Knowledge representation and ontology engineering

4. description logics

Nicolas Troquard
nicolas.troquard@gssi.it

Outline

Many application domains can be modelled without the full expressivity of FOL.

- E.g., Knowledge bases widely used in bio-medicine are simple taxonomies.

Description Logics are a family of logics:

- (typically) fragments of FOL
- (typically) decidable satisfiability, validity, entailment problems
- underlying the W3C Web Ontology Language

Example

- A juvenile disease affects only children or teenagers
- Children and teenagers are not adults
- A person is either a child, a teenager, or an adult
- Juvenile arthritis is a kind of arthritis and a juvenile disease
- Every kind of arthritis damages some joint

Formalisation in FOL:

- $\forall x.(\mathsf{JuvDisease}(x) \rightarrow \forall y.(\mathsf{Affects}(x, y) \rightarrow \mathsf{Child}(y) \vee \mathsf{Teenager}(y)))$
- $\forall x.(\mathsf{Child}(x) \vee \mathsf{Teenager}(x) \rightarrow \neg\mathsf{Adult}(x))$
- $\forall x.(\mathsf{Person}(x) \rightarrow (\mathsf{Child}(x) \vee \mathsf{Teenager}(x) \vee \mathsf{Adult}(x)))$
- $\forall x.(\mathsf{JuvArthritis}(x) \rightarrow \mathsf{Arthritis}(x) \wedge \mathsf{JuvDisease}(x))$
- $\forall x.(\mathsf{Arthritis}(x) \rightarrow \exists y.(\mathsf{Damages}(x, y) \wedge \mathsf{Joint}(y)))$

Some patterns:

- $\forall x.(\mathsf{C}(x) \rightarrow \mathsf{D}(x))$
- $\mathsf{D}(x)$ is a formula with one free variable, that can be identified with formulas with quantifiers:
  - $\forall x.(\mathsf{C}(x) \rightarrow \forall y.(\mathsf{R}(x, y) \rightarrow \mathsf{E}(y))$
  - $\forall x.(\mathsf{C}(x) \rightarrow \exists y.(\mathsf{R}(x, y) \wedge \mathsf{E}(y))$
- $\mathsf{C}(x)$ is a formula with one free variable, that can be identified with formulas with quantifiers (not in this example):
  - $\forall x.(\forall y.(\mathsf{R}(x, y) \rightarrow \mathsf{E}(y)) \rightarrow \mathsf{D}(x))$
  - $\forall x.(\exists y.(\mathsf{R}(x, y) \wedge \mathsf{E}(y)) \rightarrow \mathsf{D}(x))$

Variables are redundant

- $\forall x.(\mathsf{Child}(x) \vee \mathsf{Teenager}(x) \rightarrow \neg\mathsf{Adult}(x))$
  - Outermost universal quantification on variable x.
  - Variable x is free in both $\mathsf{Child}(x) \vee \mathsf{Teenager}(x)$ and $\neg\mathsf{Adult}(x)$.

  In DL:

  $$\mathsf{Child} \sqcup \mathsf{Teenager} \text{ is subset of } \neg\mathsf{Adult}$$

- $\forall x.(\mathsf{Arthritis}(x) \rightarrow \exists y.(\mathsf{Damages}(x, y) \wedge \mathsf{Joint}(y)))$
  - $\exists y.(\mathsf{Damages}(x, y) \wedge \mathsf{Joint}(y))$ introduces one fresh variable y.
  - $\mathsf{Damages}(x, y)$ uses the free (in subformula) variable x, and the fresh variable y.
  - $\mathsf{Joint}(y)$ uses the fresh variable y.

  In DL:

  $$\mathsf{Arthritis} \text{ is subset of } \exists\mathsf{Damages}.\mathsf{Joint}$$

- Analogously for $\forall x.(\mathsf{JuvDisease}(x) \rightarrow \forall y.(\mathsf{Affects}(x, y) \rightarrow \mathsf{Child}(y) \vee \mathsf{Teenager}(y)))$.

  In DL:

  $$\mathsf{JuvDisease} \text{ is subset of } \forall\mathsf{Affects}.(\mathsf{Child} \sqcup \mathsf{Teenager})$$

DL basic building blocks and their FOL counterparts

Concepts (unary predicates):

- $\mathsf{C} \mapsto \mathsf{C}(x)$

"Roles" (binary predicates):

- $\mathsf{R} \mapsto \mathsf{R}(x, y)$

Complex concepts:

- $\neg\mathsf{C} \mapsto \neg\mathsf{C}(x)$
- $\mathsf{C} \sqcup \mathsf{D} \mapsto \mathsf{C}(x) \vee \mathsf{D}(x)$
- $\mathsf{C} \sqcap \mathsf{D} \mapsto \mathsf{C}(x) \wedge \mathsf{D}(x)$
- $\exists\mathsf{R}.\mathsf{C} \mapsto \exists y.(\mathsf{R}(x, y) \wedge \mathsf{C}(y))$
- $\forall\mathsf{R}.\mathsf{C} \mapsto \forall y.(\mathsf{R}(x, y) \rightarrow \mathsf{C}(y))$

A decidable fragment of FOL

This is a fragment of FO2, that is First Order Logic consisting of all first-order sentences with at most two distinct variables.

- FO2 has the finite-model property and the satisfiability problem is decidable [Mortimer 1975][1]
- the satisfiability problem of FO2 is NEXPTIME-complete [Grädel, Kolaitis, Vardi 1997][2]

[1] Michael Mortimer. "On language with two variables". In: *Zeit. fur Math. Logik und Grund. der Math.* 21 (1975), pp. 135–140.

[2] Erich Grädel, Phokion G. Kolaitis, and Moshe Y. Vardi. "On the Decision Problem for Two-Variable First-Order Logic". In: *The Bulletin of Symbolic Logic* 3.1 (1997), pp. 53–69.

Outline

Basic language

(Basic) DL languages are inductively defined from:

- Unary FOL predicate symbols
- Binary FOL predicate symbols
- FOL constants
- Symbol $\top$: truth
- Symbol $\bot$: falsum
- Propositional connectives: $\neg$, $\sqcup$, $\sqcap$
- The existential and universal quantifiers: $\exists$, $\forall$
- Parentheses (and)

No functions. No predicates of arity greater than 2.

$\mathcal{ALC}$ concepts

$\mathcal{ALC}$ concepts are built inductively from a set of atomic concepts and a set of (atomic) roles, as follows:[3]

$$C ::= \top \mid \bot \mid \mathsf{A} \mid \neg C \mid C \sqcup C \mid C \sqcap C \mid \exists \mathsf{R}.C \mid \forall \mathsf{R}.C$$

where A is an atomic concept and R is an atomic role.

A concept represents a set of object. E.g.:

Child	the set of children
Child $\sqcup$ Teenager	the set of non-adult persons
Disease $\sqcap$ $\forall$Affects.Child	the set of diseases affecting only children
Disease $\sqcap$ $\exists$Affects.Child	the set of diseases affecting some children
Male $\sqcap$ $\exists$hasChild.($\exists$hasChild.$\top$)	the set of grandfathers
Female $\sqcap$ $\forall$hasChild.($\exists$hasChild.Male)	the set of women whose children all have a boy

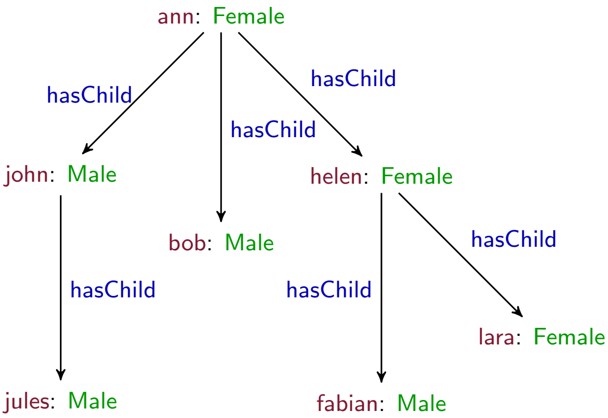
[3] $\mathcal{ALC}$ stands for Attributive Language with Concept negation.

Graphical visualisation of objects, concepts, and roles

The set of women whose all children have no children or have a boy:

$$\text{Female} \sqcap \forall \text{hasChild}.((\forall \text{hasChild}.\bot) \sqcup (\exists \text{hasChild}.\text{Male}))$$

Ann (ann) is one of them.

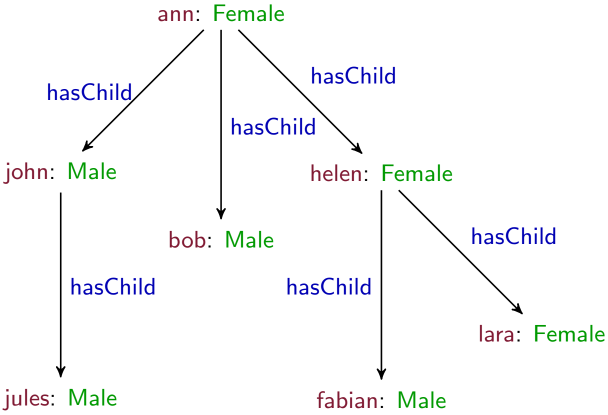


Graphical visualisation of objects, concepts, and roles

The set of women whose all children have no children or have a boy:

$$\text{Female} \sqcap \forall \text{hasChild}.((\forall \text{hasChild}.\bot) \sqcup (\exists \text{hasChild}.\text{Male}))$$

Ann (ann) is one of them.



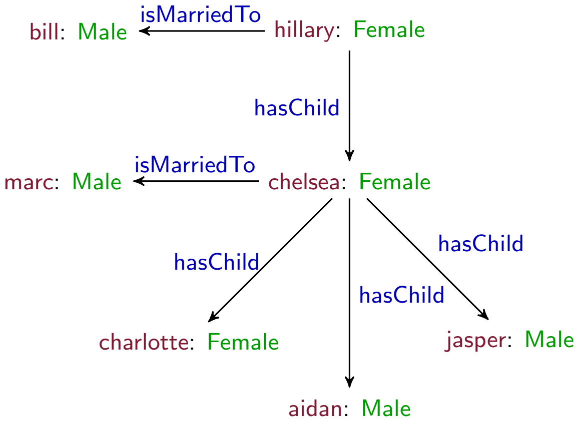
This is a knowledge graph!

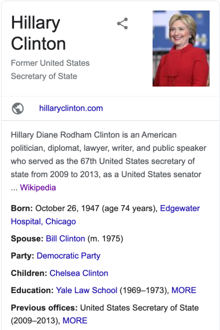
Graphical visualisation of objects, concepts, and roles

The set of women whose all children have no children or have a boy:

$$\text{Female} \sqcap \forall \text{hasChild}.((\forall \text{hasChild}.\bot) \sqcup (\exists \text{hasChild}.\text{Male}))$$

Hillary (hillary) is one of them.





General concept inclusions

Formalisation in FOL:

- $\forall x.(\text{JuvDisease}(x) \rightarrow \forall y.(\text{Affects}(x,y) \rightarrow \text{Child}(y) \vee \text{Teenager}(y)))$
- $\forall x.(\text{Child}(x) \vee \text{Teenager}(x) \rightarrow \neg\text{Adult}(x))$
- $\forall x.(\text{Person}(x) \rightarrow (\text{Child}(x) \vee \text{Teenager}(x) \vee \text{Adult}(x)))$
- $\forall x.(\text{JuvArthritis}(x) \rightarrow \text{Arthritis}(x) \wedge \text{JuvDisease}(x))$
- $\forall x.(\text{Arthritis}(x) \rightarrow \exists y.(\text{Damages}(x,y) \wedge \text{Joint}(y)))$

They all have the form:

$$\forall x.(\text{C}(x) \rightarrow \text{D}(x))$$

In DL:

$$\text{C} \sqsubseteq \text{D}$$

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

We call these statements General concept inclusions (GCIs).

General concept inclusions

Formalisation in FOL:

- $\forall x.(\text{JuvDisease}(x) \rightarrow \forall y.(\text{Affects}(x, y) \rightarrow \text{Child}(y) \vee \text{Teenager}(y)))$
- $\forall x.(\text{Child}(x) \vee \text{Teenager}(x) \rightarrow \neg\text{Adult}(x))$
- $\forall x.(\text{Person}(x) \rightarrow (\text{Child}(x) \vee \text{Teenager}(x) \vee \text{Adult}(x)))$
- $\forall x.(\text{JuvArthritis}(x) \rightarrow \text{Arthritis}(x) \wedge \text{JuvDisease}(x))$
- $\forall x.(\text{Arthritis}(x) \rightarrow \exists y.(\text{Damages}(x, y) \wedge \text{Joint}(y)))$

Formalisation in DL:

- $\text{JuvDisease} \sqsubseteq \forall \text{Affects}.(\text{Child} \sqcup \text{Teenager})$
- $\text{Child} \sqcup \text{Teenager} \sqsubseteq \neg\text{Adult}$
- $\text{Person} \sqsubseteq \text{Child} \sqcup \text{Teenager} \sqcup \text{Adult}$
- $\text{JuvArthritis} \sqsubseteq \text{Arthritis} \sqcap \text{JuvDisease}$
- $\text{Arthritis} \sqsubseteq \exists \text{Damages}.\text{Joint}$

Terminological statements

- Subtyping:

$$\text{Arthritis} \sqsubseteq \text{Disease}$$

- Definitions:

$$\text{JuvArthritis} \equiv \text{JuvDisease} \sqcap \text{Arthritis}$$

- Disjointness:

$$\text{Child} \sqsubseteq \neg\text{Adult}$$

- Covering:

$$\text{Person} \sqsubseteq \text{Child} \sqcup \text{Teenager} \sqcup \text{Adult}$$

- Domain restriction:

$$\exists\text{Affects}.\top \sqsubseteq \text{Disease}$$

- Range restriction:

$$\top \sqsubseteq \forall\text{Affects}.\text{LivingThing}$$

Axioms

Domain knowledge is represented with a set of CGIs of the form:

$$C \sqsubseteq D$$

where C and D are concepts.

They form the terminological box TBox.

Data is represented with a set of statements of the form:

$$C(a)$$

where C is a concept and a is an object/individual; and of the form

$$R(a, b) \quad ; \quad \neg R(a, b)$$

where R is a role, and a and b are individuals.

They form the assertion box ABox.

Description Logic knowledge bases

An $\mathcal{ALC}$ knowledge base is composed of:

- a TBox
- an ABox

E.g.,

TBox:

- $\mathsf{JuvDisease} \sqsubseteq \forall \mathsf{Affects}.(\mathsf{Child} \sqcup \mathsf{Teenager})$
- $\mathsf{Child} \sqcup \mathsf{Teenager} \sqsubseteq \neg \mathsf{Adult}$
- $\mathsf{Person} \sqsubseteq \mathsf{Child} \sqcup \mathsf{Teenager} \sqcup \mathsf{Adult}$
- $\mathsf{JuvArthritis} \sqsubseteq \mathsf{Arthritis} \sqcap \mathsf{JuvDisease}$
- $\mathsf{Arthritis} \sqsubseteq \exists \mathsf{Damages}.\mathsf{Joint}$

ABox:

- $\mathsf{JuvArthritis}(\mathsf{jra})$
- $\mathsf{Child}(\mathsf{johnSmith})$
- $(\mathsf{Child} \sqcup \mathsf{Teenager})(\mathsf{maryJones})$
- $\mathsf{Affects}(\mathsf{jra}, \mathsf{johnSmith})$

Outline

An interpretation for $\mathcal{ALC}$ is a tuple $\mathcal{I} = (\Delta^{\mathcal{I}}, .^{\mathcal{I}})$, where:

- $\Delta^{\mathcal{I}}$ is non-empty set; the domain of interpretation
- $.^{\mathcal{I}}$ is the interpretation function that associates:
  - every concept name A to a subset $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$
  - every role name R to a subset $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
  - every individual name a an object $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$
  - $\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$
  - $\bot^{\mathcal{I}} = \emptyset$

Given an interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, .^{\mathcal{I}})$, the meaning of a concepts is given inductively, as follows:

- $(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$
- $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$
- $(C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}}$
- $(\forall R.C)^{\mathcal{I}} = \{x \mid \forall y.(x, y) \in R^{\mathcal{I}} \rightarrow y \in C^{\mathcal{I}}\}$
- $(\exists R.C)^{\mathcal{I}} = \{x \mid \exists y.(x, y) \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$

Example

Let $\mathcal{I} = (\Delta^{\mathcal{I}}, .^{\mathcal{I}})$ be the interpretation defined as:

- $\Delta^{\mathcal{I}} = \{\mathsf{jra}, \mathsf{flu}, \mathsf{johnSmith}\}$
- $\mathsf{JuvDisease}^{\mathcal{I}} = \{\mathsf{jra}\}$
- $\mathsf{Child}^{\mathcal{I}} = \{\mathsf{johnSmith}\}$
- $\mathsf{Teenager}^{\mathcal{I}} = \emptyset$
- $\mathsf{Affects}^{\mathcal{I}} = \{(\mathsf{jra}, \mathsf{johnSmith})\}$

We have:

- $(\mathsf{JuvDisease} \sqcap \mathsf{Child})^{\mathcal{I}} = \emptyset$
- $(\mathsf{Child} \sqcup \mathsf{Teenager})^{\mathcal{I}} = \{\mathsf{johnSmith}\}$
- $(\exists \mathsf{Affects}.\top)^{\mathcal{I}} = \{\mathsf{jra}\}$
- $(\exists \mathsf{Affects}.\mathsf{Teenager})^{\mathcal{I}} = \emptyset$
- $(\exists \mathsf{Affects}.(\mathsf{Child} \sqcup \mathsf{Teenager}))^{\mathcal{I}} = \{\mathsf{jra}\}$
- $(\neg \mathsf{Child})^{\mathcal{I}} = \{\mathsf{jra}, \mathsf{flu}\}$
- $(\forall \mathsf{Affects}.\mathsf{Teenager})^{\mathcal{I}} = \{\mathsf{flu}, \mathsf{johnSmith}\}$

Given an interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, .^{\mathcal{I}})$, we define the satisfiability of axioms as:

- $\mathcal{I} \models C \sqsubseteq D$ iff $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.
- $\mathcal{I} \models C \equiv D$ iff $C^{\mathcal{I}} = D^{\mathcal{I}}$.
- $\mathcal{I} \models C(a)$ iff $a^{\mathcal{I}} \in C^{\mathcal{I}}$.
- $\mathcal{I} \models R(a, b)$ iff $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$.
- $\mathcal{I} \models \neg R(a, b)$ iff $(a^{\mathcal{I}}, b^{\mathcal{I}}) \notin R^{\mathcal{I}}$.

An $\mathcal{ALC}$ knowledge base $\mathcal{K}$ is composed of:

- a TBox $\mathcal{T}$ ($C \sqsubseteq D$ axioms)
- an ABox $\mathcal{A}$ ($C(a)$, $R(a, b)$, $\neg R(a, b)$ axioms)

An interpretation $\mathcal{I}$ is a model of the knowledge base $\mathcal{K}$ if $\mathcal{I} \models ax$ for every axiom ax in $\mathcal{T} \cup \mathcal{A}$.
A knowledge base is satisfiable if it has a model.

Example

Let $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ be the interpretation,

- $\Delta^{\mathcal{I}} = \{\mathsf{jra}, \mathsf{flu}, \mathsf{johnSmith}\}$
- $\mathsf{JuvDisease}^{\mathcal{I}} = \{\mathsf{jra}\}$
- $\mathsf{Child}^{\mathcal{I}} = \{\mathsf{johnSmith}\}$
- $\mathsf{Teenager}^{\mathcal{I}} = \emptyset$
- $\mathsf{Affects}^{\mathcal{I}} = \{(\mathsf{jra}, \mathsf{johnSmith})\}$

We have:

- $(\forall \mathsf{Affects}.(\mathsf{Child} \sqcup \mathsf{Teenager}))^{\mathcal{I}} = \Delta^{\mathcal{I}}$
- $(\forall \mathsf{Affects}.\mathsf{Teenager})^{\mathcal{I}} = \{\mathsf{flu}, \mathsf{johnSmith}\}$

Consider the knowledge base $\mathcal{K}$,

- $\mathsf{JuvDisease}(\mathsf{jra})$
- $\mathsf{Affects}(\mathsf{jra}, \mathsf{johnSmith})$
- $\mathsf{JuvDisease} \sqsubseteq \forall \mathsf{Affects}.(\mathsf{Child} \sqcup \mathsf{Teenager})$

$\mathcal{I}$ is a model of $\mathcal{K}$.

Consider the knowledge base $\mathcal{K}'$,

- $\mathsf{JuvDisease}(\mathsf{jra})$
- $\mathsf{Affects}(\mathsf{jra}, \mathsf{johnSmith})$
- $\mathsf{JuvDisease} \sqsubseteq \forall \mathsf{Affects}.\mathsf{Teenager}$

$\mathcal{I}$ is not a model of $\mathcal{K}'$.

Outline

Reasoning problems

Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be an $\mathcal{ALC}$ knowledge base, let C, D be $\mathcal{ALC}$ concepts, and let b be a named individual.

- C is satisfiable wrt. to $\mathcal{T}$ if there exists a model $\mathcal{I}$ of $\mathcal{T}$ and some $d \in \Delta^{\mathcal{I}}$ with $d \in C^{\mathcal{I}}$.
- C is subsumed by D wrt. to $\mathcal{T}$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ for every model $\mathcal{I}$ of $\mathcal{T}$.
- C and D are equivalent wrt. $\mathcal{T}$ if $C^{\mathcal{I}} = D^{\mathcal{I}}$ for every model $\mathcal{I}$ of $\mathcal{T}$.
- $\mathcal{K}$ is consistent/satisfiable if there exists a model of $\mathcal{K}$.
- b is an instance of C wrt. $\mathcal{K}$ if $b^{\mathcal{I}} \in C^{\mathcal{I}}$ for every model $\mathcal{I}$ of $\mathcal{K}$.

All reduceable to knowledge base consistency/satisfiability.

Reasoning problems (reduceability to KB consistency/satisfiability)

Reducing equivalence to subsumption:

$$\mathcal{T} \models C \equiv D \text{ iff } \mathcal{T} \models C \sqsubseteq D \text{ and } \mathcal{T} \models D \sqsubseteq C$$

Reducing subsumption to concept satisfiability:

$$\mathcal{T} \models C \sqsubseteq D \text{ iff } C \sqcap \neg D \text{ is } \textit{not} \text{ satisfiable wrt. } \mathcal{T}$$

Reducing concept satisfiability to knowledge base consistency:

$$C \text{ is satisfiable wrt. } \mathcal{T} \text{ iff } (\mathcal{T}, \{C(b)\}) \text{ is consistent.}$$

Reducing instance checking to knowledge base consistency:

$$(\mathcal{T}, \mathcal{A}) \models C(b) \text{ iff } (\mathcal{T}, \mathcal{A} \cup \{\neg C(b)\}) \text{ is } \textit{not} \text{ consistent.}$$

Exercise

- *Find a direct reduction from subsumption to knowledge base consistency.*
- *Prove the above equivalences using the definitions.*

Complexity

The problem of deciding the satisfiability of $\mathcal{ALC}$ knowledge bases is EXPTIME-complete.

Efficient tableau method in [Donini and Massacci 2000][4].

[4] Francesco M. Donini and Fabio Massacci. "EXPtime tableaux for ALC". In: *Artificial Intelligence* 124.1 (2000), pp. 87–138.

Outline

Additional concepts: Qualified number restrictions

Syntax: For every integer n, role name R, and concept C, we can also have the concept $(\leq n\, R.C)$, that refers to things have less than n R-successors that are C.
Similarly, we can have the concept $(\geq n\, R.C)$.

Let $\mathcal{DL}$ be a description logic. The set of $\mathcal{DLQ}$ concepts is the smallest set of concepts that contains all $\mathcal{DL}$ concepts and $(\leq n\, R.C)$ and $(\geq n\, R.C)$ for every $n \in \mathbb{N}$, role R, and concept C.

Semantics: Let $\#RC = \#(\{y \mid (x,y) \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\})$

$$(\leq n\, R.C)^{\mathcal{I}} = \{x \mid \#RC \leq n\} \quad ; \quad (\geq n\, R.C)^{\mathcal{I}} = \{x \mid \#RC \geq n\}$$

Example:

MildArthritis $\equiv$ Arthritis $\sqcap (\leq 2$ Damages.Joint$)$; SevereArthritis $\equiv$ Arthritis $\sqcap (\geq 5$ Damages.Joint$)$

Remarks:

- $\exists R.C$ and $(\geq 1\, R.C)$ are equivalent.
- One can define
  - $(= n\, R.C) = (\geq n\, R.C) \sqcap (\leq n\, R.C)$
  - $(> n\, R.C) = \neg(\leq n\, R.C)$
  - $(< n\, R.C) = \neg(\geq n\, R.C)$
- Functionality ($\mathcal{F}$) can be expressed with: $\top \sqsubseteq (\leq 1$ hasSSN.$\top)$

Additional roles: Inverse roles

Syntax: We can add an RBox $\mathcal{R}$, and for every role name R, we can also have the role R^- that represents the inverse of role R.

Let $\mathcal{DL}$ be a description logic. The set of $\mathcal{DLI}$ concepts is the smallest set of concepts that contains all $\mathcal{DL}$ concepts and where inverse roles can occur in all places of role names.

Semantics:

$$(R^-)^\mathcal{I} = \{(x, y) \mid (y, x) \in R^\mathcal{I}\}$$

Example (an RBox axiom):

$$\mathsf{Affects}^- \equiv \mathsf{AffectedBy}$$

Another way to capture range restrictions:

$$\exists \mathsf{Affects}^-.\top \sqsubseteq \mathsf{Person}$$

Expressing symmetry:

$$\mathsf{hasSibling} \sqsubseteq \mathsf{hasSibling}^-$$

Additional concepts: Nominals

Syntax: For every individual name a, we can also have the concept {a}, to represent the singleton containing a.

Let $\mathcal{DL}$ be a description logic. The set of $\mathcal{DLO}$ concepts is the smallest set of concepts that contains all $\mathcal{DL}$ concepts and $\{a\}$ for every individual name a.

Semantics:

$$(\{a\})^{\mathcal{I}} = \{a^{\mathcal{I}}\}$$

Example: Individuals affected by a disease that also affects johnSmith

$$\exists \mathsf{Affects}^{-}.(\mathsf{Disease} \sqcap \exists \mathsf{Affect}.\{\mathsf{johnSmith}\})$$

Simulating ABox axioms:

$$\{\mathsf{johnSmith}\} \sqsubseteq \mathsf{Child} \quad ; \quad \{\mathsf{jra}\} \sqsubseteq \exists \mathsf{Affects}.\{\mathsf{johnSmith}\} \quad ; \quad \{\mathsf{jra}\} \sqsubseteq \neg\exists \mathsf{Affects}.\{\mathsf{johnSmith}\}$$

Additional axioms: Role hierarchy

Syntax: We can add an RBox $\mathcal{R}$, and for every R and S, we can add an axiom $R \sqsubseteq S$ to represent the fact that R is subsumed by S.

Let $\mathcal{DL}$ be a description logic. A $\mathcal{DLH}$ ontology is a $\mathcal{DL}$ ontology that may contain an RBox with axioms of the form $R \sqsubseteq S$, where S and R roles in $\mathcal{DL}$.

Semantics:

$$\mathcal{I} \models R \sqsubseteq S \text{ iff } R^{\mathcal{I}} \subseteq S^{\mathcal{I}}$$

Example:

hasSister $\sqsubseteq$ hasSibling ; hasParent $\sqsubseteq$ hasAncestor

Additional roles: Chain roles

Syntax: We can add an RBox $\mathcal{R}$, and for every two roles R and S, we can also have the role $R \circ S$ that represents the composition of role R with role S.

Semantics:

$$(R \circ S)^{\mathcal{I}} = \{(x, z) \mid (x, y) \in R^{\mathcal{I}} \wedge (y, z) \in S^{\mathcal{I}}\}$$

Example:

$$\text{hasUncle} \sqsubseteq \text{hasMother} \circ \text{hasBrother}$$

Capturing transitive ($\mathcal{S}$) roles:

$$\text{hasAncestor} \circ \text{hasAncestor} \sqsubseteq \text{hasAncestor}$$

More

'Self' concepts:

$$(\exists R.Self)^{\mathcal{I}} = \{x \mid (x, x) \in R^{\mathcal{I}}\}$$

Universal role:

$$U^{\mathcal{I}} = \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$$

Identitiy ABox axioms:

$$\mathcal{I} \models a = b \text{ iff } a^{\mathcal{I}} = b^{\mathcal{I}}$$

$$\mathcal{I} \models a \neq b \text{ iff } a^{\mathcal{I}} \neq b^{\mathcal{I}}$$

Partial summary

$\mathcal{DL}$	C	R	axioms
$\mathcal{ALC}$	C, D, $\neg C$, $C \sqcap D$, $\exists R.C$	R	$C \sqsubseteq D$, $C(a)$, $R(a,b)$, $\neg R(a,b)$
$\mathcal{ALCQ}$	C, D, $\neg C$, $C \sqcap D$, $\exists R.C$, $(\bowtie n\ R.C)$	R	$C \sqsubseteq D$, $C(a)$, $R(a,b)$, $\neg R(a,b)$
$\mathcal{ALCI}$	C, D, $\neg C$, $C \sqcap D$, $\exists R.C$	R, R^-	$C \sqsubseteq D$, $C(a)$, $R(a,b)$, $\neg R(a,b)$
$\mathcal{ALCO}$	C, D, $\neg C$, $C \sqcap D$, $\exists R.C$, $\{a\}$	R	$C \sqsubseteq D$, $C(a)$, $R(a,b)$, $\neg R(a,b)$
$\mathcal{ALCH}$	C, D, $\neg C$, $C \sqcap D$, $\exists R.C$	R, S	$C \sqsubseteq D$, $C(a)$, $R(a,b)$, $\neg R(a,b)$, $R \sqsubseteq S$
$\mathcal{ALCF}$	C, D, $\neg C$, $C \sqcap D$, $\exists R.C$	R	$C \sqsubseteq D$, $C(a)$, $R(a,b)$, $\neg R(a,b)$, $\top \sqsubseteq (\leq 1\ R.\top)$
$\mathcal{ALCS}$	C, D, $\neg C$, $C \sqcap D$, $\exists R.C$	R	$C \sqsubseteq D$, $C(a)$, $R(a,b)$, $\neg R(a,b)$, $R \circ R \sqsubseteq R$
...			
$\mathcal{SROIQ}$[5]	C, D, $\neg C$, $C \sqcap D$, $\exists R.C$, $(\bowtie n\ R.C)$, $\{a\}$, $\exists R.Self$	R, S, R^-, $R \circ \ldots \circ S$, U	$C \sqsubseteq D$, $C(a)$, $R(a,b)$, $\neg R(a,b)$, $a = b$, $a \neq b$, $R \sqsubseteq S$
... and more			

[5] Ian Horrocks, Oliver Kutz, and Ulrike Sattler. "The Even More Irresistible SROIQ". In: *KR 2006*. 2006.

Some complexity results

See `http://www.cs.man.ac.uk/~ezolin/dl/`.

Complexity of reasoning problems:

$\mathcal{DL}$	complexity
$\mathcal{ALC}$	EXPTIME-complete
$\mathcal{ALCQ}$	EXPTIME-complete
$\mathcal{ALCI}$	EXPTIME-complete
$\mathcal{ALCO}$	EXPTIME-complete
$\mathcal{ALCH}$	EXPTIME-complete
$\mathcal{ALCF}$	EXPTIME-complete
$\mathcal{ALCS}$	EXPTIME-complete
$\mathcal{ALCHOQ}$	EXPTIME-complete
$\mathcal{ALCOIQ}$	NEXPTIME-complete
$\mathcal{ALCHOIQ}$	NEXPTIME-complete
$\mathcal{ALCSHOIQ} = \mathcal{SROIQ}$	N2EXPTIME-complete
...	

Outline

The need for light-weight Description Logics

The complexity of the extensions of $\mathcal{ALC}$ are still very hard computationally to reason with.

Some ontologies have a very large TBox (e.g., medical domain, SNOWMED)

Often we want to reason with a very large amount of data (e.g., diagnostic of monitoring data).

$\mathcal{EL}$, $\mathcal{EL}^{++}$

$\mathcal{EL}$ is the fragment of $\mathcal{ALC}$ without $\sqcup$ and $\neg$:

$$C ::= A \mid C \sqcap C \mid \exists R.C$$

$\mathcal{EL}^{++}$ is the minimal $\mathcal{DL}$ obtained by adding to $\mathcal{EL}$:

- $\top$ and $\bot$ concepts;
- a concept $\{a\}$ for every named individual a;
- an RBox, that can contain any $R_1 \circ \ldots \circ R_n \sqsubseteq S$.

Reasoning with $\mathcal{EL}^{++}$ is PTIME-complete [Baader, Brandt and Lutz 2005][6].

[6] Franz Baader, Sebastian Brandt, and Carsten Lutz. "Pushing the EL Envelope". In: *IJCAI-05*.

$\mathcal{RL}$

Let R be a role name, A be a concept name, a and b be two named individuals.

Subconcepts:

$$B ::= A \mid B \sqcap B \mid B \sqcup B \mid \exists R.B$$

Superconcepts:

$$C ::= A \mid \bot \mid C \sqcap C \mid \forall R.C$$

TBox:

$$B \sqsubseteq C$$

ABox:

$$A(a) \quad R(a, b)$$

$\mathcal{RL}$ reasoning can be captured by rule-based reasoning in Datalog (forward and backward chaining techniques), a subset of Prolog.

Knowledge base satisfiability with $\mathcal{RL}$ is in PTIME.

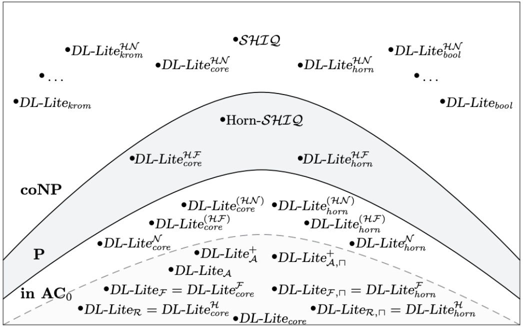
Allowed in $\mathcal{RL}$

- Subtyping:
$$\mathsf{Arthritis} \sqsubseteq \mathsf{Disease}$$
- Definitions:
$$\mathsf{JuvArthritis} \equiv \mathsf{JuvDisease} \sqcap \mathsf{Arthritis}$$
- Disjointness:
$$\mathsf{Child} \sqcap \mathsf{Adult} \sqsubseteq \bot$$
- Covering:
$$\mathsf{Person} \sqsubseteq \mathsf{Child} \sqcup \mathsf{Teenager} \sqcup \mathsf{Adult}$$
- Domain restriction:
$$\exists \mathsf{Affects}.\top \sqsubseteq \mathsf{Disease}$$
- Range restriction:
$$\top \sqsubseteq \forall \mathsf{Affects}.\mathsf{LivingThing}$$

$\mathcal{DL}$-Lite family

[Calvanese et al. 2007][7]

The $\mathcal{DL}$-lite extended family (grouped according to the data complexity of "positive existential query answering under the unique name assumption") [Artale et al. 2014][8]:



[7] Diego Calvanese et al. "Tractable Reasoning and Efficient Query Answering in Description Logics: The *DL-Lite* Family". In: *J. Autom. Reason.* 39.3 (2007), pp. 385–429.

[8] Alessandro Artale et al. "The DL-Lite Family and Relations". In: *CoRR* abs/1401.3487 (2014).

Let R be a role name, A be a concept name, a and b be two named individuals.

Subconcepts:

$$B ::= A \mid \exists P.\top$$

Superconcepts:

$$C ::= A \mid \neg A \mid A \sqcap A \mid \exists P.\top \mid \exists P.C$$

Roles:

$$P ::= R \mid R^-$$

TBox:

$$B \sqsubseteq C$$

RBox:

$$P \sqsubseteq P \qquad Dis(P, P)$$

ABox:

$$A(a) \qquad R(a, b)$$

Disallowed in $\mathcal{DL}$-Lite$_\mathcal{R}$

- Definitions:

$$\text{JuvArthritis} \equiv \text{JuvDisease} \sqcap \text{Arthritis}$$

- Covering:

$$\text{Person} \sqsubseteq \text{Child} \sqcup \text{Teenager} \sqcup \text{Adult}$$

- Transitive role:

$$\text{hasAncestor} \circ \text{hasAncestor} \sqsubseteq \text{hasAncestor}$$

- Cardinality restrictions:

$$\text{SevereArthritis} \sqsubseteq \text{Arthritis} \sqcap (\geq 5\, \text{Damages}.\text{Joint})$$

- Functional role:

$$\top \sqsubseteq (\leq 1\, \text{hasSSN}.\top)$$

- ...

Allowed in $\mathcal{DL}$-Lite$_{\mathcal{R}}$

- Subtyping:

$$\text{Arthritis} \sqsubseteq \text{Disease}$$

- Concept disjointness:

$$\text{Child} \sqsubseteq \neg\text{Adult}$$

- Role disjointness:

$$Dis(\text{hasChild}, \text{hasMother})$$

- Domain restriction:

$$\exists\text{Affects}.\top \sqsubseteq \text{Disease}$$

- Range restriction:

$$\exists\text{Affects}^{-}.\top \sqsubseteq \text{LivingThing}$$

- Symmetric role:

$$\text{hasSibling} \sqsubseteq \text{hasSibling}^{-}$$

- ...

Problem	Parameter	Computational complexity
satisfiability	size of knowledge base	NLOGSPACE
instance checking	size of data	AC^0
query answering	size of data	AC^0

Summary

- $\mathcal{EL}^{++}$ is good when focusing on TBox reasoning; tractable satsifiability.
- $\mathcal{RL}$ is good when focusing on ABox rule-based reasoning.
- $\mathcal{DL}$-Lite$_{\mathcal{R}}$ is good for query answering with limited TBox/RBox expressivity and large ABox.

$\mathcal{EL}^{++}$, $\mathcal{RL}$, and $\mathcal{DL}$-Lite$_{\mathcal{R}}$ correspond to OWL 2 profiles, as defined by the World Wide Web Consortium (W3C).

`https://www.w3.org/TR/owl2-profiles/`

They are (part of) of what we look at next.

Credits

Many slides and examples based on Ian Horrocks's KRR lectures
`https://www.cs.ox.ac.uk/people/ian.horrocks/`.
`https://www.cs.ox.ac.uk/teaching/courses/2020-2021/KRR/`