’s actually rather a Forest.)
• If attempt fails, K is unsatisfiable.
The Tableau

- Nodes represent elements of the domain of the model
  - Every node $x$ is labeled with a set $L(x)$ of class expressions. $C \in L(x)$ means: "$x$ is in the extension of $C$"

- Edges stand for role relationships:
  - Every edge $<x,y>$ is labeled with a set $L(<x,y>)$ of role names. $R \in L(<x,y>)$ means: "$(x,y)$ is in the extension of $R$"
Simple example

C(a)
C ⊆ ∃R.D
D ⊆ E

Does this entail (∃R.E)(a)?

(add ∀R.¬E(a)
and show unsatisfiability)

Contradiction!
Another example

C(a)
C ⊆ ∃R.D
D ⊆ E ∪ F
F ⊆ E

Does this entail (∃R.E)(a)?

(add ∀R.¬E(a) and show unsatisfiability)

D
¬E (because ∀R.¬E(a))
choice: (D ⊆ E ∪ F):
1. E (contradiction!)
2. F
   E (contradiction!)
Formal Definition

- Input: $K = \text{TBox} + \text{ABox}$ (in NNF)
- Output: Whether or not $K$ is satisfiable.

- A tableau is a directed labeled graph
  - nodes are individuals or (new) variable names
  - nodes $x$ are labeled with sets $L(x)$ of classes
  - edges $<x,y>$ are labeled with sets $L(<x,y>)$ of role names
Initialisation

- Make a node for every individual in the ABox.
- Every node is labeled with the corresponding class names from the ABox.
- There is an edge, labeled with R, between a and b, if R(a,b) is in the ABox.

- (If there is no ABox, the initial tableau consists of a node x with empty label.)
Example initialisation

Human $\subseteq \exists$hasParent.Human
Orphan $\subseteq$ Human $\cap \neg \exists$hasParent.Alive
Orphan(harrypotter)
hasParent(harrypotter,jamespotter)
Careful: need NNF!

\neg \text{Human} \cup \exists \text{hasParent.Human}

\neg \text{Orphan} \cup (\text{Human} \cap \forall \text{hasParent.} \neg \text{Alive})

\text{Orphan(harrypotter)}

\text{hasParent(harrypotter, jamespotter)}
Constructing the tableau

• Non-deterministically extend the tableau using the rules on the next slide.

• Terminate, if
  – there is a contradiction in a node label (i.e., it contains classes C and ¬C, or it contains \( \bot \)), or
  – none of the rules is applicable.

• If the tableau does not contain a contradiction, then the knowledge base is satisfiable.
  Or more precisely: If you can make a choice of rule applications such that no contradiction occurs and the process terminates, then the knowledge base is satisfiable.
Naive ALC tableaux rules

\( \square \)-rule: If \( C \sqcap D \in \mathcal{L}(x) \) and \( \{C, D\} \not\subseteq \mathcal{L}(x) \), then set \( \mathcal{L}(x) \leftarrow \{C, D\} \).

\( \square \)-rule: If \( C \sqcup D \in \mathcal{L}(x) \) and \( \{C, D\} \cap \mathcal{L}(x) = \emptyset \), then set \( \mathcal{L}(x) \leftarrow C \) or \( \mathcal{L}(x) \leftarrow D \).

\( \exists \)-rule: If \( \exists R.C \in \mathcal{L}(x) \) and there is no \( y \) with \( R \in \mathcal{L}(x,y) \) and \( C \in \mathcal{L}(y) \), then

1. add a new node with label \( y \) (where \( y \) is a new node label),
2. set \( \mathcal{L}(x,y) = \{R\} \), and
3. set \( \mathcal{L}(y) = \{C\} \).

\( \forall \)-rule: If \( \forall R.C \in \mathcal{L}(x) \) and there is a node \( y \) with \( R \in \mathcal{L}(x,y) \) and \( C \not\in \mathcal{L}(y) \), then set \( \mathcal{L}(y) \leftarrow C \).

TBox-rule: If \( C \) is a TBox statement and \( C \not\in \mathcal{L}(x) \), then set \( \mathcal{L}(x) \leftarrow C \).
Example

\( \neg \text{Human} \sqcup \exists \text{hasParent}.\text{Human} \)

\( \neg \text{Orphan} \sqcup (\text{Human} \sqcap \forall \text{hasParent}.\neg \text{Alive}) \)

Orphan(harrypotter)

hasParent(harrypotter,jamespotter)

\( \neg \text{Alive}(\text{jamespotter}) \)

i.e. add: \( \text{Alive}(\text{jamespotter}) \)

and search for contradiction

1. \( \neg \text{Orphan} \) (contradiction)
2. \( \text{Human} \sqcap \forall \text{hasParent}.\neg \text{Alive} \)
   
   \( \text{Human} \)
   
   \( \forall \text{hasParent}.\neg \text{Alive} \)

2. \( \neg \text{Alive} \) (contradiction)
ALC tableaux: contents

• Transformation to negation normal form
• Naive tableaux algorithm
• Tableaux algorithm with blocking
There’s a termination problem

TBox: $\exists R. T$
ABox: $T(a_1)$

• Obviously satisfiable:
  Model $M$ with domain elements $a_1^M, a_2^M, ...$
  and $R^M(a_i^M, a_{i+1}^M)$ for all $i \geq 1$
• but tableaux algorithm does not terminate!
Solution?

Actually, things repeat!
Idea: it is not necessary to expand x, since it’s simply a copy of a.

⇒ Blocking

```
 a R x R y R
 T E_R.T T E_R.T T E_R.T
```
Blocking

• $x$ is *blocked* (by $y$) if
  – $x$ is not an individual (but a variable)
  – $y$ is a predecessor of $x$ and $L(x) \subseteq L(y)$
  – or a predecessor of $x$ is blocked

Here, $x$ is blocked by $a$. 
Constructing the tableau

• Non-deterministically extend the tableau using the rules on the next slide, but only apply a rule if x is not blocked!

• Terminate, if
  – there is a contradiction in a node label (i.e., it contains classes C and \( \neg C \)), or
  – none of the rules is applicable.

• If the tableau does not contain a contradiction, then the knowledge base is satisfiable.  
  Or more precisely: If you can make a choice of rule applications such that no contradiction occurs and the process terminates, then the knowledge base is satisfiable.
Naive ALC tableaux rules

\( \square \)-rule: If \( C \sqcap D \in \mathcal{L}(x) \) and \( \{C, D\} \not\subseteq \mathcal{L}(x) \), then set \( \mathcal{L}(x) \leftarrow \{C, D\} \).

\( \square \)-rule: If \( C \sqcap D \in \mathcal{L}(x) \) and \( \{C, D\} \cap \mathcal{L}(x) = \emptyset \), then set \( \mathcal{L}(x) \leftarrow C \) or \( \mathcal{L}(x) \leftarrow D \).

\( \exists \)-rule: If \( \exists R.C \in \mathcal{L}(x) \) and there is no \( y \) with \( R \in \mathcal{L}(x, y) \) and \( C \in \mathcal{L}(y) \), then

1. add a new node with label \( y \) (where \( y \) is a new node label),
2. set \( \mathcal{L}(x, y) = \{R\} \), and
3. set \( \mathcal{L}(y) = \{C\} \).

\( \forall \)-rule: If \( \forall R.C \in \mathcal{L}(x) \) and there is a node \( y \) with \( R \in \mathcal{L}(x, y) \) and \( C \not\in \mathcal{L}(y) \), then set \( \mathcal{L}(y) \leftarrow C \).

TBox-rule: If \( C \) is a TBox statement and \( C \not\in \mathcal{L}(x) \), then set \( \mathcal{L}(x) \leftarrow C \).

Apply only if \( x \) is not blocked!
Example (0)

- Knowledge base \{\text{Human} \subseteq \exists \text{hasParent.Human}, \text{Bird(tweety)}\}
- We want to show that Human(tweety) does \textit{not} hold, i.e. that \(\neg\text{Human(tweety)}\) is entailed.
- We will not be able to show this. I.e. Human(tweety) is \textit{possible}.

- Shorter notation:
  \[ H \subseteq \exists p. H \]
  \[ B(t) \]
  \(\neg H(t)\) entailed?
Example (0)

Knowledge base \{\neg H \cup \exists p.H, B(t), H(t)\}

1. \neg H (contradiction)
2. \exists p.H

expansion stops. Cannot find contradiction!
Example (0) the other case

Knowledge base \{\neg H \cup \exists p.H, B(t), \neg H(t)\}

\[
\begin{array}{c}
t \quad p \quad x \quad p \quad y \\
\end{array}
\]

\[
\begin{array}{c}
\neg H \\
B \\
\neg H \cup \exists p.H \\
1. \quad \neg H \text{ cannot be} \\
\quad \text{added. no expansion} \\
\quad \text{in this part} \\
2. \quad \exists p.H \\
\end{array}
\]

\[
\begin{array}{c}
2.: \\
H \\
\neg H \cup \exists p.H \\
2.1: \quad \neg H \text{ (contradiction)} \\
2.2: \quad \exists p.H \\
\end{array}
\]

2.2: \quad \exists p.H \\
\text{blocked by } x

no further expansion possible – knowledge base is satisfiable!
Example(1)

Show, that

Professor $\subseteq$ (Person $\cap$ Universitymember) $\cup$ (Person $\cap$ $\neg$PhDstudent)

entails that every Professor is a Person.

Find contradiction in:

- $\neg P \cup (E \cap U) \cup (E \cap \neg S)$
- $P \cap \neg E(x)$

<table>
<thead>
<tr>
<th>$P \cap \neg E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$</td>
</tr>
<tr>
<td>$\neg E$</td>
</tr>
<tr>
<td>$\neg P \cup (E \cap U) \cup (E \cap \neg S)$</td>
</tr>
<tr>
<td>1. $\neg P$ (contradiction)</td>
</tr>
<tr>
<td>2. $(E \cap U) \cup (E \cap \neg S)$</td>
</tr>
<tr>
<td>1. $E \cap U$</td>
</tr>
<tr>
<td>$E$ (contradiction)</td>
</tr>
<tr>
<td>2. $E \cap \neg S$</td>
</tr>
<tr>
<td>$E$ (contradiction)</td>
</tr>
</tbody>
</table>
Example (2)

Show that

\[
\begin{align*}
&\text{hasChild}(\text{john}, \text{peter}) \\
&\text{hasChild}(\text{john}, \text{paul}) \\
&\text{male}(\text{peter}) \\
&\text{male}(\text{paul})
\end{align*}
\]

does not entail \(\forall\text{hasChild}.\text{male}(\text{john})\).
Example (3)

Show that the knowledge base

Bird ⊆ Flies
Penguin ⊆ Bird
Penguin ∩ Flies ⊆ ⊥
Penguin(tweety)

is unsatisfiable.

TBox:
¬B ∪ F
¬P ∪ B
¬P ∪ ¬F ∪ ⊥

P
¬P ∪ B
¬B ∪ F
¬P ∪ ¬F
1. ¬P (contradiction)
2. B
   1. ¬B (contradiction)
   2. F
      1. ¬P (contradiction)
      2. ¬F (contradiction)
Show that the knowledge base

\[ \begin{align*}
C(a) & \quad \quad C(c) \\
R(a,b) & \quad \quad R(a,c) \\
S(a,a) & \quad \quad S(c,b) \\
C & \subseteq \forall S.A \\
A & \subseteq \exists R.\exists S.A \\
A & \subseteq \exists R.C
\end{align*} \]

entails \( \exists R.\exists R.\exists S.A(a) \).
Example (4)

\[ \neg \exists R. \exists R. \exists S. a \equiv \forall R. \forall R. \forall S. \neg A \]

TBox:
\[ \neg C \sqcup \forall S. A \]
\[ \neg A \sqcup \exists R. \exists S. A \]
\[ \neg A \sqcup \exists R. C \]

Diagram:
- Node a connected by R to b and S to c
- Node b connected by R to x
- Node x connected by S to y
Contents

• Important inference problems
• Tableaux algorithm for ALC
• Tableaux algorithm for SHIQ
Tableaux Algorithm for SHIQ

• Basic idea is the same.

• Blocking rule is more complicated

• Other modifications are also needed.
Transform. to negation normal form

Given a knowledge base K.

- Replace $C \equiv D$ by $C \subseteq D$ and $D \subseteq C$.
- Replace $C \subseteq D$ by $\neg C \sqcup D$.
- Apply the equations on the next slide exhaustively.

Resulting knowledge base: NNF(K)

*Negation normal form* of K.

Negation occurs only directly in front of atomic classes.
$$\text{NNF}(C) = C \quad \text{if } C \text{ is a class name}$$

$$\text{NNF}(\neg C) = \neg C \quad \text{if } C \text{ is a class name}$$

$$\text{NNF}(\neg \neg C) = \text{NNF}(C)$$

$$\text{NNF}(C \sqcup D) = \text{NNF}(C) \sqcup \text{NNF}(D)$$

$$\text{NNF}(C \sqcap D) = \text{NNF}(C) \sqcap \text{NNF}(D)$$

$$\text{NNF}(\neg(C \sqcup D)) = \text{NNF}(\neg C) \sqcap \text{NNF}(\neg D)$$

$$\text{NNF}(\neg(C \sqcap D)) = \text{NNF}(\neg C) \sqcup \text{NNF}(\neg D)$$

$$\text{NNF}(\forall R.C) = \forall R.\text{NNF}(C)$$

$$\text{NNF}(\exists R.C) = \exists R.\text{NNF}(C)$$

$$\text{NNF}(\neg \forall R.C) = \exists R.\text{NNF}(\neg C)$$

$$\text{NNF}(\neg \exists R.C) = \forall R.\text{NNF}(\neg C)$$

$$\text{NNF}(\leq n \ R.C) = \leq n \ R.\text{NNF}(C)$$

$$\text{NNF}(\geq n \ R.C) = \geq n \ R.\text{NNF}(C)$$

$$\text{NNF}(\neg \leq n \ R.C) = \geq (n+1)R.\text{NNF}(C)$$

$$\text{NNF}(\neg \geq n \ R.C) = \leq (n-1)R.\text{NNF}(C), \text{ where } \leq (-1)R.C = \bot$$

K and NNF(K) have the same models (are logically equivalent).
Formal Definition

• A tableau is a directed labeled graph
  – nodes are individuals or (new) variable names
  – nodes \( x \) are labeled with sets \( L(x) \) of classes
  – edges \( <x,y> \) are labeled
    • either with sets \( L(<x,y>) \) of role names or inverse role names
    • or with the symbol \( = \) (for equality)
    • or with the symbol \( \neq \) (for inequality)
Initialisation

- Make a node for every individual in the ABox. These nodes are called *root nodes*.
- Every node is labeled with the corresponding class names from the ABox.
- There is an edge, labeled with $R$, between $a$ and $b$, if $R(a,b)$ is in the ABox.
- There is an edge, labeled $\neq$, between $a$ and $b$ if $a \neq b$ is in the ABox.
- There are no $=$ relations (yet).
Notions

• We write $S^{-}$ as $S$.
• If $R \in L(<x,y>)$ and $R \sqsubseteq S$ (where $R,S$ can be inverse roles), then
  – $y$ is an $S$-successor of $x$ and
  – $x$ is an $S$-predecessor of $y$.
• If $y$ is an $S$-successor or an $S^{-}$-predecessor of $x$, then $y$ is an neighbor of $x$.
• Ancestor is the transitive closure of Predecessor.
• x is *blocked* by y if x, y are not root nodes and
  – the following hold: ["x is directly blocked"]
    • no ancestor of x is blocked
    • there are predecessors y', x' of x
    • y is a successor of y' and x is a successor of x'
    • L(x) = L(y) and L(x') = L(y')
    • L(<x',x>) = L(<y',y>)
  – or the following holds: ["x is indirectly blocked"]
    • an ancestor of x is blocked or
    • x is successor of some y with L(<y,x>) = ∅
Constructing the tableau

- Non-deterministically extend the tableau using the rules on the next slide.

- Terminate, if
  - there is a contradiction in a node label, i.e.,
    - it contains $\perp$ or classes $C$ and $\neg C$ or
    - it contains a class $\leq nR.C$ and
      x also has $(n+1)$ R-successors $y_i$ and $y_i \neq y_j$ (for all $i \neq j$)
  - or none of the rules is applicable.

- If the tableau does not contain a contradiction, then the knowledge base is satisfiable.
Or more precisely: If you can make a choice of rule applications such that no contradiction occurs and the process terminates, then the knowledge base is satisfiable.
SHIQ Tableaux Rules

\(\Box\)-rule: If \(x\) is not indirectly blocked, \(C \square D \in \mathcal{L}(x)\), and \(\{C, D\} \not\subset \mathcal{L}(x)\), then set \(\mathcal{L}(x) \leftarrow \{C, D\}\).

\(\Box\)-rule: If \(x\) is not indirectly blocked, \(C \sqcup D \in \mathcal{L}(x)\) and \(\{C, D\} \cap \mathcal{L}(x) = \emptyset\), then set \(\mathcal{L}(x) \leftarrow C\) or \(\mathcal{L}(x) \leftarrow D\).

\(\exists\)-rule: If \(x\) is not blocked, \(\exists R.C \in \mathcal{L}(x)\), and there is no \(y\) with \(R \in \mathcal{L}(x, y)\) and \(C \in \mathcal{L}(y)\), then

1. add a new node with label \(y\) (where \(y\) is a new node label),
2. set \(\mathcal{L}(x, y) = \{R\}\) and \(\mathcal{L}(y) = \{C\}\).

\(\forall\)-rule: If \(x\) is not indirectly blocked, \(\forall R.C \in \mathcal{L}(x)\), and there is a node \(y\) with \(R \in \mathcal{L}(x, y)\) and \(C \not\in \mathcal{L}(y)\), then set \(\mathcal{L}(y) \leftarrow C\).

TBox-rule: If \(x\) is not indirectly blocked, \(C\) is a TBox statement, and \(C \not\in \mathcal{L}(x)\), then set \(\mathcal{L}(x) \leftarrow C\).
**trans-rule:** If \( x \) is not indirectly blocked, \( \forall S.C \in \mathcal{L}(x) \), \( S \) has a transitive subrole \( R \), and \( x \) has an \( R \)-neighbor \( y \) with \( \forall R.C \not\in \mathcal{L}(y) \), then set \( \mathcal{L}(y) \leftarrow \forall R.C \).

**choose-rule:** If \( x \) is not indirectly blocked, \( \leq n S.C \in \mathcal{L}(x) \) or \( \geq n S.C \in \mathcal{L}(x) \), and there is an \( S \)-neighbor \( y \) of \( x \) with \( \{ C, \text{NNF}(\neg C) \} \cap \mathcal{L}(y) = \emptyset \), then set \( \mathcal{L}(y) \leftarrow C \) or \( \mathcal{L}(y) \leftarrow \text{NNF}(\neg C) \).

**\( \geq \)-rule:** If \( x \) is not blocked, \( \geq n S.C \in \mathcal{L}(x) \), and there are no \( n \) \( S \)-neighbors \( y_1, \ldots, y_n \) of \( x \) with \( C \in \mathcal{L}(y_i) \) and \( y_i \not\sim y_j \) for \( i, j \in \{1, \ldots, n\} \) and \( i \neq j \), then

1. create \( n \) new nodes with labels \( y_1, \ldots, y_n \) (where the labels are new),
2. set \( \mathcal{L}(x, y_i) = \{ S \} \), \( \mathcal{L}(y_i) = \{ C \} \), and \( y_i \not\sim y_j \) for all \( i, j \in \{1, \ldots, n\} \) with \( i \neq j \).
\textbf{\(\leq\)-rule:} If \(x\) is not indirectly blocked, \(\leq n S.C \in \mathcal{L}(x)\), there are more than \(n\) \(S\)-neighbors \(y_i\) of \(x\) with \(C \in \mathcal{L}(y_i)\), and \(x\) has two \(S\)-neighbors \(y, z\) such that \(y\) is neither a root node nor an ancestor of \(z\), \(y \not\approx z\) does not hold, and \(C \in \mathcal{L}(y) \cap \mathcal{L}(z)\), then

1. set \(\mathcal{L}(z) \leftarrow \mathcal{L}(y)\),
2. if \(z\) is an ancestor of \(x\), then \(\mathcal{L}(z, x) \leftarrow \{\text{Inv}(R) \mid R \in \mathcal{L}(x, y)\}\),
3. if \(z\) is not an ancestor of \(x\), then \(\mathcal{L}(x, z) \leftarrow \mathcal{L}(x, y)\),
4. set \(\mathcal{L}(x, y) = \emptyset\), and
5. set \(u \not\approx z\) for all \(u\) with \(u \not\approx y\).

\textbf{\(\leq\)-root-rule:} If \(\leq n S.C \in \mathcal{L}(x)\), there are more than \(n\) \(S\)-neighbors \(y_i\) of \(x\) with \(C \in \mathcal{L}(y_i)\), and \(x\) has two \(S\)-neighbors \(y, z\) which are both root nodes, \(y \not\approx z\) does not hold, and \(C \in \mathcal{L}(y) \cap \mathcal{L}(z)\), then

1. set \(\mathcal{L}(z) \leftarrow \mathcal{L}(y)\),
2. for all directed edges from \(y\) to some \(w\), set \(\mathcal{L}(z, w) \leftarrow \mathcal{L}(y, w)\),
3. for all directed edges from some \(w\) to \(y\), set \(\mathcal{L}(w, z) \leftarrow \mathcal{L}(w, y)\),
4. set \(\mathcal{L}(y) = \mathcal{L}(w, y) = \mathcal{L}(y, w) = \emptyset\) for all \(w\),
5. set \(u \not\approx z\) for all \(u\) with \(u \not\approx y\), and
6. set \(y \approx z\).
Example (1): cardinalities

Show, that

\[ \text{hasChild}(\text{john}, \text{peter}) \]
\[ \text{hasChild}(\text{john}, \text{paul}) \]
\[ \text{male}(\text{peter}) \]
\[ \text{male}(\text{paul}) \]
\[ \leq 2 \text{hasChild}. \top(\text{john}) \]

does not entail \[ \forall \text{hasChild}. \text{male}(\text{john}) \].

\[ \neg \forall \text{hasChild}. \text{male} \equiv \exists \text{hasChild}. \neg \text{male} \]

now apply \[ \leq \]
Example (1): cardinalities

Show, that

\[ \text{hasChild(john, peter)} \]
\[ \text{hasChild(john, paul)} \]
\[ \text{male(peter)} \]
\[ \text{male(paul)} \]
\[ \leq 2 \text{hasChild.} \top(john) \]

does not entail \( \forall \text{hasChild.male(john)}. \)

\[ \neg \forall \text{hasChild.male} \equiv \exists \text{hasChild.} \neg \text{male} \]
Example (1): cardinalities – again

Show, that

\[
\begin{align*}
&\text{hasChild}(\text{john, peter}) \\
&\text{hasChild}(\text{john, paul}) \\
&\text{male}(\text{peter}) \\
&\text{male}(\text{paul}) \\
&\leq 2\text{hasChild.} \top (\text{john}) \text{ and peter} \neq \text{paul} \\
\end{align*}
\]

does not entail \(\forall\text{hasChild.male(}\text{john})\).

\[
\neg \forall\text{hasChild.male} \equiv \exists\text{hasChild.} \neg \text{male}
\]

now apply \(\leq\)

can backtrack only between x and peter – also leads to contradiction
Example (2): cardinalities

Show, that

\[ \geq 2\text{hasSon}. \top \text{(john)} \]
entails

\[ \geq 2\text{hasChild}. \top \text{(john)}. \]

\[
\neg \geq 2\text{hasSon}. \top \equiv \leq 1\text{hasChild}. \top
\]

\text{hasSon} \sqsubseteq \text{hasChild}

\[ \geq 2\text{hasSon}. \top \leq 1\text{hasChild}. \top \]

\text{hasSon-neighbors are also hasChild-neighbors,}
\text{tableau terminates with contradiction}
Example (3): choose

$\geq 3\text{hasSon}(\text{john})$

$\leq 2\text{hasSon.male}(\text{john})$

Is this contradictory?

No, because the following tableau is complete.
Example (4): inverse roles

\[ \exists \text{hasChild} . \text{human}(\text{john}) \]

human \(\subseteq\) \(\forall\text{hasParent} . \text{human} \)

\(\exists \text{hasChild} \sqsubseteq \neg \text{hasParent} \)

zu zeigen: \(\text{human}(\text{john})\)

\(\exists \text{hasChild} . \text{human} \)
\(\neg \text{human} \)
\(\text{human} \)

\(\text{john} \quad \text{hasChild} \quad \rightarrow \quad x\)

\(\text{human} \)
\(\neg \text{human} \)
\(\forall \text{hasParent} . \text{human} \)

\(\text{john} \) is hP\(^{-}\)-predecessor of \(x\), hence hP-neighbor of \(x\)
Example (5): Transitivity and Blocking

human ⊆ ∃hasFather.⊤
human ⊆ ∀hasAncestor.human
hasFather ⊆ hasAncestor       Trans(hasAncestor)
human(john)

Does this entail ≤1hasFather.⊤(john)?
Negation: ≥2hasFather.⊤(john)
Example (5): Transitivity and Blocking

\[
\text{human} \subseteq \exists \text{hasFather}. \top \\
\text{hasFather} \subseteq \text{hasAncestor} \\
\forall \text{hasAncestor}. \text{human}(\text{john}) \\
\text{human}(\text{john}) \\
\forall \text{hasAncestor}. \text{human}(\text{john}) \\
\geq 2 \text{hasFather}. \top (\text{john})
\]

\[
\begin{align*}
\text{john} & \quad \text{hF} \\
h & \quad \text{hF} \\
\forall \text{hA.h} & \quad \exists \text{hF}. \top \\
\forall \text{hA.h} & \quad \exists \text{hF}. \top
\end{align*}
\]

\[
\begin{align*}
\text{x} & \quad \text{hF} \\
x_1 & \quad \text{hF} \\
x_2 & \quad \text{hF}
\end{align*}
\]

\[
x_2 \text{ now blocked by } x_1: \\
\text{Pair } (x_1, x_2) \text{ repeats } (x, x_1)
\]

same as branch above
Example (6): Pairwise Blocking

\[ \neg C \cap (\leq 1F) \cap \exists F^-.D \cap \forall R^-.(\exists F^-).D, \] where
\[ D = C \cap (\leq 1F) \cap \exists F^- \neg C, \ Trans(R), \text{ and } F \subseteq R, \]
is not satisfiable.

Without pairwise blocking, z would be blocked, which shouldn't happen:
Expansion of \( \exists F^- \neg C \) yields \( \neg C \) for node y as required.
Example (7): Dynamic Blocking


with C = ∀R⁻.(∀P⁻.(∀S⁻.¬A)) and Trans(P), is not satisfiable.

Part of the tableau:

At this stage, z would be blocked by y (assuming the presence of another pair). However, when C from v is expanded, z becomes unblocked, which is necessary in order to label w with C which in turn labels x with ¬A, yielding the required contradiction.
Tableaux Reasoners

- Fact++
  - http://owl.man.ac.uk/factplusplus/

- Pellet

- RacerPro
  - http://www.sts.tu-harburg.de/~r.f.moeller/racer/
Contents

• OWL – Basic Ideas
• OWL As the Description Logic SROIQ(D)
• Different Perspectives on OWL
• OWL Semantics
• OWL Profiles
• Proof Theory
• Tools
OWL tools (incomplete listing)

Reasoner:
- **OWL 2 DL:**
  - Pellet http://clarkparsia.com/pellet/
  - HerMiT http://www.hermit-reasoner.com/
- **OWL 2 EL:**
  - CEL http://code.google.com/p/cel/
- **OWL 2 RL:**
  - essentially any rule engine
- **OWL 2 QL:**
  - essentially any SQL engine (with a bit of query rewriting on top)

Editors:
- Protégé
- NeOn Toolkit
- TopBraid Composer
Main References

• W3C OWL Working Group, OWL 2 Web Ontology Language: Document Overview. http://www.w3.org/TR/owl2-overview/

• Pascal Hitzler, Markus Krötzsch, Bijan Parsia, Peter Patel-Schneider, Sebastian Rudolph, OWL 2 Web Ontology Language: Primer. http://www.w3.org/TR/owl2-primer/

Main References – Textbooks


Further References

- DL complexity calculator: http://www.cs.man.ac.uk/~ezolin/dl/


Thanks!

OWL 2 and Rules
–
Optional Part, If Enough Time
Main References:


Contents

• Motivation: OWL and Rules
• Preliminaries: Datalog

• More rules than you ever need: SWRL
• Retaining decidability I: DL-safety
• Retaining decidability II: DL Rules

• The rules hidden in OWL 2: SROIQ Rules
• Retaining tractability I: OWL 2 EL Rules
• Retaining tractability II: DLP 2

• Retaining tractability III: ELP

Extending OWL with Rules

Rules inside OWL

putting it all together
Contents

• Motivation: OWL and Rules
• Preliminaries: Datalog

• More rules than you ever need: SWRL
• Retaining decidability I: DL-safety
• Retaining decidability II: DL Rules

• The rules hidden in OWL 2: SROIQ Rules
• Retaining tractability I: OWL 2 EL Rules
• Retaining tractability II: DLP 2

• Retaining tractability III: ELP

Extending OWL with Rules
Rules inside OWL
putting it all together
Motivation: OWL and Rules

- Rules (mainly, logic programming) as alternative ontology modelling paradigm.
- Similar tradition, and in use in practice (e.g. F-Logic)

- Ongoing: W3C RIF working group
  - Rule Interchange Format
  - based on Horn-logic
  - language standard forthcoming 2009

- Seek: Integration of rules paradigm with ontology paradigm
  - Here: Tight Integration in the tradition of OWL
  - Foundational obstacle: reasoning efficiency / decidability [naive combinations are undecidable]
Contents

• Motivation: OWL and Rules
• Preliminaries: Datalog

• More rules than you ever need: SWRL
• Retaining decidability I: DL-safety
• Retaining decidability II: DL Rules

• The rules hidden in OWL 2: SROIQ Rules
• Retaining tractability I: OWL 2 EL Rules
• Retaining tractability II: DLP 2

• Retaining tractability III: ELP

Intro

Extending OWL with Rules

Rules inside OWL

putting it all together
Preliminaries: Datalog

• Essentially Horn-rules without function symbols

general form of the rules:

\[ p_1(x_1,\ldots,x_n) \land \cdots \land p_m(y_1,\ldots,y_k) \rightarrow q(z_1,\ldots,z_j) \]

body \rightarrow head

semantics either as in predicate logic or as Herbrand semantics (see next slide)

• decidable
• polynomial data complexity (in number of facts)
• combined (overall) complexity: ExpTime
• combined complexity is P if the number of variables per rule is globally bounded
Datalog semantics example

- Example:
  \[ p(x) \rightarrow q(x) \]
  \[ q(x) \rightarrow r(x) \]
  \[ \rightarrow p(a) \]

- Predicate logic semantics:
  \[ (\forall x)(p(x) \rightarrow r(x)) \]
  and
  \[ (\forall x)(\neg r(x) \rightarrow \neg p(x)) \]
  are logical consequences

- Herbrand semantics
  those on the left are not logical consequences

  q(a) and r(a)
  are logical consequences

  material implication:
  apply only to known constants
Contents

• Motivation: OWL and Rules
• Preliminaries: Datalog

• More rules than you ever need: SWRL
• Retaining decidability I: DL-safety
• Retaining decidability II: DL Rules

• The rules hidden in OWL 2: SROIQ Rules
• Retaining tractability I: OWL 2 EL Rules
• Retaining tractability II: DLP 2

• Retaining tractability III: ELP

Intro

Extending OWL with Rules

Rules inside OWL

putting it all together
More rules than you ever need: SWRL

- Union of OWL DL with (binary) function-free Horn rules (with binary Datalog rules)

- undecidable
- no native tools available

- rather an overarching formalism

- see http://www.w3.org/Submission/SWRL/
**SWRL example (running example)**

NutAllergic(sebastian)
NutProduct(peanutOil)
\( \exists \text{orderedDish}. \text{ThaiCurry}(\text{sebastian}) \)

ThaiCurry \( \subseteq \exists \text{contains}.\{\text{peanutOil}\} \)
\( \top \subseteq \forall \text{orderedDish}. \text{Dish} \)

NutAllergic(x) \( \land \) NutProduct(y) \( \rightarrow \) dislikes(x,y)
dislikes(x,z) \( \land \) Dish(y) \( \land \) contains(y,z) \( \rightarrow \) dislikes(x,y)
orderedDish(x,y) \( \land \) dislikes(x,y) \( \rightarrow \) Unhappy(x)
SWRL example (running example)

NutAllergic(sebastian)
NutProduct(peanutOil)
∃orderedDish.ThaiCurry(sebastian)

ThaiCurry ⊆ ∃contains.\{peanutOil\}

\text{T} ⊆ ∀orderedDish.Dish

NutAllergic(x) \land NutProduct(y) \rightarrow dislikes(x,y)
dislikes(x,z) \land Dish(y) \land contains(y,z) \rightarrow dislikes(x,y)
orderedDish(x,y) \land dislikes(x,y) \rightarrow Unhappy(x)

Conclusions:
dislikes(sebastian,peanutOil)
SWRL example (running example)

NutAllergic(sebastian)
NutProduct(peanutOil)
∃orderedDish.ThaiCurry(sebastian)

ThaiCurry ⊑ ∃contains.{peanutOil}
T ⊑ ∀orderedDish.Dish

orderedDish rdfs:range Dish.

NutAllergic(x) ∧ NutProduct(y) → dislikes(x,y)
dislikes(x,z) ∧ Dish(y) ∧ contains(y,z) → dislikes(x,y)
orderedDish(x,y) ∧ dislikes(x,y) → Unhappy(x)

Conclusions:
dislikes(sebastian,peanutOil)
orderedDish(sebastian,y_s)
ThaiCurry(y_s)
Dish(y_s)
NutAllergic(sebastian)
NutProduct(peanutOil)
\exists orderedDish.ThaiCurry(sebastian)

\text{ThaiCurry} \subseteq \exists \text{contains}\{\text{peanutOil}\}
T \subseteq \forall \text{orderedDish.Dish}

\text{NutAllergic}(x) \land \text{NutProduct}(y) \rightarrow \text{dislikes}(x,y)
dislikes(x,z) \land \text{Dish}(y) \land \text{contains}(y,z) \rightarrow \text{dislikes}(x,y)
\text{orderedDish}(x,y) \land \text{dislikes}(x,y) \rightarrow \text{Unhappy}(x)

Conclusions:
dislikes(sebastian,peanutOil)     \text{contains}(y_s,peanutOil)
\text{orderedDish}(sebastian,y_s)
\text{ThaiCurry}(y_s)
\text{Dish}(y_s)
**SWRL example (running example)**

- NutAllergic(sebastian)
- NutProduct(peanutOil)
- \(\exists\text{orderedDish.ThaiCurry}(\text{sebastian})\)

ThaiCurry \(\subseteq \exists\text{contains.}\{\text{peanutOil}\}\)

\(\top \subseteq \forall\text{orderedDish.Dish}\)

\[
\text{NutAllergic}(x) \land \text{NutProduct}(y) \rightarrow \text{dislikes}(x,y) \\
\text{dislikes}(x,z) \land \text{Dish}(y) \land \text{contains}(y,z) \rightarrow \text{dislikes}(x,y) \\
\text{orderedDish}(x,y) \land \text{dislikes}(x,y) \rightarrow \text{Unhappy}(x)
\]

**Conclusions:**
- dislikes(sebastian,peanutOil)
- orderedDish(sebastian,\(y_s\))
- ThaiCurry(\(y_s\))
- Dish(\(y_s\))
- contains(\(y_s\),peanutOil)
- dislikes(sebastian,\(y_s\))
SWRL example (running example)

NutAllergic(sebastian)
NutProduct(peanutOil)
\(\exists \text{orderedDish}. \text{ThaiCurry(sebastian)}\)

\(\text{ThaiCurry} \subseteq \exists \text{contains.\{peanutOil\}}\)
\(\top \subseteq \forall \text{orderedDish}. \text{Dish}\)

NutAllergic(x) \land NutProduct(y) \rightarrow dislikes(x,y)
dislikes(x,z) \land Dish(y) \land contains(y,z) \rightarrow dislikes(x,y)
orderedDish(x,y) \land dislikes(x,y) \rightarrow Unhappy(x)

Conclusions:
dislikes(sebastian,peanutOil)
orderedDish(sebastian,y_s)
ThaiCurry(y_s)
Dish(y_s)
contains(y_s,peanutOil)
dislikes(sebastian,y_s)
Unhappy(sebastian)
SWRL example (running example)

NutAllergic(sebastian)
NutProduct(peanutOil)
\( \exists \text{orderedDish}. \text{ThaiCurry}(sebastian) \)

\text{ThaiCurry} \sqsubseteq \exists \text{contains}.\{\text{peanutOil}\}
\top \sqsubseteq \forall \text{orderedDish}. \text{Dish}

NutAllergic(x) \land NutProduct(y) \rightarrow \text{dislikes}(x,y)
dislikes(x,z) \land \text{Dish}(y) \land \text{contains}(y,z) \rightarrow \text{dislikes}(x,y)
\text{orderedDish}(x,y) \land \text{dislikes}(x,y) \rightarrow \text{Unhappy}(x)

Conclusion: \text{Unhappy}(sebastian)
Contents

- Motivation: OWL and Rules
- Preliminaries: Datalog

- More rules than you ever need: SWRL
- Retaining decidability I: DL-safety
- Retaining decidability II: DL Rules

- The rules hidden in OWL 2: SROIQ Rules
- Retaining tractability I: OWL 2 EL Rules
- Retaining tractability II: DLP 2

- Retaining tractability III: ELP

Extending OWL with Rules

Rules inside OWL

putting it all together
Retaining decidability I: DL-safety

• Reinterpret SWRL rules: Rules apply only to individuals which are explicitly given in the knowledge base.
  – Herbrand-style way of interpreting them

• OWL DL + DL-safe SWRL is decidable
• Native support e.g. by KAON2 and Pellet

DL-safe SWRL example

\[
\begin{align*}
&\text{NutAllergic(sebastian)} \\
&\text{NutProduct(peanutOil)} \\
&\exists \text{orderedDish}. \text{ThaiCurry(sebastian)} \\
&\text{ThaiCurry} \sqsubseteq \exists \text{contains}. \{\text{peanutOil}\} \\
&\top \sqsubseteq \forall \text{orderedDish}. \text{Dish} \\
&\begin{cases}
&\text{NutAllergic}(x) \land \text{NutProduct}(y) \rightarrow \text{dislikes}(x,y) \\
&\text{dislikes}(x,z) \land \text{Dish}(y) \land \text{contains}(y,z) \rightarrow \text{dislikes}(x,y) \\
&\text{orderedDish}(x,y) \land \text{dislikes}(x,y) \rightarrow \text{Unhappy}(x)
\end{cases}
\end{align*}
\]

Unhappy(sebastian) cannot be concluded
DL-safe SWRL example

NutAllergic(sebastian)
NutProduct(peanutOil)
\( \exists orderedDish. \text{ThaiCurry}(sebastian) \)

\text{ThaiCurry} \subseteq \exists \text{contains.\{peanutOil\}}
\top \subseteq \forall \text{orderedDish. Dish}

\begin{align*}
\text{DL-safe} & \left\{ \right. \\
\text{NutAllergic}(x) \land \text{NutProduct}(y) & \rightarrow \text{dislikes}(x,y) \\
\text{dislikes}(x,z) \land \text{Dish}(y) \land \text{contains}(y,z) & \rightarrow \text{dislikes}(x,y) \\
\text{orderedDish}(x,y) \land \text{dislikes}(x,y) & \rightarrow \text{Unhappy}(x)
\end{align*}

Conclusions:
\text{dislikes}(sebastian,peanutOil)
\text{orderedDish}(sebastian,y_s)
\text{ThaiCurry}(y_s)
\text{Dish}(y_s)
\text{contains}(y_s,peanutOil)
\text{dislikes}(sebastian,y_s)
Contents

- Motivation: OWL and Rules
- Preliminaries: Datalog

- More rules than you ever need: SWRL
- Retaining decidability I: DL-safety
- Retaining decidability II: DL Rules

- The rules hidden in OWL 2: SROIQ Rules
- Retaining tractability I: OWL 2 EL Rules
- Retaining tractability II: DLP 2

- Retaining tractability III: ELP
Retaining decidability II: DL Rules

- General idea:
  Find out which rules can be encoded in OWL (2 DL) anyway

- Man(x) ∧ hasBrother(x,y) ∧ hasChild(y,z) → Uncle(x)
  - Man ∩ ∃hasBrother.∃hasChild. ⊨ ⊆ Uncle

- ThaiCurry(x) → ∃contains.FishProduct(x)
  - ThaiCurry ⊨∃contains.FishProduct

- kills(x,x) → suicide(x)
  - ∃kills.Self ⊨ suicide

Note: with these two axioms,

\( \text{suicide} \) is basically the same as \( \text{kills} \)
DL Rules: more examples

- NutAllergic(x) \land NutProduct(y) \rightarrow dislikes(x,y)
  - NutAllergic \equiv \exists nutAllergic.Self
  NutProduct \equiv \exists nutProduct.Self
  nutAllergic \circ \cup \circ nutProduct \sqsubseteq dislikes

- dislikes(x,z) \land Dish(y) \land contains(y,z) \rightarrow dislikes(x,y)
  - Dish \equiv \exists dish.Self
    dislikes \circ contains^\perp \circ dish \sqsubseteq dislikes

- worksAt(x,y) \land University(y) \land supervises(x,z) \land PhDStudent(z) \rightarrow professorOf(x,z)
  - \exists worksAt.University \equiv \exists worksAtUniversity.Self
    PhDStudent \equiv \exists phDStudent.Self
    worksAtUniversity \circ supervises \circ phDStudent \sqsubseteq professorOf
DL Rules: definition

- Tree-shaped bodies
- First argument of the conclusion is the root

C(x) \land R(x,a) \land S(x,y) \land D(y) \land T(y,a) \rightarrow E(x)

- C \sqcap \exists R.\{a\} \sqcap \exists S.(D \sqcap \exists T.\{a\}) \subseteq E

[duplicating nominals is ok]
DL Rules: definition

- Tree-shaped bodies
- First argument of the conclusion is the root

\[
C(x) \land R(x,a) \land S(x,y) \land D(y) \land T(y,a) \rightarrow V(x,y)
\]

\[
C \sqcap \exists R \{a\} \sqsubseteq \exists R1.\text{Self}
\]

\[
D \sqcap \exists T \{a\} \sqsubseteq \exists R2.\text{Self}
\]

\[
R1 \circ S \circ R2 \sqsubseteq V
\]
DL Rules: definition

- Tree-shaped bodies
- First argument of the conclusion is the root
- Complex classes are allowed in the rules
  
  - Mouse(x) ∧ ∃hasNose.TrunkLike(y) → smallerThan(x, y)
  
  - ThaiCurry(x) → ∃contains.FishProduct(x)

  Note: This allows to reason with unknowns (unlike Datalog)

  - Allowed class constructors depend on the chosen underlying description logic!
DL Rules: definition

Given a description logic $\mathcal{D}$, the language $\mathcal{D}$ Rules consists of

- all axioms expressible in $\mathcal{D}$,
- plus all rules with
  - tree-shaped bodies, where
  - the first argument of the conclusion is the root, and
  - complex classes from $\mathcal{D}$ are allowed in the rules.
  - <plus possibly some restrictions concerning e.g. the use of simple roles – depending on $\mathcal{D}$>
Contents

• Motivation: OWL and Rules
• Preliminaries: Datalog

• More rules than you ever need: SWRL
• Retaining decidability I: DL-safety
• Retaining decidability II: DL Rules

• The rules hidden in OWL 2: SROIQ Rules
• Retaining tractability I: OWL 2 EL Rules
• Retaining tractability II: DLP 2

• Retaining tractability III: ELP

Extending OWL with Rules

Rules inside OWL

putting it all together
The rules hidden in OWL 2: SROIQ Rules

• N2ExpTime complete

• In fact, SROIQ Rules can be translated into SROIQ i.e. they don't add expressivity.

  Translation is polynomial.

• SROIQ Rules are essentially helpful syntactic sugar for OWL 2.
SROIQ Rules example

NutAllergic(sebastian)
NutProduct(peanutOil)
\exists orderedDish.\text{ThaiCurry}(sebastian)

\text{ThaiCurry} \sqsubseteq \exists \text{contains.}\{\text{peanutOil}\}
\top \sqsubseteq \forall \text{orderedDish.}\text{Dish}

\text{NutAllergic}(x) \land \text{NutProduct}(y) \rightarrow \text{dislikes}(x,y)
\text{dislikes}(x,z) \land \text{Dish}(y) \land \text{contains}(y,z) \rightarrow \text{dislikes}(x,y)
\text{orderedDish}(x,y) \land \text{dislikes}(x,y) \rightarrow \text{Unhappy}(x)

!not a SROIQ Rule!
SROIQ Rules normal form

- Each SROIQ Rule can be written ("linearised") such that
  - the body-tree is linear,
  - if the head is of the form $R(x,y)$, then $y$ is the leaf of the tree, and
  - if the head is of the form $C(x)$, then the tree is only the root.

- $\text{worksAt}(x,y) \land \text{University}(y) \land \text{supervises}(x,z) \land \text{PhDStudent}(z) \rightarrow \text{professorOf}(x,z)$
- $\exists \text{worksAt.University}(x) \land \text{supervises}(x,z) \land \text{PhDStudent}(z) \rightarrow \text{professorOf}(x,z)$

- $C(x) \land R(x,a) \land S(x,y) \land D(y) \land T(y,a) \rightarrow V(x,y)$
  - $(C \cap \exists R\{a}\})(x) \land S(x,y) \land (D \cap \exists T\{a}\})(y) \rightarrow V(x,y)$
Contents

• Motivation: OWL and Rules
• Preliminaries: Datalog

• More rules than you ever need: SWRL
• Retaining decidability I: DL-safety
• Retaining decidability II: DL Rules

• The rules hidden in OWL 2: SROIQ Rules
• Retaining tractability I: OWL 2 EL Rules
• Retaining tractability II: DLP 2

• Retaining tractability III: ELP

Intro
Extending OWL with Rules
Rules inside OWL
putting it all together
Retaining tractability I: OWL 2 EL Rules

- EL++ Rules are PTime complete
- EL++ Rules offer expressivity which is not readily available in EL++.

\[
\text{OWL 2 EL} \quad \supset \text{ExpTime, tractable}
\]

\[
\text{OWL 2 EL Rules} \quad \subset \text{SROIQ Rules}
\]
OWL 2 EL Rules: normal form

- Every EL++ Rule can be converted into a normal form, where
  - occurring classes in the rule body are either atomic or nominals,
  - all variables in a rule's head occur also in its body, and
  - rule heads can only be of one of the forms $A(x)$, $\exists R.A(x)$, $R(x,y)$, where $A$ is an atomic class or a nominal or $\top$ or $\bot$.

- Translation is polynomial.

- $\exists \text{worksAt.University}(x) \land \text{supervises}(x,z) \land \text{PhDStudent}(z) \rightarrow \text{professorOf}(x,z)$

- $\exists \text{worksAt}(x,y) \land \text{University}(y) \land \text{supervises}(x,z) \land \text{PhDStudent}(z) \rightarrow \text{professorOf}(x,z)$

- $\text{ThaiCurry}(x) \rightarrow \exists \text{contains.FishProduct}(x)$
OWL 2 EL Rules in a nutshell

Essentially, OWL 2 EL Rules is

- Binary Datalog with tree-shaped rule bodies,
- extended by
  - occurrence of nominals as atoms and
  - existential class expressions in the head.

- The existentials really make the difference.

- Arguably the better alternative to OWL 2 EL (aka EL++)?
  - (which is covered anyway)
Contents

• Motivation: OWL and Rules
• Preliminaries: Datalog
  
• More rules than you ever need: SWRL
• Retaining decidability I: DL-safety
• Retaining decidability II: DL Rules
  
• The rules hidden in OWL 2: SROIQ Rules
• Retaining tractability I: OWL 2 EL Rules
• Retaining tractability II: DLP 2
  
• Retaining tractability III: ELP
  
Extending OWL with Rules
Rules inside OWL
putting it all together
Retaining tractability II: DLP 2

• DLP 2 is
  – DLP (aka OWL 2 RL) extended with
  – DL rules, which use
    • left-hand-side class expressions in the bodies and
    • right-hand-side class expressions in the head.

• Polynomial transformation into 5-variable Horn rules.

• PTime.

• Quite a bit more expressive than DLP / OWL 2 RL ...
Contents

- Motivation: OWL and Rules
- Preliminaries: Datalog
  
- More rules than you ever need: SWRL
- Retaining decidability I: DL-safety
- Retaining decidability II: DL Rules
  
- The rules hidden in OWL 2: SROIQ Rules
- Retaining tractability I: OWL 2 EL Rules
- Retaining tractability II: DLP 2
  
- Retaining tractability III: ELP

Extending OWL with Rules

Rules inside OWL

putting it all together
Putting it all together:

- ELP is
  - OWL 2 EL Rules +
  - a generalisation of DL-safety +
  - variable-restricted DL-safe Datalog +
  - role conjunctions (for simple roles).

- PTime complete.
- Contains OWL 2 EL and OWL 2 RL.
- Covers variable-restricted Datalog.
DL-safe variables

- A generalisation of DL-safety.
- DL-safe variables are special variables which bind only to named individuals (like in DL-safe rules).
- DL-safe variables can replace individuals in EL++ rules.

\[
C(x) \land R(x,x_s) \land S(x,y) \land D(y) \land T(y,x_s) \rightarrow E(x)
\]

with \(x_s\) a safe variable is allowed, because
\[
C(x) \land R(x,a) \land S(x,y) \land D(y) \land T(y,a) \rightarrow E(x)
\]
is an EL++ rule.

- duplicating nominals is ok

\[
\begin{align*}
C(x) & \land R(x,a) \\
S(x,y) & \land D(y) \\
T(x,a) & \rightarrow E(x)
\end{align*}
\]
Variable-restricted DL-safe Datalog

- n-Datalog is Datalog, where the number of variables occurring in rules is globally bounded by $n$.

- Complexity of n-Datalog is PTime (for fixed $n$)
  - (but exponential in $n$)

- In a sense, this is cheating.
- In another sense, this means that using a few DL-safe Datalog rules together with an EL++ rules knowledge base shouldn't really be a problem in terms of reasoning performance.
Role conjunctions

- \( \text{orderedDish}(x,y) \land \text{dislikes}(x,y) \rightarrow \text{Unhappy}(x) \)

- In fact, role conjunctions can also be added to OWL 2 DL without increase in complexity.

Retaining tractability III: ELP

- $\text{ELP}_n$ is
  - OWL 2 EL Rules generalised by DL-safe variables +
  - DL-safe Datalog rules with at most $n$ variables +
  - role conjunctions (for simple roles).

- PTime complete (for fixed $n$).
  - exponential in $n$
- Contains OWL 2 EL and OWL 2 RL.
- Covers all Datalog rules with at most $n$ variables. (!)
ELP example

NutAllergic(sebastian)
NutProduct(peanutOil)
∃orderedDish.ThaiCurry(sebastian)

ThaiCurry ⊆ ∃contains.{peanutOil}
T ⊆ ∀orderedDish.Dish

[okay]  NutAllergic(x) ∧ NutProduct(y) → dislikes(x,y)
dislikes(x,z) ∧ Dish(y) ∧ contains(y,z) → dislikes(x,y)
orderedDish(x,y) ∧ dislikes(x,y) → Unhappy(x)

[okay – role conjunction]

not an EL++ rule
ELP example

- dislikes(x,z) \land \text{Dish}(y) \land \text{contains}(y,z) \rightarrow \text{dislikes}(x,y)
as SROIQ rule translates to

\begin{align*}
\text{Dish} & \equiv \exists \text{dish}. \text{Self} \\
\text{dislikes} \circ \text{contains} & \preceq \text{dislikes}
\end{align*}

but we don't have inverse roles in ELP!

- solution: make z a DL-safe variable:

\begin{align*}
\text{dislikes}(x,\neg z) \land \text{Dish}(y) \land \text{contains}(y,\neg z) & \rightarrow \text{dislikes}(x,y)
\end{align*}

this is fine 😊
DL-safe SWRL example

NutAllergic(sebastian)
NutProduct(peanutOil)
\exists orderedDish.ThaiCurry(sebastian)

\exists \text{ThaiCurry} \sqsubseteq \exists \text{contains.}\{\text{peanutOil}\}
\top \sqsubseteq \forall \text{orderedDish.Dish}

\text{NutAllergic}(x) \land \text{NutProduct}(y) \rightarrow \text{dislikes}(x,y)
\text{dislikes}(x,\neg z) \land \text{Dish}(y) \land \text{contains}(y,\neg z) \rightarrow \text{dislikes}(x,y)
\text{orderedDish}(x,y) \land \text{dislikes}(x,y) \rightarrow \text{Unhappy}(x)

Conclusions:
\text{dislikes}(sebastian,peanutOil)
\text{contains}(y_{s},peanutOil)
\text{orderedDish}(sebastian,y_{s})
\text{dislikes}(sebastian,y_{s})
\text{ThaiCurry}(y_{s})
\text{Dish}(y_{s})
ELP example

NutAllergic(sebastian)
NutProduct(peanutOil)
\exists \text{orderedDish}. \text{ThaiCurry}(sebastian)

\text{ThaiCurry} \subseteq \exists \text{contains.}\{\text{peanutOil}\}
\top \subseteq \forall \text{orderedDish}. \text{Dish}

\text{NutAllergic}(x) \land \text{NutProduct}(y) \rightarrow \text{dislikes}(x,y)
\text{dislikes}(x,!z) \land \text{Dish}(y) \land \text{contains}(y,!z) \rightarrow \text{dislikes}(x,y)
\text{orderedDish}(x,y) \land \text{dislikes}(x,y) \rightarrow \text{Unhappy}(x)

Conclusion: \text{Unhappy}(sebastian)
ELP Reasoner ELLY

- Implementation currently being finalised.
- Based on IRIS Datalog reasoner.
- In cooperation with STI Innsbruck (Barry Bishop, Daniel Winkler, Gulay Unel).
The Big Picture

- ELP
- OWL 2 EL
- OWL 2 EL Rules
- OWL 2
  - = SROIQ Rules
- >ExpTime
  - tractable
Closed World and ELP

• There's an extension of ELP using (non-monotonic) closed-world reasoning – based on a well-founded semantics for hybrid MKNF knowledge bases.

The Big Picture II

- **ELP**
- **hybrid ELP (local closed world)**
- **OWL 2 EL**
- **OWL 2 EL Rules**
- **OWL 2 = SROIQ Rules**
- **>ExpTime**
- **tractable**
- **data-tractable**
Thanks!

References OWL and Rules


• http://www.w3.org/Submission/SWRL/

References OWL and Rules


See also our books


(Grab a flyer.)