Description Logics: a Nice Family of Logics — Reasoning in the Lightweight DL $\mathcal{EL}$ —

Uli Sattler$^1$  Thomas Schneider$^2$

$^1$School of Computer Science, University of Manchester, UK
$^2$Department of Computer Science, University of Bremen, Germany

ESSLLI, 9 August 2012
Plan for today

1. What is $\mathcal{EL}$?
2. Normalisation
3. A simple poly-time reasoning algorithm
And now . . .

1. What is $\mathcal{EL}$?

2. Normalisation

3. A simple poly-time reasoning algorithm
EL is a restriction of ALC . . .

- that allows only conjunction and existential restrictions
- that is used to represent, e.g., medical knowledge
- whose standard reasoning problems are tractable
  i.e., there is a worst-case poly-time algorithm for deciding subsumption etc.

- whose extension $\mathcal{EL}^{++}$ with other features, namely:
  - domain and range restrictions
  - concept and role assertions
  - nominals
  - concrete domains
  - remains tractable and is a profile of OWL
Syntax and semantics of $\mathcal{EL}$

### Concepts

For $C$, $D$ concepts and $R$ a role name:

<table>
<thead>
<tr>
<th>Constructor</th>
<th>Syntax</th>
<th>Example</th>
<th>Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>top</td>
<td>$\top$</td>
<td></td>
<td>$\Delta^I$</td>
</tr>
<tr>
<td>conjunction</td>
<td>$C \sqcap D$</td>
<td>Human $\sqcap$ Male</td>
<td>$C^I \sqcap D^I$</td>
</tr>
<tr>
<td>exist. restr.</td>
<td>$\exists r. C$</td>
<td>$\exists$ hasChild.Human</td>
<td>${x \mid \exists y. (x, y) \in r^I \land y \in C^I}$</td>
</tr>
</tbody>
</table>

### Axioms

- $C \sqsubseteq D$
- $C \equiv D$ as a shortcut for “$C \sqsubseteq D, D \sqsubseteq C$”
What is $\mathcal{EL}$?

**What a tiny logic !?**

✅ We can say in $\mathcal{EL}$

Hand $\sqsubseteq \exists\text{hasPart}.\text{Finger}$

❌ but we can’t say

Hand $\sqsubseteq =5\text{hasPart}.\text{Finger}$

Finger $\sqsubseteq \exists\text{hasPart}^-\text{.Hand}$

❌ We’d like to say, but can’t

MildFlu $\equiv \text{Flu} \sqcap \forall\text{symptom. Triv}$

✅ all we can say (in $\mathcal{EL}^{++}$) is

MildFlu $\sqsubseteq \text{Flu}$

MildFlu $\sqcap \exists\text{symptom. Fever} \sqsubseteq \bot$

Fever $\sqcap \text{Triv} \sqsubseteq \bot$

$\mathcal{EL}^{++}$ is used in some large-scale ontologies

e.g. bio-medical domain, terminologies: SNOMED, GALEN, GO (see References)
\( \mathcal{EL}^{(+)} \) is not so tiny – an example ontology

\[
\begin{align*}
\text{Endocardium} & \sqsubseteq \text{Tissue} \sqcap \exists \text{cont-in. HeartWall} \sqcap \\
& \quad \exists \text{cont-in. HeartValve} \\
\text{HeartWall} & \sqsubseteq \text{BodyWall} \sqcap \exists \text{part-of. Heart} \\
\text{HeartValve} & \sqsubseteq \text{BodyValve} \sqcap \exists \text{part-of. Heart} \\
\text{Endocarditis} & \sqsubseteq \text{Inflammation} \sqcap \exists \text{has-loc. Endocardium} \\
\text{Inflammation} & \sqsubseteq \text{Disease} \sqcap \exists \text{acts-on. Tissue} \\
\text{Heartdisease} & \sqcap \exists \text{has-loc. HeartValve} \sqsubseteq \text{CriticalDisease} \\
\text{Heartdisease} & \equiv \text{Disease} \sqcap \exists \text{has-loc. Heart} \\
\end{align*}
\]

\( \mathcal{EL}^{+} \)

\[
\begin{align*}
\text{part-of} \circ \text{part-of} & \sqsubseteq \text{part-of} \\
\text{part-of} & \sqsubseteq \text{cont-in} \\
\text{has-loc} \circ \text{cont-in} & \sqsubseteq \text{has-loc}
\end{align*}
\]

Taken from [Baader et al. 2006]
Satisfiability and subsumption

**Satisfiability + coherence are trivial:** every $\mathcal{EL}$-TBox is coherent

because ?
Satisfiability and subsumption

**Satisfiability + coherence are trivial:** every $\mathcal{EL}$-TBox is coherent

- $\mathcal{I}$ with $A^\mathcal{I} = \Delta^\mathcal{I}$ and $r^\mathcal{I} = \Delta^\mathcal{I} \times \Delta^\mathcal{I}$, for all concept names $A$ and role names $r$, satisfies every $\mathcal{EL}$ axiom

- $(\mathcal{I}$ with $A^\mathcal{I} = r^\mathcal{I} = \emptyset$ doesn’t – why?)

**Subsumption ?**
Satisfiability and subsumption

Satisfiability + coherence are trivial: every $\mathcal{EL}$-TBox is coherent

- $\mathcal{I}$ with $A^\mathcal{I} = \Delta^\mathcal{I}$ and $r^\mathcal{I} = \Delta^\mathcal{I} \times \Delta^\mathcal{I}$, for all concept names $A$ and role names $r$, satisfies every $\mathcal{EL}$ axiom
- $(\mathcal{I}$ with $A^\mathcal{I} = r^\mathcal{I} = \emptyset$ doesn’t – why?)

Subsumption isn’t:
does the following TBox entail $A \subseteq B$? $A' \subseteq B'$?

$\exists r.A \subseteq \exists r.B$

$A' \equiv \exists r.\exists r.A$

$B' \equiv \exists r.\exists r.B$

(Without negation, they are no longer interreducible.)
Roadmap

**Goal:** present a decision procedure for subsumption in $\mathcal{EL}$

**Outline:**
- Normalisation procedure
- Decision procedure
  (simple, naïve, without optimisations)
And now . . .

1. What is $\mathcal{EL}$?

2. Normalisation

3. A simple poly-time reasoning algorithm
Normal form

... keeps the reasoning procedure simple

Definition
An $\mathcal{EL}$ ontology is in normal form if all axioms have these forms:

$$A_1 \cap A_2 \sqsubseteq B$$
$$A \sqsubseteq B$$
$$A \sqsubseteq \exists r.B$$
$$\exists r.A \sqsubseteq B$$

$A_{(i)}, B$: atomic concepts or $\top$  \hspace{1cm} $r$: role
The normalisation procedure

... applies **normalisation rules** to axioms in a given TBox $\mathcal{T}$
- each rule transforms an axiom into one or several shorter ones
- old axiom is removed from $\mathcal{T}$; new axioms are added
The normalisation rules

<table>
<thead>
<tr>
<th>NF1</th>
<th>Input</th>
<th>$C \equiv D$</th>
<th>Output</th>
<th>$C \sqsubseteq D$</th>
<th>$D \sqsubseteq C$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NF2</td>
<td>Input</td>
<td>$C \sqsubseteq D$</td>
<td>Output</td>
<td>$C \sqsubseteq A$</td>
<td>$A \sqsubseteq D$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NF3</td>
<td>Input</td>
<td>$\exists r. C \sqsubseteq D$</td>
<td>Output</td>
<td>$C \sqsubseteq A$</td>
<td>$\exists r. A \sqsubseteq D$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NF4</td>
<td>Input</td>
<td>$C \sqcap D \sqsubseteq E$</td>
<td>Output</td>
<td>$C \sqsubseteq A$</td>
<td>$A \sqcap D \sqsubseteq E$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NF5</td>
<td>Input</td>
<td>$B \sqsubseteq \exists r. C$</td>
<td>Output</td>
<td>$B \sqsubseteq \exists r. A$</td>
<td>$A \sqsubseteq C$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NF6</td>
<td>Input</td>
<td>$B \sqsubseteq C \sqcap D$</td>
<td>Output</td>
<td>$B \sqsubseteq C$</td>
<td>$B \sqsubseteq D$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

-arbitrary concepts
-complex concepts
-atomic concept
-fresh atomic concept
The normalisation procedure

Given TBox $\mathcal{T}$, apply NF1–NF7 axiom-wise until none can be applied

The result $\mathcal{T}'$ contains new atomic concepts $A_1, \ldots, A_k$ and is of size linear in the size of $\mathcal{T}$

**Lemma**

- For every model $\mathcal{I} \models \mathcal{T}$, there is a model $\mathcal{J} \models \mathcal{T}'$ such that $\mathcal{X}^\mathcal{J} = \mathcal{X}^\mathcal{I}$ for all $X \not\in \{A_1, \ldots, A_k\}$.
- For every model $\mathcal{I} \models \mathcal{T}'$, it holds that $\mathcal{I} \models \mathcal{T}$.

**Consequence:** $\mathcal{T}'$ is equivalent to $\mathcal{T}$ w.r.t. subsumption:

$\mathcal{T} \models C \sqsubseteq D \iff \mathcal{T}' \models C \sqsubseteq D$

for all $C, D$ that don’t use the $A_i$

**Details and Example:** see [Suntisrivaraporn 2005, pg. 37–39]
And now . . .

1. What is $\mathcal{EL}$?

2. Normalisation

3. A simple poly-time reasoning algorithm
Initial assumptions

Input: TBox $\mathcal{T}$, atomic concepts $A, B$

Question: does $\mathcal{T} \models A \sqsubseteq B$ hold?

Assumption of $A, B$ being atomic is no real restriction:

$$\mathcal{T} \models C \sqsubseteq D$$

$\iff$

$$\mathcal{T} \cup \{A \equiv C, B \equiv D\} \models A \sqsubseteq B$$

Shorter notation: $A \sqsubseteq_\mathcal{T} B$ abbreviates $\mathcal{T} \models A \sqsubseteq B$
Deciding subsumptions via subsumer sets

**Subsumer** of $A$: a concept name $B$ (including $\top$) with $A \sqsubseteq_{\mathcal{T}} B$

**Subsumer set** $S(A)$: set that contains subsumers of $A$

**Representation of subsumer sets**: in a labelled graph $G(\mathcal{T})$
- Nodes of $G(\mathcal{T}) = \text{concept names (including } \top \text{) in } \mathcal{T}$
- Label of node $A$: $S(A)$
  - $B$ in label $S(A)$ means $A \sqsubseteq_{\mathcal{T}} B$
- Label of edge $(A, B)$: set $R(A, B)$ of roles
  - $r \in R(A, B)$ means $A \sqsubseteq_{\mathcal{T}} \exists r.B$

**Outline of the procedure**:
1. Set $S(A) = \{ A, \top \}$ for every $A$
2. Monotonically build $G(\mathcal{T})$
   by exhaustively applying completion rules
3. Check whether $B \in S(A)$
The completion rules (1)

**Completion rule R1 for node X**

If
\[ A_1 \cap A_2 \subseteq B \text{ in } T \]
and \( \{A_1, A_2\} \subseteq S(X) \) and \( B \notin S(X) \)
then
\[ S(X) := S(X) \cup \{B\} \]

**Completion rule R2 for node X**

If
\[ A \sqsubseteq \exists r.B \text{ in } T \]
and \( A \in S(X) \) and \( r \notin R(X, B) \)
then
\[ R(X, B) := R(X, B) \cup \{r\} \]
The completion rules (2)

Completion rule R3 for nodes $X$, $Y$

If $\exists r. A \sqsubseteq B$ in $\mathcal{T}$
and $r \in R(X, Y)$ and $A \in S(Y)$ and $B \notin S(X)$

then $S(X) := S(X) \cup \{B\}$
The “naïve” subsumption algorithm [Baader et al. 2006]

Algorithm 1

**Input:** $\mathcal{EL}$ ontology $\mathcal{T}$, atomic classes $A, B$

**Output:** yes if $A \sqsubseteq T B$, no otherwise

For each atomic concept $A$ in $\mathcal{T}$ plus $\mathcal{T}$ 

create a node with label $\{A, T\}$

while some rule is applicable to some axiom in $\mathcal{T}$ 

choose axiom $\alpha$ and rule $R_i$ applicable to $\alpha$
apply $R_i$ to $\alpha$

if $B \in S(A)$ then output yes
else output no
Summary

Algorithm 1 . . .

- terminates in time polynomial in the size of $\mathcal{T}$

**Corollary**

Subsumption in $\mathcal{EL}$ can be decided in polynomial time.

- constructs a **canonical model** of $\mathcal{T}$
- is **sound** and **complete**: outputs yes iff $A \sqsubseteq_{\mathcal{T}} B$
- works “one-pass”: computes all $A \sqsubseteq_{\mathcal{T}} B$ at once
- is still slow for big ontologies:
  - **crux** = search for applicable rules
Extensions

Smarter versions of Algorithm 1 . . .

- are goal-oriented:
  - only apply rules that are necessary for (dis)proving $A \sqsubseteq_T B$
- are implemented in the reasoner CEL for the extension $\mathcal{EL}^{++}$
- can be extended even to the Horn fragment of $SHIQ$

For details see [Baader et al. 2005, Baader et al. 2006, Kazakov 2009].
Homework

You’re cordially invited to apply the normalisation procedure to the TBox

\[ T = \{ A \sqsubseteq B \sqcap \exists r.C, \]
\[ C \sqsubseteq \exists s.D, \]
\[ \exists r.\exists s. T \sqcap B \sqsubseteq D \} \]

and then check whether it entails

\[ A \sqsubseteq D. \]

That’s all for today. Thanks!
Bio-medical ontologies

- **SNOMED**, the systematized nomenclature of human and veterinary medicine
  

- **GALEN**
  
  [http://www.opengalen.org](http://www.opengalen.org)

- **Go**, the Gene Ontology
  
F. Baader.
Terminological cycles in a description logic with existential restrictions.
http://lat.inf.tu-dresden.de/research/papers.html#2003

F. Baader, S. Brandt, and C. Lutz.
Pushing the $\mathcal{EL}$ envelope.

Efficient reasoning in $\mathcal{EL}^+$. 
In *Description Logics*, volume 189 of *CEUR Workshop Proc.*, 2006.
S. Brandt.
Polynomial time reasoning in a description logic with existential restrictions, GCI axioms, and – what else?
http://www.cs.man.ac.uk/~sbrandt/papers.html

Y. Kazakov:
Consequence-Driven Reasoning for Horn SHIQ Ontologies.

B. Suntisrivaraporn.
Optimization and Implementation of Subsumption Algorithms for the Description Logic $\mathcal{EL}$ with Cyclic TBoxes and General Concept Inclusion Axioms.
http://lat.inf.tu-dresden.de/research/papers.html#2005