

Introduction to  
Logic in Computer Science: Autumn 2006  
Description Logics Crash Course

Andreas Witzel

Institute for Logic, Language and Computation  
University of Amsterdam

# Why Description Logics?

The goal of AI is to create systems with intelligent behaviour.

Crucial for this is

- ▶ acquiring knowledge about the world / application domain
- ▶ representing the knowledge
- ▶ reasoning with the knowledge

Conclusion: We need a logic formalism to represent knowledge and facilitate reasoning.

# Outline

In this lecture, we will introduce one example of such a **description logic** formalism, see how to reason with it, and discuss some properties and modifications. The topics are:

The Attributive Language with Complement ( $\mathcal{ALC}$ )

Reasoning with  $\mathcal{ALC}$

$\mathcal{ALC}$  with Number Restrictions ( $\mathcal{ALCN}$ )

Computational Issues

Summary

# Outline

## The Attributive Language with Complement ( $\mathcal{ALC}$ )

Syntax

Semantics

Terminology (TBox)

Assertions (ABox)

Reasoning with  $\mathcal{ALC}$

$\mathcal{ALC}$  with Number Restrictions ( $\mathcal{ALCN}$ )

Computational Issues

Summary

# Syntax of $\mathcal{ALC}$

Let  $N_C$  and  $N_R$  be disjoint sets of **concept names** and **role names**.

$\mathcal{ALC}$  **concept terms** are defined inductively:

1. Each concept name  $A \in N_C$  is an  $\mathcal{ALC}$  concept term
2. If  $C, D$  are  $\mathcal{ALC}$  concept terms and  $r \in N_R$  is a role name, then the following are also  $\mathcal{ALC}$  concept terms:
  - ▶  $\perp, \top$  (Bottom, Top)
  - ▶  $C \sqcap D, C \sqcup D, \neg C$  (Boolean Operators)
  - ▶  $\forall r.C, \exists r.C$  (value restriction and existential restriction)

## Example

- ▶  $\text{Human} \sqcap \neg \text{Female}$  describes “male”
- ▶  $\exists \text{has-child}.\top$  describes “parent”
- ▶  $\text{Human} \sqcap \forall \text{has-child}.\text{Female}$  describes “humans having only daughters”

## Semantics of $\mathcal{ALC}$

An **interpretation**  $\mathcal{J}$  over a non-empty domain  $\Delta^{\mathcal{J}}$  assigns

- ▶ to each concept name  $A \in N_C$  a subset  $A^{\mathcal{J}} \subseteq \Delta^{\mathcal{J}}$
- ▶ to each role name  $r \in N_R$  a binary relation  $r^{\mathcal{J}} \subseteq \Delta^{\mathcal{J}} \times \Delta^{\mathcal{J}}$

and is inductively extended to all  $\mathcal{ALC}$  concept terms:

$$\perp^{\mathcal{J}} := \emptyset$$

$$\top^{\mathcal{J}} := \Delta^{\mathcal{J}}$$

$$(\neg C)^{\mathcal{J}} := \Delta^{\mathcal{J}} \setminus C^{\mathcal{J}}$$

$$(C \sqcap D)^{\mathcal{J}} := C^{\mathcal{J}} \cap D^{\mathcal{J}}$$

$$(C \sqcup D)^{\mathcal{J}} := C^{\mathcal{J}} \cup D^{\mathcal{J}}$$

$$(\forall r.C)^{\mathcal{J}} := \left\{ d \in \Delta^{\mathcal{J}} \mid \forall e \in \Delta^{\mathcal{J}}. (d, e) \in r^{\mathcal{J}} \Rightarrow e \in C^{\mathcal{J}} \right\}$$

$$(\exists r.C)^{\mathcal{J}} := \left\{ d \in \Delta^{\mathcal{J}} \mid \exists e \in \Delta^{\mathcal{J}}. (d, e) \in r^{\mathcal{J}} \wedge e \in C^{\mathcal{J}} \right\}$$

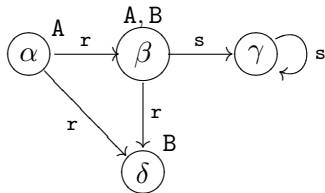
# Semantics of $\mathcal{ALC}$

## Example

For a language with concept names A, B and role names r, s, we consider an example interpretation  $\mathcal{I}$  with  $\Delta^{\mathcal{I}} := \{\alpha, \beta, \gamma, \delta\}$  and

$$A^{\mathcal{I}} := \{\alpha, \beta\} \quad r^{\mathcal{I}} := \{(\alpha, \beta), (\alpha, \delta), (\beta, \delta)\}$$

$$B^{\mathcal{I}} := \{\beta, \delta\} \quad s^{\mathcal{I}} := \{(\beta, \gamma), (\gamma, \gamma)\}$$



$$(\neg A)^{\mathcal{I}} = \{\gamma, \delta\}$$

$$(\forall r.A)^{\mathcal{I}} = \{\gamma, \delta\}$$

$$(\neg A \sqcap B)^{\mathcal{I}} = \{\delta\}$$

$$(\exists s.\neg A)^{\mathcal{I}} = \{\beta, \gamma\}$$

$$(\forall r.B)^{\mathcal{I}} = \{\alpha, \beta, \gamma, \delta\}$$

$$(\exists s.A)^{\mathcal{I}} = \emptyset$$

## Terminology (TBox): Syntax

In a realistic setting there will commonly be **complex concepts**, i.e. concepts built from simpler ones. We want to be able to give them abbreviating names, in this way defining the **terminology** of the setting.

A **TBox** allows us to do just that. It contains (finitely many) pairs of concept names and complex concept terms defining them.

### Example (TBox)

$$\left[ \begin{array}{l} \text{Male} \doteq \neg \text{Female} \\ \text{Woman} \doteq \text{Human} \sqcap \text{Female} \\ \text{Man} \doteq \text{Human} \sqcap \text{Male} \\ \text{Mother} \doteq \text{Woman} \sqcap \exists \text{has-child.Human} \\ \text{Father} \doteq \text{Man} \sqcap \exists \text{has-child.Human} \end{array} \right]$$

**Note:** Multiple definitions and cycles are not allowed!



## Terminology (TBox): Semantics

An interpretation  $\mathcal{J}$  is a **model of a TBox  $\mathcal{T}$**  iff we have:

$$A^{\mathcal{J}} = C^{\mathcal{J}} \quad \text{for all } A \doteq C \in \mathcal{T}$$

Two TBoxes are **equivalent** iff they have the same models.

For every TBox  $\mathcal{T}$  there is an equivalent **unfolded TBox  $\hat{\mathcal{T}}$**  where only primitive concept names (i.e. names which are not themselves being defined in  $\hat{\mathcal{T}}$ ) occur on right-hand sides.

An interpretation of the primitive concept names and role names in  $\mathcal{T}$  can be **uniquely extended** to a model of  $\mathcal{T}$ .

### Example

$$\left[ \begin{array}{l} \text{Woman} \doteq \text{Human} \sqcap \text{Female} \\ \text{Mother} \doteq \text{Human} \sqcap \text{Female} \sqcap \exists \text{has-child.Human} \end{array} \right]$$

**Note:** This may result in exponential blowup!

## Assertions (ABox): Syntax

Having fixed the terminology of a setting, we may want to assign names to individuals and make assertions about them.

That is what an **ABox** is used for. It contains (again finitely many) assertions over **individual names**  $N_I$  (disjoint from  $N_C$  and  $N_R$ ).

- $a : C$  (concept assertion)
- $(a, b) : r$  (role assertion)
- $a \neq b$  (distinct individuals)  $\rightsquigarrow$  generalized ABox

Example ( $N_I = \{\text{gunther, gundula, gisbert}\}$ )

$$\left\{ \begin{array}{l} \text{gunther} : \text{Man} \\ \text{gundula} : \text{Woman} \\ \text{gisbert} : \text{Man} \\ (\text{gunther}, \text{gisbert}) : \text{has-child} \\ (\text{gundula}, \text{gisbert}) : \text{has-child} \end{array} \right\}$$

## Assertions (ABox): Semantics

An interpretation  $\mathcal{I}$  now additionally assigns, to each individual name  $a \in N_I$ , an element  $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$ .

**Remark:** Normally, we make the **unique name assumption**, implicitly requiring that  $a^{\mathcal{I}} \neq b^{\mathcal{I}}$  for any two (unequal) individual names  $a$  and  $b$ . However, later on we need to use generalized ABoxes, where these statements are made explicitly.

An interpretation  $\mathcal{I}$  is a **model of an ABox  $\mathcal{A}$**  iff

$$\begin{array}{ll} a^{\mathcal{I}} \in C^{\mathcal{I}} & \text{for all } a : C \in \mathcal{A}, \text{ and} \\ (a^{\mathcal{I}}, b^{\mathcal{I}}) \in r^{\mathcal{I}} & \text{for all } (a, b) : r \in \mathcal{A} \end{array}$$

Given a TBox  $\mathcal{T}$  and an ABox  $\mathcal{A}$ , one is often interested in **common models** of  $\mathcal{T}$  and  $\mathcal{A}$ .

# Example

$$\left[ \begin{array}{l} \text{Male} \doteq \neg \text{Female} \\ \text{Woman} \doteq \text{Human} \sqcap \text{Female} \\ \text{Man} \doteq \text{Human} \sqcap \text{Male} \\ \text{Mother} \doteq \text{Woman} \sqcap \exists \text{has-child.Human} \\ \text{Father} \doteq \text{Man} \sqcap \exists \text{has-child.Human} \end{array} \right]$$

$$\left\{ \begin{array}{l} \text{gunther} : \text{Man} \\ \text{gundula} : \text{Woman} \\ \text{gisbert} : \text{Man} \\ (\text{gunther}, \text{gisbert}) : \text{has-child} \\ (\text{gundula}, \text{gisbert}) : \text{has-child} \end{array} \right\}$$

$$\Delta^J := \text{Human}^J := \{ind_1, ind_2, ind_3\}$$

$$\text{Female}^J := \{ind_2\}$$

$$\text{has-child}^J := \{(ind_1, ind_3), (ind_2, ind_3)\}$$

$$\text{gunther}^J := ind_1$$

$$\text{gundula}^J := ind_2$$

$$\text{gisbert}^J := ind_3$$

This uniquely defines the remaining concepts:

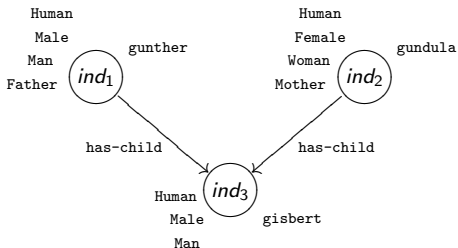
$$\text{Male}^J = \{ind_1, ind_3\}$$

$$\text{Woman}^J = \{ind_2\}$$

$$\text{Man}^J = \{ind_1, ind_3\}$$

$$\text{Mother}^J = \{ind_2\}$$

$$\text{Father}^J = \{ind_1\}$$



# Outline

The Attributive Language with Complement ( $\mathcal{ALC}$ )

Reasoning with  $\mathcal{ALC}$

Reasoning Tasks

Tableau Rules

$\mathcal{ALC}$  with Number Restrictions ( $\mathcal{ALCN}$ )

Computational Issues

Summary

# Terminological Reasoning Tasks

**Purpose:** Extract implicit terminological knowledge from explicitly given one

**Terminological reasoning:**

- ▶ *Satisfiability*: is there an interpretation  $\mathcal{I}$  with  $C^{\mathcal{I}} \neq \emptyset$ ?
- ▶ *Satisfiability wrt  $\mathcal{T}$* : is there a model  $\mathcal{I}$  of  $\mathcal{T}$  with  $C^{\mathcal{I}} \neq \emptyset$ ?
- ▶ *Subsumption ( $C \sqsubseteq D$ )*: is  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$  in all interpretations  $\mathcal{I}$ ?
- ▶ *Subsumption wrt  $\mathcal{T}$  ( $C \sqsubseteq_{\mathcal{T}} D$ )*: is  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$  in all models  $\mathcal{I}$  of  $\mathcal{T}$ ?

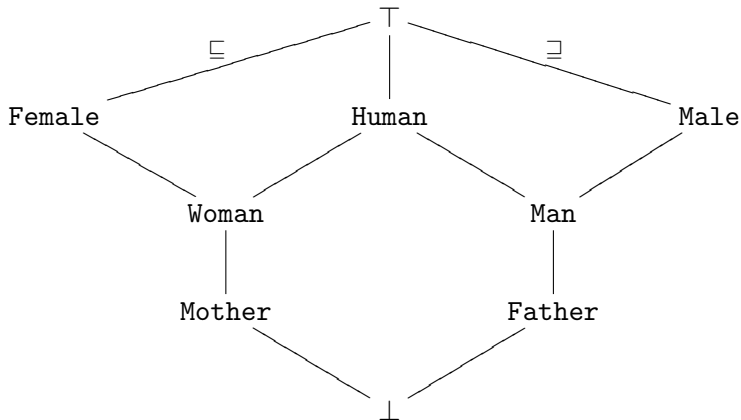
## Example

With respect to the TBox from the earlier examples, we have

- ▶  $\text{Mother} \sqcap \text{Man}$  is not satisfiable (it can be unfolded to  $\dots \sqcap \text{Female} \sqcap \dots \sqcap \neg \text{Female} \sqcap \dots$ )
- ▶  $\text{Mother} \sqcap \exists \text{has-child.} \neg \text{Human}$  is satisfiable (this would be a Woman with a Human and a  $\neg$ Human child...)
- ▶  $\text{Mother} \sqsubseteq_{\mathcal{T}} \text{Human}$
- ▶  $\text{Man} \not\sqsubseteq_{\mathcal{T}} \text{Father}$

## TBox Classification

Description Logic systems usually offer various services, for example computing the **concept hierarchy** (wrt the subsumption relation) of all occurring concept names.



# Assertional Reasoning Tasks

**Purpose:** Extract implicit knowledge about individuals from explicitly given one

**Assertional reasoning:**

- ▶ *Consistency*: does  $\mathcal{A}$  have a model?
- ▶ *Consistency wrt  $\mathcal{T}$* : do  $\mathcal{A}$  and  $\mathcal{T}$  have a common model?
- ▶ *Instance wrt  $\mathcal{A}$  ( $\mathcal{A} \models a : C$ )*: is  $a^{\mathcal{J}} \in C^{\mathcal{J}}$  in all models  $\mathcal{J}$  of  $\mathcal{A}$ ?
- ▶ *Instance wrt  $\mathcal{A}$  and  $\mathcal{T}$  ( $\mathcal{T}, \mathcal{A} \models a : C$ )*: is  $a^{\mathcal{J}} \in C^{\mathcal{J}}$  in all common models  $\mathcal{J}$  of  $\mathcal{A}$  and  $\mathcal{T}$ ?

## Example

With respect to  $\mathcal{A}$  and  $\mathcal{T}$  from the earlier examples, we have

- ▶  $\mathcal{A}$  is consistent and consistent wrt  $\mathcal{T}$
- ▶  $\mathcal{A} \models \text{gundula} : \text{Woman}$
- ▶  $\mathcal{A} \not\models \text{gundula} : \text{Human}$
- ▶  $\mathcal{T}, \mathcal{A} \models \text{gundula} : \text{Human}$



# Realization and Retrieval

Services offered by Description Logic systems in the context of ABoxes include:

- ▶ **Realization** of an individual name  $a$ : What are the (minimal) concept names of which  $a$  is an instance?

$\text{gunther} \rightsquigarrow \text{Father, Man, Human, Male}$  (all)

$\text{gunther} \rightsquigarrow \text{Father}$  (minimal)

- ▶ **Retrieval**: Which individuals are instances of  $C$ ?

$\text{Mother} \rightsquigarrow \text{gundula}$

$\text{Man} \sqcap \neg\text{Father} \rightsquigarrow \text{gisbert}$

## Reducing Reasoning Tasks

Instead of writing reasoners for each single reasoning task, one can try to **reduce** some tasks to other ones. For example,

*Is concept  $C$  satisfiable?*

can be reduced as follows:

$\Leftrightarrow$  *Is there a model  $\mathcal{J}$  such that  $C^{\mathcal{J}}$  is non-empty?*

$\Leftrightarrow$  *Is it not the case that for all models  $\mathcal{J}$ ,  $C^{\mathcal{J}}$  is empty?*

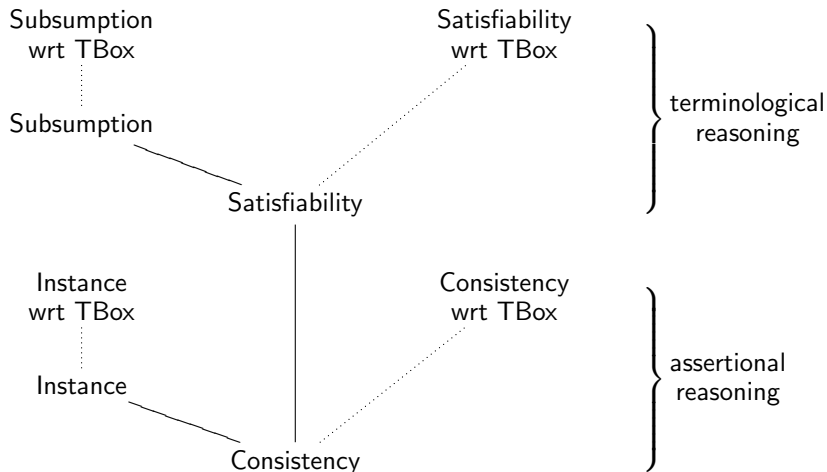
$\Leftrightarrow$  *Is concept  $C$  not subsumed by the empty concept  $\perp$ ?*

That is,  $C$  is satisfiable iff  $C \not\sqsubseteq \perp$ .

Similarly,

- ▶  $C \sqsubseteq D$  iff  $C \sqcap \neg D$  is unsatisfiable
- ▶  $C$  is satisfiable iff the ABox  $\{a : C\}$  is consistent
- ▶  $\mathcal{A} \models a : C$  iff  $\mathcal{A} \cup \{a : \neg C\}$  is inconsistent

# The Reduction Hierarchy



# Tableau Algorithm

We present a tableau algorithm to decide **ABox consistency**.

Given an ABox  $\mathcal{A}$  with unique name assumption, we

- ▶ drop the unique name assumption and add  $a \neq b$  for all  $a, b \in N_I$
- ▶ transform the resulting ABox into Negation Normal Form (i.e. negation occurs only in front of concept names)

We then exhaustively apply the **tableau rules** on the following slide.  $\mathcal{A}$  is consistent iff there remains an open branch. As with earlier Tableau systems, such a branch can be used to **build a model** of the ABox.

In order to guarantee **termination**, rules cannot be applied if the resulting formulas are already on the branch under consideration (note the special case with  $\exists$ ).

# Tableau Rules

$\sqcap$  and  $\sqcup$  rules:

$$\frac{a : C \sqcap D}{\begin{array}{l} a : C \\ a : D \end{array}}$$

$$\frac{a : C \sqcup D}{\begin{array}{l|l} a : C & a : D \end{array}}$$

$\forall$  and  $\exists$  rules:

$$\frac{\begin{array}{l} (a, b) : r \\ a : \forall r.C \end{array}}{b : C}$$

$$\frac{a : \exists r.C}{\begin{array}{l} (a, c) : r \\ c : C \end{array}}$$

Closure rules:

$$\frac{\begin{array}{l} a : C \\ a : \neg C \end{array}}{\times}$$

$$\frac{a \neq a}{\times}$$

In the  $\exists$  rule,  $c$  is a new individual name unused in  $\mathcal{A}$ .

Furthermore, the  $\exists$  rule is not applicable if there is any individual name  $b$  such that  $(a, b) : r$  and  $b : C$  are already on the branch.

# Outline

The Attributive Language with Complement ( $\mathcal{ALC}$ )

Reasoning with  $\mathcal{ALC}$

$\mathcal{ALC}$  with Number Restrictions ( $\mathcal{ALCN}$ )

Syntax and Semantics

Tableau Rules

Computational Issues

Summary

## Number Restrictions: Syntax and Semantics

Sometimes it is important not only to make statements about existence of roles, but also about their quantity.  $\mathcal{ALCN}$  introduces **number restrictions** for any role name  $r \in N_R$ :

- ▶  $(\geq n r)$  (at least restriction)

Semantics:  $(\geq n r)^J := \{d \in \Delta^J \mid |\{e \in \Delta^J \mid (d, e) \in r^J\}| \geq n\}$

- ▶  $(\leq n r)$  (at most restriction)

Semantics:  $(\leq n r)^J := \{d \in \Delta^J \mid |\{e \in \Delta^J \mid (d, e) \in r^J\}| \leq n\}$

$\mathcal{ALCN}$  **concept terms** are all concept terms built analogously to  $\mathcal{ALC}$  concept terms, where now additionally these two new basic concept terms can be used.

### Example

- ▶  $(\leq 3 \text{ has-child}) \sqcap \text{Woman}$  describes “mother of at most 3 children”
- ▶  $\exists \text{has-child} . (\geq 2 \text{ has-child})$  describes “grandparent of at least 2 siblings”

## Number Restrictions: Tableau Rules

The tableau rules for number restrictions are a bit more involved. In particular, the  $\leq$  rule requires renaming of individuals along the branch, which isn't really compatible with our notation. It also requires constraints on the rule application order to preserve termination.

For these reasons, we only consider the  $\geq$  rule here.

$\geq$  rule:

$$\frac{a : (\geq n r)}{(a, c_1) : r}$$

...

$$(a, c_n) : r$$
$$c_i \neq c_j \quad (1 \leq i < j \leq n)$$

Again, the  $c_i$  are new individual names unused in  $\mathcal{A}$ . Furthermore, the rule is not applicable if there are individual names  $b_1, \dots, b_n$  such that all  $(a, b_i) : r$  and all  $b_i \neq b_j$  are already on the branch.



# Outline

The Attributive Language with Complement ( $\mathcal{ALC}$ )

Reasoning with  $\mathcal{ALC}$

$\mathcal{ALC}$  with Number Restrictions ( $\mathcal{ALCN}$ )

**Computational Issues**

$\mathcal{ALC}$  is hard

Useful Restrictions

Summary

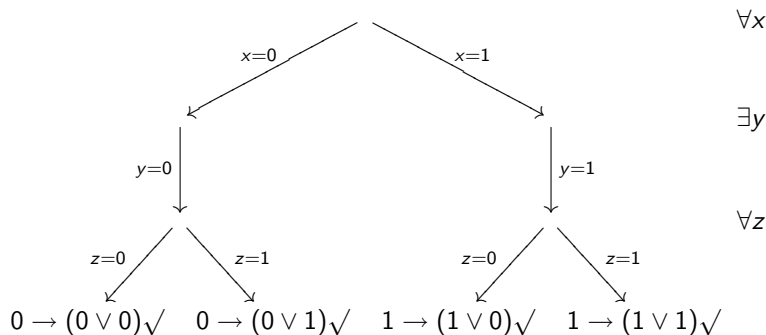
# Quantified Boolean Formulas (QBF)

**Idea:** Reduce QBF validity to  $\mathcal{ALC}$  concept satisfiability, thus showing **PSPACE**-hardness ('in **PSPACE**' also holds, but not shown here).

**Reminder from last lecture:** A QBF is a propositional formula preceded by either  $\forall x$  or  $\exists x$  for each occurring propositional variable  $x$ . For example:

$$\forall x \exists y \forall z. (x \rightarrow (y \vee z))$$

A QBF is valid iff it has a quantifier tree:



## Reducing QBF validity to $\mathcal{ALC}$ concept satisfiability

**Idea:** Use  $\mathcal{ALC}$  concepts to describe quantifier trees

**Given:** QBF  $Q = Q_1x_1 \dots Q_nx_n.\varphi$ .

**Find:**  $C_Q$  which is satisfiable iff  $Q$  has a quantifier tree

**Concepts:**  $X_1, \dots, X_n$  ( $\hat{=}$  the  $x_i$ )

$L_1, \dots, L_n$  ( $\hat{=}$  nodes of level  $i$  in the tree)

**Roles:**  $r$  ( $\hat{=}$  edges from nodes to children)

We define  $C_Q := L_1 \sqcap \forall r.(L_2 \sqcap \forall r.(L_3 \sqcap \dots \forall r.(L_n \sqcap \forall r.C_\varphi) \dots))$ ,  
where:

- ▶  $L_i := D_i \sqcap \begin{cases} \exists r.\top & \text{if } Q_i = \exists \\ \exists r.X_i \sqcap \exists r.\neg X_i & \text{if } Q_i = \forall \end{cases}$
- ▶  $D_i := \prod_{j < i} (X_j \Rightarrow \forall r.X_j) \sqcap (\neg X_j \Rightarrow \forall r.\neg X_j)$
- ▶  $Y \Rightarrow Z := \neg Y \sqcup Z$
- ▶  $C_\varphi$  is obtained from  $\varphi$  by replacing all  $x_i$  by  $X_i$ ,  $\wedge$  by  $\sqcap$ ,  $\vee$  by  $\sqcup$

## Useful Restrictions

If we want to guarantee reasoning tasks to be tractable, we can consider **sub-Boolean** fragments of  $\mathcal{ALC}$ . These typically allow for conjunction, but prohibit disjunction and/or negation.

Two examples for such logics are:

- ▶  $\mathcal{FL}_0$ , featuring  $\sqcap$ ,  $\forall$  and  $\top$
- ▶  $\mathcal{EL}$ , featuring  $\sqcap$ ,  $\exists$  and  $\top$

**Satisfiability** for such logics without negation is often trivial (i.e. all concept terms are satisfiable). Therefore, we are now interested in other notions such as **least common subsumers** (LCS).

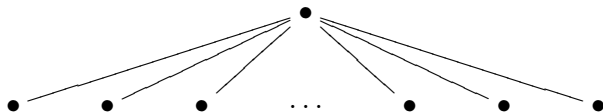
LCS are intended to describe the commonalities of concept terms. We call  $E$  the least common subsumer of  $C$  and  $D$  if

- (i)  $C \sqsubseteq E$  and  $D \sqsubseteq E$
- (ii)  $E \sqsubseteq F$  for all  $F$  with  $C \sqsubseteq F$  and  $D \sqsubseteq F$

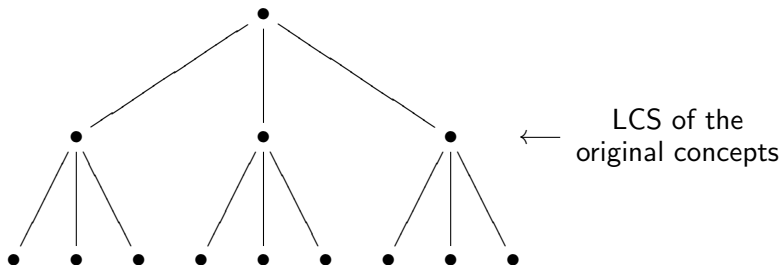
In  $\mathcal{FL}_0$  and  $\mathcal{EL}$ , the LCS always exists.

## Least Common Subsumer Application Example

The LCS can be used to **add more structure** to “flat” knowledge bases by introducing “meaningful” intermediate concepts into the concept hierarchy.



becomes



# Outline

The Attributive Language with Complement ( $\mathcal{ALC}$ )

Reasoning with  $\mathcal{ALC}$

$\mathcal{ALC}$  with Number Restrictions ( $\mathcal{ALCN}$ )

Computational Issues

**Summary**

## Summary

We have introduced and discussed Description Logics, in particular

- ▶ Syntax and Semantics of  $\mathcal{ALC}$  as an exemplary language
- ▶ TBoxes and ABoxes for terminology and assertions
- ▶ Common reasoning tasks
- ▶ Tableau rules
- ▶ Computational complexity of  $\mathcal{ALC}$
- ▶ An extension and two restrictions of  $\mathcal{ALC}$

Related topics include

- ▶ Historical approaches (semantic networks, frames, non-monotonic inheritance networks)
- ▶ First description logic system (KL-ONE)
- ▶ Modal Logic ( $\mathcal{ALC}$  is a syntactic variant of  $K_n$ )
- ▶ State of the art (OWL, FaCT, RACER, KAON 2)

Based on “Logic-based Knowledge Representation” by F. Baader.

<http://lat.inf.tu-dresden.de/teaching/ss2006/lbkr/>