Chapter 6  

The $\mathcal{EL}$ family

The DL $\mathcal{EL}$ has the constructors

- existential restriction: $\exists r.C$;  
- conjunction: $C \sqcap D$;  
- the top concept: $\top$.

no value restriction
no disjunction
no negation, no $\bot$

Every $\mathcal{EL}$ concept is satisfiable w.r.t. any $\mathcal{EL}$ TBox and thus satisfiability is not an interesting problem.

Subsumption in $\mathcal{EL}$ is non-trivial, and cannot be reduced to satisfiability in $\mathcal{EL}$.

We show that subsumption w.r.t. general TBoxes in $\mathcal{EL}$ can be decided in polynomial time.

Note: for the dual DL $\mathcal{FL}_0$, which uses $\forall r.C$ in place of $\exists r.C$, subsumption w.r.t. general TBoxes is ExpTime-complete.
6.1 Subsumption in $\mathcal{EL}$ w.r.t. general TBoxes

Without loss of generality we assume that the concepts tested for subsumption are concept names:

Lemma 6.1

Let $\mathcal{T}$ be a general $\mathcal{EL}$ TBox, $C, D \in \mathcal{EL}$ concepts, and $A, B$ concept names not occurring in $\mathcal{T}$ or $C, D$. Then

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

Proof: blackboard.

In addition, we assume that the TBox $\mathcal{T}$ is in normal form, i.e., all GCIs in $\mathcal{T}$ have one of the following forms:

$$A \sqsubseteq B, \quad A_1 \sqcap A_2 \sqsubseteq B, \quad A \sqsubseteq \exists r.B, \quad \text{or} \quad \exists r.A \sqsubseteq B,$$

where $A, A_1, A_2, B$ are concept names or the top concept $\top$ and $r$ is a role name.
Normalisation of an $\mathcal{EL}$ TBox

One can transform a given TBox into a normalised one by applying the following normalisation rules:

- **NF0** \( \hat{D} \sqsubseteq \hat{E} \rightarrow \hat{D} \sqsubseteq A, \; A \sqsubseteq \hat{E} \)
- **NF1_r** \( C \sqcap \hat{D} \sqsubseteq B \rightarrow \hat{D} \sqsubseteq A, \; C \sqcap A \sqsubseteq B \)
- **NF1_ℓ** \( \hat{D} \sqcap C \sqsubseteq B \rightarrow \hat{D} \sqsubseteq A, \; A \sqcap C \sqsubseteq B \)
- **NF2** \( \exists r. \hat{D} \sqsubseteq B \rightarrow \hat{D} \sqsubseteq A, \; \exists r. A \sqsubseteq B \)
- **NF3** \( B \sqsubseteq \exists r. \hat{D} \rightarrow A \sqsubseteq \hat{D}, \; B \sqsubseteq \exists r. A \)
- **NF4** \( B \sqsubseteq D \sqcap E \rightarrow B \sqsubseteq D, \; B \sqsubseteq E \)

where $C, D, E$ denote arbitrary $\mathcal{EL}$ concepts,
\( \hat{D}, \hat{E} \) denote $\mathcal{EL}$ concepts that are neither concept names nor $\top$,
$B$ is a concept name, and
$A$ is a new concept name.
Normalisation example

\[
\begin{align*}
\text{NF0} & \quad \hat{D} \sqsubseteq \hat{E} \quad \rightarrow \quad \hat{D} \sqsubseteq A, \quad A \sqsubseteq \hat{E} \\
\text{NF1}_r & \quad C \cap \hat{D} \sqsubseteq B \quad \rightarrow \quad \hat{D} \sqsubseteq A, \quad C \cap A \sqsubseteq B \\
\text{NF1}_\ell & \quad \hat{D} \cap C \sqsubseteq B \quad \rightarrow \quad \hat{D} \sqsubseteq A, \quad A \cap C \sqsubseteq B \\
\text{NF2} & \quad \exists r. \hat{D} \sqsubseteq B \quad \rightarrow \quad \hat{D} \sqsubseteq A, \quad \exists r. A \sqsubseteq B \\
\text{NF3} & \quad B \sqsubseteq \exists r. \hat{D} \quad \rightarrow \quad A \sqsubseteq \hat{D}, \quad B \sqsubseteq \exists r. A \\
\text{NF4} & \quad B \sqsubseteq D \cap E \quad \rightarrow \quad B \sqsubseteq D, \quad B \sqsubseteq E
\end{align*}
\]

\[
\begin{align*}
\exists r. A \cap \exists r. \exists s. A & \sqsubseteq A \cap B \quad \rightsquigarrow_{\text{NF0}} \quad \exists r. A \cap \exists r. \exists s. A \sqsubseteq B_0, \quad B_0 \sqsubseteq A \cap B, \\
\exists r. A \cap \exists r. \exists s. A & \sqsubseteq B_0 \quad \rightsquigarrow_{\text{NF1}_r} \quad \exists r. A \sqsubseteq B_1, \quad B_1 \cap \exists r. \exists s. A \sqsubseteq B_0, \\
B_1 \cap \exists r. \exists s. A & \sqsubseteq B_0 \quad \rightsquigarrow_{\text{NF1}_\ell} \quad \exists r. \exists s. A \sqsubseteq B_2, \quad B_1 \cap B_2 \sqsubseteq B_0, \\
\exists r. \exists s. A & \sqsubseteq B_2 \quad \rightsquigarrow_{\text{NF2}} \quad \exists s. A \sqsubseteq B_3, \quad \exists r. B_3 \sqsubseteq B_2, \\
B_0 \sqsubseteq A \cap B & \quad \rightsquigarrow_{\text{NF4}} \quad B_0 \sqsubseteq A, \quad B_0 \sqsubseteq B.
\end{align*}
\]
Normalisation terminates

Lemma 6.2

Any $\mathcal{EL}$ TBox $\mathcal{T}$ can be transformed into a normalised $\mathcal{EL}$ TBox $\mathcal{T}'$ by a linear number of applications of the normalisation rules.

In addition, the size of the resulting TBox $\mathcal{T}'$ is linear in the size of $\mathcal{T}$.

Proof: Show that the abnormality degree of a TBox decreases with each rule application
Abnormal occurrence of a concept $\widehat{D}$ within a general $\mathcal{E}\mathcal{L}$ TBox:

(i) $\widehat{D}$ is the left-hand side of a GCI $\widehat{D} \sqsubseteq \widehat{E}$ where $\widehat{D}$, $\widehat{E}$ are neither concept names nor $\top$; or

(ii) $\widehat{D}$ is neither concept name nor $\top$, and this occurrence is under a conjunction or an existential restriction operator; or

(iii) the occurrence of $\widehat{D}$ is under a conjunction operator on the right-hand side of a GCI.

The abnormality degree of a general $\mathcal{E}\mathcal{L}$ TBox is the number of abnormal occurrences of a concept in this TBox:

- the abnormality degree of a TBox is bounded by the size of the TBox,
- a TBox with abnormality degree 0 is normalised.

Proof continued on blackboard.
Normalisation

original TBox $\mathcal{T}$ $\xrightarrow{\text{appropriate semantic relationship}}$ normalised TBox $\mathcal{T'}$

subsumption hierarchy for the concept names occurring in $\mathcal{T}$ yields classification of $\mathcal{T'}$

Note:

$\mathcal{T}$ and $\mathcal{T'}$ are not equivalent in the sense that they have the same models due to the introduction of new concept names by the normalisation rules.

However, $\mathcal{T'}$ is a conservative extension of $\mathcal{T}$.
Definition 6.3

For a given general $\mathcal{EL}$ TBox $\mathcal{T}_0$, its signature $\text{sig}(\mathcal{T}_0)$ consists of the concept and role names occurring in the GCIs of $\mathcal{T}_0$.

Given general $\mathcal{EL}$ TBoxes $\mathcal{T}_1$ and $\mathcal{T}_2$, we say that $\mathcal{T}_2$ is a conservative extension of $\mathcal{T}_1$ if

- $\text{sig}(\mathcal{T}_1) \subseteq \text{sig}(\mathcal{T}_2)$,

- every model of $\mathcal{T}_2$ is a model of $\mathcal{T}_1$, and

- for every model $\mathcal{I}_1$ of $\mathcal{T}_1$ there exists a model $\mathcal{I}_2$ of $\mathcal{T}_2$ such that $\mathcal{I}_1$ and $\mathcal{I}_2$ coincide on $\text{sig}(\mathcal{T}_1) \cup \{\top\}$, i.e.,
  - $\Delta^{\mathcal{I}_1} = \Delta^{\mathcal{I}_2}$,
  - $A^{\mathcal{I}_1} = A^{\mathcal{I}_2}$ for all concept names $A \in \text{sig}(\mathcal{T}_1)$, and
  - $r^{\mathcal{I}_1} = r^{\mathcal{I}_2}$ for all role names $r \in \text{sig}(\mathcal{T}_1)$.
Conservative extension properties

The notion of a conservative extension is transitive:

\(\mathcal{T}_2\) conservative extension of \(\mathcal{T}_1\)

\(\mathcal{T}_3\) conservative extension of \(\mathcal{T}_2\)

\(\mathcal{T}_3\) conservative extension of \(\mathcal{T}_1\)

Lemma 6.4

Let \(\mathcal{T}_1\) and \(\mathcal{T}_2\) be general \(\mathcal{EL}\) TBoxes such that \(\mathcal{T}_2\) is a conservative extension of \(\mathcal{T}_1\), and \(C, D\) \(\mathcal{EL}\) concepts containing only concept and role names from \(\text{sig}(\mathcal{T}_1)\).

Then \(\mathcal{T}_1 \models C \sqsubseteq D\) iff \(\mathcal{T}_2 \models C \sqsubseteq D\).

Proof: blackboard.
Conservative extension

Proposition 6.5

Assume that $\mathcal{T}_2$ is obtained from $\mathcal{T}_1$ by applying one of the normalisation rules. Then $\mathcal{T}_2$ is a conservative extension of $\mathcal{T}_1$.

Proof: blackboard.

Corollary 6.6

Let $\mathcal{T}$ be a general $\mathcal{EL}$ TBox and $\mathcal{T}'$ the normalised TBox obtained from $\mathcal{T}$ using the normalisation rules, as described in the proof of Lemma 6.2. Then we have

$$\mathcal{T} \models A \sqsubseteq B \iff \mathcal{T}' \models A \sqsubseteq B$$

for all concept names $A, B \in \text{sig}(\mathcal{T})$.

Conservative extension

application

subsumption hierarchy for the concept names occurring in $\mathcal{T}$ yields classification of $\mathcal{T}'$

Corollary 6.6

Let $\mathcal{T}$ be a general $\mathcal{EL}$ TBox and $\mathcal{T}'$ the normalised TBox obtained from $\mathcal{T}$ using the normalisation rules, as described in the proof of Lemma 6.2.

Then we have

$$\mathcal{T} \models A \sqsubseteq B \iff \mathcal{T}' \models A \sqsubseteq B$$

for all concept names $A, B \in \text{sig}(\mathcal{T})$.

Classification procedure for $\mathcal{EL}$

We assume that the input TBox $\mathcal{T}$ is a general $\mathcal{EL}$ TBox in normal form. The procedure starts with the GCIs in $\mathcal{T}$ and adds implied GCIs using appropriate inference rules.

All the GCIs generated in this way are of a specific form:

Definition 6.7

A $\mathcal{T}$-sequent is a GCI of the form

$$A \sqsubseteq B, \quad A_1 \sqcap A_2 \sqsubseteq B, \quad A \sqsubseteq \exists r.B, \quad \text{or} \quad \exists r.A \sqsubseteq B,$$

where $A, A_1, A_2, B$ are concept names in $\text{sig}(\mathcal{T})$ or the top concept $\top$, and $r$ is a role name in $\text{sig}(\mathcal{T})$.

Note:  
- The overall number of $\mathcal{T}$-sequents is polynomial in the size of $\mathcal{T}$.
- Every GCI in $\mathcal{T}$ is a $\mathcal{T}$-sequent.
- A set of $\mathcal{T}$-sequents consists of GCIs, and thus is a TBox.
### Classification rules for $\mathcal{EL}$

**CR1** \[ A \sqsubseteq A \]

**CR2** \[ A \sqsubseteq \top \]

**CR3** \[ A_1 \sqsubseteq A_2, A_2 \sqsubseteq A_3 \quad \Rightarrow \quad A_1 \sqsubseteq A_3 \]

**CR4** \[ A \sqsubseteq A_1, A \sqsubseteq A_2, A_1 \sqcap A_2 \sqsubseteq B \quad \Rightarrow \quad A \sqsubseteq B \]

**CR5** \[ A \sqsubseteq \exists r. A_1, A_1 \sqsubseteq B_1, \exists r. B_1 \sqsubseteq B \quad \Rightarrow \quad A \sqsubseteq B \]

The rules given above are, of course, not concrete rules, but rule schemata.

Concrete instance: replace meta-variables $A, A_1, A_2, B, B_1$ by concrete $\mathcal{EL}$ concepts and meta-variable $r$ by a concrete role name.

*Only instantiations are allowed for which all the GCIs occurring in the rule are $\mathcal{T}$-sequents!*
6.1.2 The Classification Procedure

Let $T$ be a general EL TBox in normal form. We start with the GCIs in $T$ and add implied GCIs using appropriate inference rules. All the GCIs generated in this way are of a specific form.

Definition 6.7. A $T$-sequent is a GCI of the form $A \sqsubseteq B$, $A_1 \sqcap A_2 \sqsubseteq B$, $A \sqsubseteq \exists r.B$, or $\exists r.A \sqcup B$, where $A, A_1, A_2, B$ are concept names in $\text{sig}(T)$ or $\top$, and $r$ is a role name in $\text{sig}(T)$.

Obviously, the overall number of $T$-sequents is polynomial in the size of $T$, and every GCI in $T$ is a $T$-sequent. A set of $T$-sequents consists of GCIs, and thus is a TBox. Inspired by its use in sequent calculi, we employ the name sequent rather than GCI to emphasize the fact that new $T$-sequents can be derived using inference rules. The prefix $T$ specifies the original TBox and restricts $T$-sequents to being normalised GCIs containing only concept and role names from $\text{sig}(T)$.

Given the normalised input TBox $T$, we define the current TBox $T'$ to be initially $T$, and add new $T$-sequents to $T'$ by applying the classification rules of Figure 6.2. The rules given in this figure are, of course, not concrete rules, but rule schemata. To build a concrete instance of such a rule schema, the meta-variables $A, A_1, A_2, B, B_1$ must be replaced by a concrete EL concept and the meta-variable $r$ by a concrete role name. However, it is important to note that only instantiations are allowed for which all the GCIs occurring in the rule are $T$-sequents.

A rule instance obtained in this way is then to be read as follows: if all the $T$-sequents above the line occur in the current TBox $T'$, add the $T$-sequent below the line to $T'$ unless it already belongs to $T'$.

### Classification rules for $\mathcal{EL}$

<table>
<thead>
<tr>
<th>Rule (CR)</th>
<th>Premise</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR1</td>
<td>$A \sqsubseteq A$</td>
<td>$A \sqsubseteq A$</td>
</tr>
<tr>
<td>CR2</td>
<td>$A \sqsubseteq \top$</td>
<td>$A \sqsubseteq \top$</td>
</tr>
<tr>
<td>CR3</td>
<td>$A_1 \sqsubseteq A_2$, $A_2 \sqsubseteq A_3$</td>
<td>$A_1 \sqsubseteq A_3$</td>
</tr>
<tr>
<td>CR4</td>
<td>$A \sqsubseteq A_1$, $A \sqsubseteq A_2$, $A_1 \sqcap A_2 \sqsubseteq B$</td>
<td>$A \sqsubseteq B$</td>
</tr>
<tr>
<td>CR5</td>
<td>$A \sqsubseteq \exists r.A_1$, $A_1 \sqsubseteq B_1$, $\exists r.B_1 \sqsubseteq B$</td>
<td>$A \sqsubseteq B$</td>
</tr>
</tbody>
</table>
### Classification rules for EL

<table>
<thead>
<tr>
<th>Rule</th>
<th>Premises</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR1</td>
<td>$A \sqsubseteq A$</td>
<td></td>
</tr>
<tr>
<td>CR2</td>
<td>$A \sqsubseteq T$</td>
<td></td>
</tr>
<tr>
<td>CR3</td>
<td>$A_1 \sqsubseteq A_2$, $A_2 \sqsubseteq A_3$</td>
<td>$A_1 \sqsubseteq A_3$</td>
</tr>
<tr>
<td>CR4</td>
<td>$A \sqsubseteq A_1$, $A \sqsubseteq A_2$, $A_1 \cap A_2 \sqsubseteq B$</td>
<td>$A \sqsubseteq B$</td>
</tr>
<tr>
<td>CR5</td>
<td>$A \sqsubseteq \exists r. A_1$, $A_1 \sqsubseteq B_1$, $\exists r. B_1 \sqsubseteq B$</td>
<td>$A \sqsubseteq B$</td>
</tr>
</tbody>
</table>

#### Example 6.8

Let $\mathcal{T}_1 = \{ A \sqsubseteq \exists r. A, \exists r. B \sqsubseteq B_1, \top \sqsubseteq B, A \sqsubseteq B_2, B_1 \cap B_2 \sqsubseteq C \}$. Let $\mathcal{T}_2 = \{ A \sqsubseteq \exists r. A, \exists r. A \sqsubseteq B \}$. 

© Franz Baader
### Classification rules for $\mathcal{EL}$

<table>
<thead>
<tr>
<th>Rule</th>
<th>Premise</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR1</td>
<td>$A \sqsubseteq A$</td>
<td>$A \sqsubseteq A$</td>
</tr>
<tr>
<td>CR2</td>
<td>$A \sqsubseteq \top$</td>
<td>$A \sqsubseteq \top$</td>
</tr>
<tr>
<td>CR3</td>
<td>$A_1 \sqsubseteq A_2$, $A_2 \sqsubseteq A_3$</td>
<td>$A_1 \sqsubseteq A_3$</td>
</tr>
<tr>
<td>CR4</td>
<td>$A \sqsubseteq A_1$, $A \sqsubseteq A_2$, $A_1 \sqsubseteq A_2 \sqsubseteq B$</td>
<td>$A \sqsubseteq B$</td>
</tr>
<tr>
<td>CR5</td>
<td>$A \sqsubseteq \exists r.A_1$, $A_1 \sqsubseteq B_1$, $\exists r.B_1 \sqsubseteq B_1$</td>
<td>$A \sqsubseteq B$</td>
</tr>
</tbody>
</table>

#### Saturation of $\mathcal{T}$:
- apply the classification rules exhaustively to the input TBox $\mathcal{T}$
- the resulting TBox $\mathcal{T}^*$ is called the saturated TBox

#### Lemma 6.9

The saturated TBox $\mathcal{T}^*$ is uniquely determined by $\mathcal{T}$, and it can be computed by a polynomial number of rule applications.

Proof: blackboard.
Classification procedure for $\mathcal{EL}$

To show that polynomial-time saturation of $\mathcal{EL}$ TBoxes yields a polynomial-time classification procedure, it is sufficient to prove the following equivalence:

$$\mathcal{T} \models A \sqsubseteq B \text{ iff } A \sqsubseteq B \in \mathcal{T}^*$$

Soundness of the classification procedure (i.e., the if-direction of the equivalence) is an easy consequence of the next lemma:

**Lemma 6.10 (Soundness)**

If all the GCIs in $\mathcal{T}'$ follow from $\mathcal{T}$ and the $\mathcal{T}$-sequents above the line of one of the rules belong to $\mathcal{T}'$

then the $\mathcal{T}$-sequent below the line also follows from $\mathcal{T}$

*Proof: blackboard.*
Classification procedure for $\mathcal{L}$

$\mathcal{T} \models A \subseteq B$ iff $A \subseteq B \in \mathcal{T}^*$

Completeness: instead of showing the only-if direction of the equivalence directly, we prove its contrapositive:

if $A \subseteq B \not\in \mathcal{T}^*$ then $\mathcal{T} \not\models A \subseteq B$.

For this purpose, we use $\mathcal{T}^*$ to construct a canonical model of $\mathcal{T}$ that

- does not satisfy the GCI $A \subseteq B$
- in case $A \subseteq B \not\in \mathcal{T}^*$. 
Definition 6.11

Let $\mathcal{T}$ be a general $\mathcal{EL}$ TBox in normal form and $\mathcal{T}^*$ the saturated TBox obtained by exhaustive application of the classification rules.

The canonical interpretation $\mathcal{I}_{\mathcal{T}^*}$ induced by $\mathcal{T}^*$ is defined as follows:

$\Delta_{\mathcal{I}_{\mathcal{T^*}}} = \{ A \mid A \text{ is a concept name in } \text{sig}(\mathcal{T}) \} \cup \{ \top \}$;

$A_{\mathcal{I}_{\mathcal{T^*}}} = \{ B \in \Delta_{\mathcal{I}_{\mathcal{T^*}}} \mid B \sqsubseteq A \in \mathcal{T}^* \}$ for all concept names $A \in \text{sig}(\mathcal{T})$;

$r_{\mathcal{I}_{\mathcal{T^*}}} = \{(A, B) \in \Delta_{\mathcal{I}_{\mathcal{T^*}}} \times \Delta_{\mathcal{I}_{\mathcal{T^*}}} \mid A \sqsubseteq \exists r . B \in \mathcal{T}^* \}$ for all role names $r \in \text{sig}(\mathcal{T})$.

Note:  
- By definition, we have $B \in A_{\mathcal{I}_{\mathcal{T^*}}}$ iff $B \sqsubseteq A \in \mathcal{T}^*$ for all concept names $A \in \text{sig}(\mathcal{T})$.
- The same is actually true for $A = \top$. 

© Franz Baader
Canonical model

Lemma 6.12

The canonical interpretation induced by $T^*$ is a model of the saturated TBox $T^*$.

Proof: blackboard.

Lemma 6.13 (Completeness)

Let $T$ be a general $E\mathcal{L}$ TBox in normal form and $T^*$ the saturated TBox obtained by exhaustive application of the classification rules. Then

$$T \models A \sqsubseteq B \text{ implies } A \sqsubseteq B \in T^*.$$

Proof: blackboard.

Theorem 6.14

Subsumption in $E\mathcal{L}$ w.r.t. general TBoxes is decidable in polynomial time.

Proof: blackboard.
6.2 Subsumption in $\mathcal{ELI}$ w.r.t. general TBoxes

Inverse roles: if $r$ is a role, then $r^{-}$ denotes its inverse

$$ (r^{-})^I := \{(e, d) \mid (d, e) \in r^I\} $$

As usual, we will use $r^{-}$ to denote $s$ if $r = s^{-}$ for a role name $s$.

In contrast to the case of $\mathcal{EL}$, subsumption in $\mathcal{ELI}$ w.r.t. general TBoxes is no longer polynomial, but $\text{ExpTime}$-complete.

One reason for the higher complexity of subsumption in $\mathcal{ELI}$ is that it can express a restricted form of value restrictions, and thus comes close to $\mathcal{FL}_{0}$:

$$ \exists r^{-}. C \sqsubseteq D \quad \text{has the same models as} \quad C \sqsubseteq \forall r. D $$

In the following, we will show the $\text{ExpTime}$-upper bound.
Normalisation of an $\mathcal{ELI}$ TBox

We say that the general $\mathcal{ELI}$ TBox $\mathcal{T}$ is in i.normal form (or i.normalised) if all its GCIs are of one of the following forms:

\[ A \subseteq B, \quad A_1 \sqcap A_2 \subseteq B, \quad A \subseteq \exists r. B, \quad \text{or} \quad A \subseteq \forall r. B, \]

where $A, A_1, A_2, B$ are concept names or the top-concept $\top$ and $r$ is a role name or the inverse of a role name.

Corollary 6.15

Given a general $\mathcal{ELI}$ TBox $\mathcal{T}$, we can compute in polynomial time an i.normalised $\mathcal{ELI}$ TBox $\mathcal{T'}$ that is a conservative extension of $\mathcal{T}$.

In particular, we have

\[ \mathcal{T} \models A \subseteq B \quad \text{iff} \quad \mathcal{T'} \models A \subseteq B \]

for all concept names $A, B \in \text{sig}(\mathcal{T})$.

Proof: blackboard.
Classification procedure for $\mathcal{ELI}$

We assume that the input TBox $\mathcal{T}$ is a general $\mathcal{ELI}$ TBox in i.normal form.

The higher complexity of subsumption in $\mathcal{ELI}$ necessitates the use of an extended notion of sequents:

**Definition 6.16**

A $\mathcal{T}$-i.sequent is an expression of the form

\[ K \sqsubseteq \{A\}, \quad K \sqsubseteq \exists r.K', \quad \text{or} \quad K \sqsubseteq \forall r.\{A\}, \]

where $K, K'$ are sets of concept names in $\text{sig}(\mathcal{T})$, $A$ is a concept name in $\text{sig}(\mathcal{T})$, and $r$ is a role name in $\text{sig}(\mathcal{T})$ or the inverse of a role name in $\text{sig}(\mathcal{T})$.

**Note:**
- The overall number of $\mathcal{T}$-i.sequents is exponential in the size of $\mathcal{T}$.
- A set in a $\mathcal{T}$-i.sequent stands for the conjunction of its element.

empty conjunction is $\top$
Classification procedure for $\mathcal{ELI}$

We assume that the input TBox $\mathcal{T}$ is a general $\mathcal{ELI}$ TBox in i.normal form.

The higher complexity of subsumption in $\mathcal{ELI}$ necessitates the use of an extended notion of sequents:

**Definition 6.16**

A $\mathcal{T}$-i.sequent is an expression of the form

$$K \subseteq \{A\}, \quad K \subseteq \exists r.K', \quad \text{or} \quad K \subseteq \forall r.\{A\},$$

where $K, K'$ are sets of concept names in $\text{sig}(\mathcal{T})$, $A$ is a concept name in $\text{sig}(\mathcal{T})$, and $r$ is a role name in $\text{sig}(\mathcal{T})$ or the inverse of a role name in $\text{sig}(\mathcal{T})$.

**Note:**
- The overall number of $\mathcal{T}$-i.sequents is exponential in the size of $\mathcal{T}$.
- A set in a $\mathcal{T}$-i.sequent stands for the conjunction of its element.
- $\mathcal{T}$-i.sequents are GCIs, and a set of $\mathcal{T}$-i.sequents is a general $\mathcal{ELI}$ TBox.
- Every GCI in the i.normalised TBox $\mathcal{T}$ is either equivalent to a $\mathcal{T}$-i.sequent or a tautology, i.e., satisfied in every interpretation.
Classification rules for $\mathcal{ELI}$

\[
\begin{align*}
i.CR1 & \quad K \sqsubseteq \{A\} & \text{if } A \in K \text{ and } K \text{ occurs in } \mathcal{T}' \\
i.CR2 & \quad M \sqsubseteq \{B\} \text{ for all } B \in K & K \sqsubseteq C \quad M \sqsubseteq C & \text{if } M \text{ occurs in } \mathcal{T}' \\
i.CR3 & \quad M_2 \sqsubseteq \exists r.M_1 & M_1 \sqsubseteq \forall r^{-}.\{A\} & M_2 \sqsubseteq \{A\} \\
i.CR4 & \quad M_1 \sqsubseteq \exists r.M_2 & M_1 \sqsubseteq \forall r.\{A\} & M_1 \sqsubseteq \exists r.(M_2 \cup \{A\})
\end{align*}
\]

The rules given above are, again, not concrete rules, but rule schemata.

Concrete instance: replace $K, M, M_1, M_2$ by sets of concept names in $\text{sig}(\mathcal{T})$, $A$ by a concept name in $\text{sig}(\mathcal{T})$, $r$ by a role name or inverse of a role name in $\text{sig}(\mathcal{T})$, $C$ by any admissible right-hand side of a $\mathcal{T}$-i.sequent.
Classification rules

In i.CR1, only instantiations are allowed for which $K$ actually occurs explicitly in some $\mathcal{T}$-i.sequent in the current TBox $\mathcal{T}'$.

Reason:
Otherwise, the procedure would always generate an exponential number of $\mathcal{T}$-i.sequents.

The analogous restriction on $M$ in rule i.CR2 is needed in the case where $K = \emptyset$.

Condition “$M \sqsubseteq \{B\}$ for all $B \in K$” trivially satisfied for all sets $M$.
Classification rules

\[ \text{i.CR1} \quad \frac{K \sqsubseteq \{A\}}{\text{if } A \in K \text{ and } K \text{ occurs in } T'} \]

\[ \text{i.CR2} \quad \frac{M \sqsubseteq \{B\} \text{ for all } B \in K \quad K \sqsubseteq C}{M \sqsubseteq C} \quad \text{if } M \text{ occurs in } T' \]

Example 6.17

\[ T = \{A \sqsubseteq B\} \cup \{A_i \sqsubseteq A_i \mid 1 \leq i \leq n\} \]

We have \( T \models M \sqcup \{A\} \sqsubseteq \{B\} \) for all (exponentially many) sets \( \emptyset \neq M \sqsubseteq \{A_1, \ldots, A_n\} \).

None of these \( T \text{-i.sequents} \) is actually generated by the rules when applied to \( T' = \{\{A\} \sqsubseteq \{B\}\} \cup \{\{A_i\} \sqsubseteq \{A_i\} \mid 1 \leq i \leq n\} \).
**Classification rules**

### Rule Schemata

- **i.CR1**
  \[
  K \sqsubseteq \{A\} \quad \text{if } A \in K \text{ and } K \text{ occurs in } T'
  \]

- **i.CR2**
  \[
  M \sqsubseteq \{B\} \quad \text{for all } B \in K \quad K \sqsubseteq C
  \]
  \[
  \Rightarrow M \sqsubseteq C
  \]
  \[
  \text{if } M \text{ occurs in } T'
  \]

### Example 6.18 (i.CR1 and i.CR2 in action)

\[
T = \{A \sqsubseteq \exists r.(A_1 \cap A_2 \cap A_3), \exists r.(A_1 \cap A_2) \sqsubseteq B\}
\]

Blackboard.
Classification rules

Due to the occurrence restrictions, the rules i.CR1 and i.CR2 cannot introduce new sets of concept names into $T'$.

The same is obviously true (without any restriction) for i.CR3.

In contrast, rule i.CR4 can generate new sets, and thus may cause an exponential blowup.
Classification rules

\[
\begin{align*}
\text{i.CR1} & \quad K \subseteq \{A\} \quad \text{if } A \in K \text{ and } K \text{ occurs in } T' \\
\text{i.CR2} & \quad M \subseteq \{B\} \text{ for all } B \in K \quad K \subseteq C \\
\text{if } M \text{ occurs in } T' \\
\text{i.CR3} & \quad M_2 \subseteq \exists r. M_1 \quad M_1 \subseteq \forall r^- \{A\} \\
\quad M_2 \subseteq \{A\} \\
\text{i.CR4} & \quad M_1 \subseteq \exists r. M_2 \quad M_1 \subseteq \forall r. \{A\} \\
\text{if } M_1 \subseteq \exists r. (M_2 \cup \{A\})
\end{align*}
\]

\underline{Example 6.19 (exponential blowup)}

\[\mathcal{T} := \{A \sqsubseteq \exists r. T\} \cup \{\exists r^- . A \sqsubseteq A_i \mid i = 1, \ldots, n\}\]

i.normalisation: \[\mathcal{T}' := \{\{A\} \sqsubseteq \exists r. \emptyset\} \cup \{\{A\} \sqsubseteq \forall r. \{A_i\} \mid i = 1, \ldots, n\}\]
Classification algorithm

i. Saturation of $\mathcal{T}$:
- apply the classification rules exhaustively to the input TBox $\mathcal{T}$
- the resulting TBox $\mathcal{T}^*$ is called the i.saturated TBox

The i.saturated TBox $\mathcal{T}^*$ is again uniquely determined by $\mathcal{T}$.

**Proposition 6.20 (soundness and completeness)**

For all concept names $A, B$ in $\text{sig}(\mathcal{T})$ such that $\{A\}$ occurs in $\mathcal{T}^*$ we have

$$\mathcal{T} \models A \sqsubseteq B \iff \{A\} \sqsubseteq \{B\} \in \mathcal{T}^*.$$  

Condition $\{A\}$ occurs in $\mathcal{T}^*$:

can easily be satisfied by adding $A \sqsubseteq A$ to the input TBox.

$\mathcal{T}$-i.sequent $\{A\} \sqsubseteq \{A\}$
Classification algorithm

Soundness, i.e. the if direction of Proposition 6.20, is an easy consequence of the next lemma and the fact that any GCI in $\mathcal{T}$ follows from $\mathcal{T}$.

**Lemma 6.21** (soundness)

Assume that

- all the GCIs in $\mathcal{T}'$ follow from $\mathcal{T}$ and
- the $\mathcal{T}$-i.sequents above the line of one of the classification rules belong to $\mathcal{T}'$.

Then the $\mathcal{T}$-i.sequent below the line also follows from $\mathcal{T}$.

*Proof: blackboard.*
To show completeness, i.e. the only-if direction of Proposition 6.20, we construct an appropriate canonical interpretation.

**Definition 6.22** (canonical interpretation)

Let $\mathcal{T}$ be a general $\mathcal{ELI}$ TBox in i.normal form and $\mathcal{T}^*$ the i.saturated TBox obtained by exhaustive application of the classification rules.

The canonical interpretation $\mathcal{I}_{\mathcal{T}^*}$ induced by $\mathcal{T}^*$ is defined as follows:

- $\Delta^{\mathcal{I}_{\mathcal{T}^*}} = \{ M \mid M \text{ is a set of concept names in } \text{sig}(\mathcal{T}) \text{ that occurs in } \mathcal{T}^* \}$,
- $A^{\mathcal{I}_{\mathcal{T}^*}} = \{ M \in \Delta^{\mathcal{I}_{\mathcal{T}^*}} \mid M \sqsubseteq \{ A \} \in \mathcal{T}^* \}$,
- $s^{\mathcal{I}_{\mathcal{T}^*}} = \{ (M, N) \in \Delta^{\mathcal{I}_{\mathcal{T}^*}} \times \Delta^{\mathcal{I}_{\mathcal{T}^*}} \mid M \sqsubseteq \exists s . N \in \mathcal{T}^* \text{ and } N \text{ is maximal, i.e., there is no } N' \supseteq N \text{ such that } M \sqsubseteq \exists s . N' \in \mathcal{T}^* \} \cup \{ (N, M) \in \Delta^{\mathcal{I}_{\mathcal{T}^*}} \times \Delta^{\mathcal{I}_{\mathcal{T}^*}} \mid M \sqsubseteq \exists s^- . N \in \mathcal{T}^* \text{ and } N \text{ is maximal, i.e., there is no } N' \supseteq N \text{ such that } M \sqsubseteq \exists s^- . N' \in \mathcal{T}^* \}$.
Lemma 6.23

Let $r$ be a role name or the inverse of a role name. Then

$$r^\mathcal{I}_{\tau^*} = \{(M, N) \in \Delta^\mathcal{I}_{\tau^*} \times \Delta^\mathcal{I}_{\tau^*} \mid M \sqsubseteq \exists r . N \in \mathcal{T}^*, \ N \text{ maximal}\} \cup \{(N, M) \in \Delta^\mathcal{I}_{\tau^*} \times \Delta^\mathcal{I}_{\tau^*} \mid M \sqsubseteq \exists r^- . N \in \mathcal{T}^*, \ N \text{ maximal}\}.$$

*Proof: blackboard.*

Lemma 6.24

The canonical interpretation induced by $\mathcal{T}^*$ is a model of the i.saturated TBox $\mathcal{T}^*$.

*Proof: blackboard.*
Classification algorithm completeness

Example (maximality condition needed)

Consider Example 6.19, where all the $\mathcal{T}$-i.sequents

$$\{A\} \subseteq \exists r.M \text{ for } M \subseteq \{A_1, \ldots, A_n\}$$

belong to $\mathcal{T}^*$.

We have $(\{A\}, \{A_1, \ldots, A_n\}) \in r^{\mathcal{T}_r}$,

but $(\{A\}, M) \notin r^{\mathcal{T}_r}$, for any strict subset $M \subset \{A_1, \ldots, A_n\}$.

In fact, such a role relationship would violate one of the GCIs

$$\{A\} \subseteq \forall r.\{A_i\}.$$
Lemma 6.25 (completeness)

Let $A, B$ in $\text{sig}(\mathcal{T})$ be such that $\{A\}$ occurs in $\mathcal{T}^*$. Then $\mathcal{T} \models A \sqsubseteq B$ implies $\{A\} \sqsubseteq \{B\} \in \mathcal{T}^*$.

Proof: blackboard.

Theorem 6.26

Subsumption in $\mathcal{ELI}$ w.r.t. general TBoxes is decidable in exponential time.

Proof: blackboard.
We can show that the algorithm for $\mathcal{ELT}$ runs in polynomial time if it receives a general $\mathcal{EL}$ TBox as input.

$\mathcal{ELT}$-i.sequents are $\mathcal{T}$-i.sequents satisfying the following restrictions:

1. the only sets occurring in them are the empty set and singleton sets,
2. value restrictions in these $\mathcal{T}$-i.sequents are only w.r.t. inverses of role names;
3. existential restrictions in these $\mathcal{T}$-i.sequents are only w.r.t. role names.

If we start with an $\mathcal{EL}$ TBox $\mathcal{T}_0$, then the corresponding i.normalised TBox $\mathcal{T}$ (written as a set of $\mathcal{T}$-i.sequents) contains only $\mathcal{ELT}$-i.sequents.
Classification algorithm for $\mathcal{ELI}$ applied to $\mathcal{EL}$

Lemma 6.27

There are only polynomially many $\mathcal{EL}$-$\mathcal{T}$-i.sequents in the size of $\mathcal{T}$.

In addition, applying a classification rule for $\mathcal{ELI}$ to a set $\mathcal{T}'$ of $\mathcal{EL}$-$\mathcal{T}$-i.sequents yields a set of $\mathcal{EL}$-$\mathcal{T}$-i.sequents.

Proof: blackboard.

Proposition 6.28

The subsumption algorithm for $\mathcal{ELI}$ yields a polynomial-time decision procedure for subsumption in $\mathcal{EL}$.

Proof: blackboard.
Classification algorithm for $\mathcal{ELI}$ is exponential

In Example 6.29, the i.saturated TBox $\mathcal{T}^*$ contains exponentially many $\mathcal{T}$-i.sequents.

In the following example, one needs to derive exponentially many $\mathcal{T}$-i.sequents before the consequence $\{A\} \subseteq \{B\}$ can be derived.

**Example 6.29 (unavoidable exponential blowup)**

\[
\begin{align*}
\{A\} & \subseteq \{\overline{X}_i\} \text{ for } 0 \leq i \leq n - 1, \\
\emptyset & \subseteq \exists r.\emptyset, \\
\{\overline{X}_i, X_0, \ldots, X_{i-1}\} & \subseteq \forall r.\{X_i\} \text{ for } 0 \leq i \leq n - 1, \\
\{X_i, X_0, \ldots, X_{i-1}\} & \subseteq \forall r.\{\overline{X}_i\} \text{ for } 0 \leq i \leq n - 1, \\
\{\overline{X}_i, \overline{X}_j\} & \subseteq \forall r.\{\overline{X}_i\} \text{ for } 0 \leq j < i \leq n - 1, \\
\{X_i, \overline{X}_j\} & \subseteq \forall r.\{X_i\} \text{ for } 0 \leq j < i \leq n - 1, \\
\{X_0, \ldots, X_{n-1}\} & \subseteq \{B\}, \\
\{B\} & \subseteq \forall r^{-}.\{B\}.
\end{align*}
\]