The description logic $\mathcal{ALC}$

We have investigated (the terminological part of) two types of **lightweight** description logics:

- $\mathcal{EL}$ which has been designed to represent large-scale ontologies such as SNOMED CT and in which terminological reasoning is tractable;
- the DL-Lite family which has been designed to access data using conceptual models such as ER and UML diagrams and in which corresponding querying tasks are often tractable.

We now consider the basic **expressive** description logic $\mathcal{ALC}$. All other expressive description logics are defined as extensions of $\mathcal{ALC}$.

Unfortunately, reasoning in $\mathcal{ALC}$ is not tractable!

$\mathcal{ALC}$ (syntax)

- **Language for $\mathcal{ALC}$ concepts (classes)**
  - concept names $A_0, A_1, \ldots$
  - role names $r_0, r_1, \ldots$
  - the concept $\top$ (often called "thing")
  - the concept $\bot$ (stands for the empty class)
  - the concept constructor $\sqcap$ (often called intersection, conjunction, or simply "and").
  - the concept constructor $\exists$ (often called existential restriction).
  - the concept constructor $\forall$ (often called value restriction).
  - the concept constructor $\sqcup$ (often called union, disjunction, or simply "or").
  - the concept constructor $\neg$ (often called complement or negation).

$\mathcal{ALC}$

$\mathcal{ALC}$ concepts are defined inductively as follows:

- All concept names, $\top$ and $\bot$ are $\mathcal{ALC}$ concepts;
- if C is a $\mathcal{ALC}$ concept, then $\neg C$ is a $\mathcal{ALC}$ concept;
- if C and D are $\mathcal{ALC}$ concepts and r is a role names, then

$$(C \sqcap D), \quad (C \sqcup D), \quad \exists r.C, \quad \forall r.C$$

  are $\mathcal{ALC}$ concepts.

A **$\mathcal{ALC}$ concept-inclusion** is of the form

$$C \sqsubseteq D,$$

where C, D are $\mathcal{ALC}$ concepts.

Examples of $\mathcal{ALC}$ concepts

- **Person** $\sqcap$ $\forall$**hasChild.Male** (everybody whose children are all male);
- **Person** $\sqcap$ $\forall$**hasChild.Male** $\sqcap$ $\exists$**hasChild.**$\top$ (everybody who has a child and whose children are all male).
- **Living_being** $\sqcap$ $\neg$**Human_being** (all living beings that are not human beings);
- **Student** $\sqcap$ $\neg\exists$**interested_in.Mathematics** (all students not interested in mathematics);
- **Student** $\sqcap$ $\forall$**drinks.tea** (all students who only drink tea).
- $\exists$**hasChild.Male** $\sqcup$ $\forall$**hasChild.**$\bot$ (everybody who has a son or no child).

Description logics: $\mathcal{ALC}$ (semantics)

Interpretations are defined as before:

- Recall that an **interpretation** is a structure $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ in which
  - $\Delta^{\mathcal{I}}$ is the **domain** (a non-empty set)
  - $\cdot^{\mathcal{I}}$ is an **interpretation function** that maps:
    - every concept name A to a subseteq $A^{\mathcal{I}}$ of $\Delta^{\mathcal{I}}$ $(A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}})$
    - every role name r to a binary relation $r^{\mathcal{I}}$ over $\Delta^{\mathcal{I}}$ $(r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}})$
- interpretation of **complex concepts** in $\mathcal{I}$:

  (C, D are concepts and r a role name)

  - $(\top)^{\mathcal{I}} = \Delta^{\mathcal{I}}$ and $(\bot)^{\mathcal{I}} = \emptyset$
  - $(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$
  - $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$ and $(C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}}$
  - $(\forall r.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid$ for all $y \in \Delta^{\mathcal{I}}$ with $(x, y) \in r^{\mathcal{I}}$ we have $y \in C^{\mathcal{I}}\}$
  - $(\exists r.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid$ exists $y \in \Delta^{\mathcal{I}}$ such that $(x, y) \in r^{\mathcal{I}}$ and $y \in C^{\mathcal{I}}\}$

Example

Let $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ be defined by setting

- $\Delta^{\mathcal{I}} = \{a, b, c, d\}$;
- $A^{\mathcal{I}} = \{b, d\}$, $B^{\mathcal{I}} = \{c\}$;
- $r^{\mathcal{I}} = \{(a, b), (a, c)\}$, $s^{\mathcal{I}} = \{(a, b), (a, d)\}$.

Then

- $(\forall r.A)^{\mathcal{I}} = \{b, c, d\}$, $(\forall s.A)^{\mathcal{I}} = \{a, b, c, d\}$;
- $(\exists r.A \sqcap \forall r.A)^{\mathcal{I}} = \emptyset$, $(\exists s.A \sqcap \forall s.A)^{\mathcal{I}} = \{a\}$;
- $(\exists r.B \sqcap \exists r.A)^{\mathcal{I}} = \{a\}$, $(\exists r.(A \sqcap B))^{\mathcal{I}} = \emptyset$;
- $(\forall r.\neg A)^{\mathcal{I}} = \{b, c, d\}$, $(\forall s.\neg A)^{\mathcal{I}} = \{b, c, d\}$.

Examples of equivalent concepts (classes)

For all interpretations $\mathcal{I}$ and all concepts C, D and roles r the following holds:

- $(\neg\neg C)^{\mathcal{I}} = C^{\mathcal{I}}$;
- $(\forall r.C)^{\mathcal{I}} = (\neg\exists r.\neg C)^{\mathcal{I}}$;
- $(\neg(C \sqcap D))^{\mathcal{I}} = (\neg C \sqcup \neg D)^{\mathcal{I}}$;
- $(\neg(C \sqcup D))^{\mathcal{I}} = (\neg C \sqcap \neg D)^{\mathcal{I}}$;
- $(\neg\exists r.C)^{\mathcal{I}} = (\forall r.\neg C)^{\mathcal{I}}$;
- $(\neg\forall r.C)^{\mathcal{I}} = (\exists r.\neg C)^{\mathcal{I}}$;
- $(C \sqcap \neg C)^{\mathcal{I}} = \bot^{\mathcal{I}} = \emptyset$;
- $(C \sqcup \neg C)^{\mathcal{I}} = \top^{\mathcal{I}} = \Delta^{\mathcal{I}}$.

Concept inclusions and TBoxes

- A **$\mathcal{ALC}$-concept inclusion** is an expression

$$C \sqsubseteq D,$$

  where C and D are $\mathcal{ALC}$-concepts.

- A **$\mathcal{ALC}$-TBox** is a finite set $\mathcal{T}$ of $\mathcal{ALC}$-concept inclusions. $\mathcal{ALC}$-terminologies and acyclic terminologies are defined following the definition for $\mathcal{EL}$.

Example

Let $\mathcal{T} = \{A \sqsubseteq \exists r.B\}$. Then

$$\mathcal{T} \not\models A \sqsubseteq \forall r.B.$$

To see this, construct an interpretation $\mathcal{I}$ such that

- $\mathcal{I} \models \mathcal{T}$;
- $\mathcal{I} \not\models A \sqsubseteq \forall r.B$.

Let $\mathcal{I}$ be defined by

- $\Delta^{\mathcal{I}} = \{a, b, c\}$;
- $A^{\mathcal{I}} = \{a\}$;
- $r^{\mathcal{I}} = \{(a, b), (a, c)\}$;
- $B^{\mathcal{I}} = \{b\}$.

Then $A^{\mathcal{I}} = \{a\} \subseteq \{a\} = (\exists r.B)^{\mathcal{I}}$ and so $\mathcal{I} \models \mathcal{T}$. But $A^{\mathcal{I}} \not\subseteq \{b, c\} = (\forall r.B)^{\mathcal{I}}$ and so $\mathcal{I} \not\models A \sqsubseteq \forall r.B$.

Example

Let $\mathcal{T} = \{A \sqsubseteq \forall r.B\}$. Then

$$\mathcal{T} \not\models A \sqsubseteq \exists r.B.$$

To see this, construct an interpretation $\mathcal{I}$ such that

- $\mathcal{I} \models \mathcal{T}$;
- $\mathcal{I} \not\models A \sqsubseteq \exists r.B$.

Let $\mathcal{I}$ be defined by

- $\Delta^{\mathcal{I}} = \{a\}$;
- $A^{\mathcal{I}} = \{a\}$;
- $r^{\mathcal{I}} = \emptyset$;
- $B^{\mathcal{I}} = \emptyset$.

Then $A^{\mathcal{I}} = \{a\} \subseteq \{a\} = (\forall r.B)^{\mathcal{I}}$ and so $\mathcal{I} \models \mathcal{T}$. But $A^{\mathcal{I}} \not\subseteq \emptyset = (\exists r.B)^{\mathcal{I}}$ and so $\mathcal{I} \not\models A \sqsubseteq \exists r.B$.

Domain and Range Restrictions in $\mathcal{ALC}$

Recall that

$$\exists r.\top \sqsubseteq C$$

states that the domain of r is contained in C. This inclusion is in $\mathcal{ALC}$.

Recall that, on the other hand,

$$\exists r^{-}.\top \sqsubseteq C$$

states that the range of r is contained in C. This inclusion is not in $\mathcal{ALC}$. We can express such a range restriction in $\mathcal{ALC}$, however, as

$$\top \sqsubseteq \forall r.C$$

Modelling in $\mathcal{ALC}$: Disjoint Classes

In $\mathcal{EL}$ we cannot represent that two (or more) classes are disjoint (have no common elements). In $\mathcal{ALC}$ we can state this in many different ways.

'Vegetable, Meat, Seafood, and Cheese are mutually disjoint' can be represented by the inclusions

$$\textbf{Vegetable} \sqcap \textbf{Meat} \sqsubseteq \bot, \quad \textbf{Vegetable} \sqcap \textbf{Seafood} \sqsubseteq \bot, \quad \textbf{Vegetable} \sqcap \textbf{Cheese} \sqsubseteq \bot,$$

$$\textbf{Meat} \sqcap \textbf{Seafood} \sqsubseteq \bot, \quad \textbf{Meat} \sqcap \textbf{Cheese} \sqsubseteq \bot, \quad \textbf{Seafood} \sqcap \textbf{Cheese} \sqsubseteq \bot,$$

Equivalently, we could write $\textbf{Vegetable} \sqsubseteq \neg\textbf{Meat}$, etc.

Note, however, that

$$\textbf{Vegetable} \sqcap \textbf{Meat} \sqcap \textbf{Seafood} \sqcap \textbf{Cheese} \sqsubseteq \bot$$

is a weaker assertion stating that nothing is Vegetable, Meat, Seafood, and Cheese at the same time. So there could still be something that is Meat and Cheese.

Modelling in $\mathcal{ALC}$: typical mistake for $\forall$

Assume we state that the domain of hasTopping is pizza (only pizza's have a topping):

$$\exists \mathbf{hasTopping}.\top \sqsubseteq \mathbf{Pizza}$$

and we add that ice cream cones have a topping that is ice cream:

$$\mathbf{IceCreamCone} \sqsubseteq \exists \mathbf{hasTopping}.\mathbf{IceCream}$$

then, if we assume that ice cream cones and pizzas are disjoint

$$\mathbf{Pizza} \sqcap \mathbf{IceCreamCone} \sqsubseteq \bot,$$

we obtain that the class IceCreamCone is empty!

Extending $\mathcal{ALC}$

We discuss the extension of $\mathcal{ALC}$ by

- qualified number restrictions;
- inverse roles;
- transitive roles;
- roles inclusions;
- nominals.

Extending $\mathcal{ALC}$ by Qualified Number Restrictions

Qualified number restrictions: if C is a concept, r a role, and n a number, then

$$(\leq n\, r.C), \quad (\geq n\, r.C)$$

are concepts. If $\mathcal{S}$ is a set, then we denote by $|\mathcal{S}|$ the number of its elements.

The interpretation of qualified number restrictions is given by

- $(\leq n\, r.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \ |\{y \in \Delta^{\mathcal{I}} \mid (x, y) \in r^{\mathcal{I}} \text{ and } y \in C^{\mathcal{I}}\}| \leq n \ \}$
- $(\geq n\, r.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \ |\{y \in \Delta^{\mathcal{I}} \mid (x, y) \in r^{\mathcal{I}} \text{ and } y \in C^{\mathcal{I}}\}| \geq n \}$

Examples

- (≥ 3 **hasChild.Male**) is the class of all objects having at least three children who are male.
- (≤ 2 **hasChild.Male**) is the class of all objects having at most two children who are male.

Extending $\mathcal{ALC}$

We have seen **unqualified** number restrictions in DL-Lite. Recall that unqualified number restrictions are of the form

- $(\leq n\ r\ \top)$, and do not admit qualifications using an arbitrary concept C.

DL-Lite does not admit such qualifications because terminological reasoning would become ExpTime-hard.

Extending $\mathcal{ALC}$ by inverse roles

Inverse roles: If r is a role name, then r^- is a role, called the inverse of r. The interpretation of inverse roles is given by

- $(r^-)^\mathcal{I} = \{(y,x) \in \Delta^\mathcal{I} \times \Delta^\mathcal{I} \mid (x,y) \in r^\mathcal{I}\}$.

r^- can occur in all places in which the role name r can occur.

Examples

- $\exists \mathbf{has\_child}^-.\mathbf{Gardener}$ is the class of all objects having a parent who is a gardener.
- $(\geq 3\mathbf{parent}^-.\mathbf{Gardener})$ is the class of all objects having at least three children who are gardeners.

We have seen inverse roles in DL-Lite. There are no inverse roles in $\mathcal{EL}$. In fact, adding inverse roles to $\mathcal{EL}$ would make reasoning ExpTime-hard.

Extending $\mathcal{ALC}$ by transitive roles and role hierarchies

Transitive roles: One can add *transitive*(r) to a TBox to state that the relation r is transitive. Thus,

- $\mathcal{I} \models$ *transitive*(r) if, and only if, $r^{\mathcal{I}}$ is transitive, i.e., for all x, y, $z \in \Delta^{\mathcal{I}}$ such that $(x, y) \in r^{\mathcal{I}}$ and $(y, z) \in r^{\mathcal{I}}$ we have $(x, z) \in r^{\mathcal{I}}$.

Examples

- The role "is part of" is often regarded as transitive.

Role hierarchies: one can add a role inclusion $r \sqsubseteq s$ to a TBox to state that r is included in s. Thus,

- $\mathcal{I} \models r \sqsubseteq s \quad \text{iff} \quad r^{\mathcal{I}} \subseteq s^{\mathcal{I}}$.

Example:

- hasSon $\sqsubseteq$ hasChild

Extending $\mathcal{ALC}$ by Nominals

Sometimes we want to use concepts/classes consisting of exactly one object or a finite set of objects. To enable the construction of such concepts, $\mathcal{ALC}$ has been extended by nominals.

Nominals: We use a, b, etc. to denote individual names. Individual names denote elements of the domain of interpretations. They are names for individual objects (not for classes or relations). Thus, we extend interpretations $\mathcal{I}$ to interpret individual names by setting $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$, $b^{\mathcal{I}} \in \Delta^{\mathcal{I}}$, etc.

For every individual name a, we call $\{a\}$ a nominal. For individual names $a_1, \ldots, a_n$, we call $\{a_1, \ldots, a_n\}$ a nominal set.

In every interpretation $\mathcal{I}$:

- $\{a\}^{\mathcal{I}} = \{a^{\mathcal{I}}\}$;
- $\{a_1, \ldots, a_n\}^{\mathcal{I}} = \{a_1^{\mathcal{I}}, \ldots, a_n^{\mathcal{I}}\}$.

Extending $\mathcal{ALC}$

In $\mathcal{ACC}$ extended with nominals we can use the expressions $\{a\}$ and $\{a_1, \ldots, a_n\}$ as concepts.

Examples:

- $\exists$**citizen_of.{France}** (citizens of France).
- $\exists$**citizen_of.{France, Ireland}** (citizens of France or Ireland).
- $\exists$**has_colour.{Green}** (all green objects).
- $\exists$**student_of.{Liverpool_University}** (students of Liverpool University).
- One can also define the concept **Colour** by giving a list of all colours:

  $$\textbf{Colour} \equiv \{\textbf{red}, \textbf{yellow}, \ldots, \textbf{green}\}$$

  and give a value restriction for the role **has_colour** by

  $$\top \sqsubseteq \forall \textbf{has\_colour}.\textbf{Colour}.$$

The expressive Description Logics $\mathcal{SHOIQ}$

The extension of $\mathcal{ALC}$ with the constructors

- qualified number restrictions,
- inverse roles,
- role hierarchies,
- transitive roles,
- and nominals

is called $\mathcal{SHOIQ}$. It is the underlying description logic of the Web Ontology Language OWL-DL we will discuss later. Standard reasoning systems (FACT, RACER, Pellet) for $\mathcal{SHOIQ}$ are based on tableau procedures similar to the one discussed for $\mathcal{ALC}$. Similar to $\mathcal{ALC}$, terminological reasoning in $\mathcal{SHOIQ}$ is decidable, but not tractable (it is ExpTime hard).

Interpretation as Graph

Definition (Interpretation as Graph)

Interpretations of $\mathcal{ALC}$ can be represented as graphs, with edges labelled by roles and nodes labelled by sets of concept names. More precisely, in such a graph

- each node corresponds to an element in the domain of the interpretation and it is labelled with all the concept names to which this element belongs in the interpretation;
- an edge with label r between two nodes says that the correspondsing two elements of the interpretation are related by the role r.

Bisimulation

Definition (Bisimulation)

Let $\mathcal{I}_1$ and $\mathcal{I}_2$ be interpretations. The relation $\bigotimes \subseteq \Delta^{\mathcal{I}_1} \times \Delta^{\mathcal{I}_2}$ is a <u>bisimulation</u> between $\mathcal{I}_1$ and $\mathcal{I}_2$ if:

(i) $d_1 \bigotimes d_2$ implies $\underline{d_1 \in A^{\mathcal{I}_1} \text{ iff } d_2 \in A^{\mathcal{I}_2}}$, for any $d_1 \in \Delta^{\mathcal{I}_1}$, $d_2 \in \Delta^{\mathcal{I}_2}$, and A any concept name;

(ii) $d_1 \bigotimes d_2$ and $(d_1, d_1') \in r^{\mathcal{I}_1}$ implies the existence of $d_2' \in \Delta^{\mathcal{I}_2}$ such that $\underline{d_1' \bigotimes d_2' \text{ and } (d_2, d_2') \in r^{\mathcal{I}_2}}$, for any $d_1, d_1' \in \Delta^{\mathcal{I}_1}$, $d_2 \in \Delta^{\mathcal{I}_2}$, and r any role name;

(iii) $d_1 \bigotimes d_2$ and $(d_2, d_2') \in r^{\mathcal{I}_2}$ implies the existence of $d_1' \in \Delta^{\mathcal{I}_1}$ such that $\underline{d_1' \bigotimes d_2' \text{ and } (d_1, d_1') \in r^{\mathcal{I}_1}}$, for any $d_1 \in \Delta^{\mathcal{I}_1}$, $d_2, d_2' \in \Delta^{\mathcal{I}_2}$, and r any role name;

Given $d_1 \in \Delta^{\mathcal{I}_1}$ and $d_2 \in \Delta^{\mathcal{I}_2}$, we define $(\mathcal{I}_1, d_1) \sim (\mathcal{I}_2, d_2)$ if there is a bisimulation $\bigotimes$ between $\mathcal{I}_1$ and $\mathcal{I}_2$ such that $d_1 \bigotimes d_2$, and say that $d_1 \in \mathcal{I}_1$ is <u>bisimilar</u> to $d_2 \in \mathcal{I}_2$.

Properties of Bisimulation for $\mathcal{ALC}$

Theorem

If $(\mathcal{I}_1, d_1) \sim (\mathcal{I}_2, d_2)$, then the following holds for all $\mathcal{ALC}$ concepts C:

$$d_1 \in C^{\mathcal{I}_1} \textit{ if and only if } d_2 \in C^{\mathcal{I}_2}.$$

Proof.

Since $(\mathcal{I}_1, d_1) \sim (\mathcal{I}_2, d_2)$, there is a bisimulation $\bigotimes$ between $\mathcal{I}_1$ and $\mathcal{I}_2$ such that $d_1 \bigotimes d_2$. We prove the theorem by induction on the structure of C. Since, up to equivalence, any $\mathcal{ALC}$ concept can be constructed using only the constructors conjunction, negation, and existential quantification, we consider only these constructors in the induction step. The base case is the one where C is a concept name. $\square$

Properties of Bisimulation for $\mathcal{ALC}$

Proof.

- Assume that $C = \mathsf{A}$. Then $d_1 \in \mathsf{A}^{\mathcal{I}_1}$

$$\text{if and only if } d_2 \in \mathsf{A}^{\mathcal{I}_2}$$

  is an immediate consequence of $d_1 \bigotimes d_2$.
- Assume that $C = D \sqcap E$. Then

$$\begin{aligned} d_1 \in (D \sqcap E)^{\mathcal{I}_1} &\text{ if and only if } d_1 \in D^{\mathcal{I}_1} \text{ and } d_1 \in E^{\mathcal{I}_1}, \\ &\text{ if and only if } d_2 \in D^{\mathcal{I}_2} \text{ and } d_2 \in E^{\mathcal{I}_2}, \\ &\text{ if and only if } d_2 \in (D \sqcap E)^{\mathcal{I}_2}, \end{aligned}$$

  where the first and third equivalences are due to the semantics of conjunction, and the second is due to the induction hypothesis applied to D and E.

□