Ontologies and Description Logic

Isabelle Bloch (with contributions of Natalia Diaz)

LIP6, Sorbonne Université - LTCI, Télécom Paris

isabelle.bloch@sorbonne-universite.fr, isabelle.bloch@telecom-paris.fr
Definition of an ontology

- In Philosophy: part of metaphysics, science of “being”. Studies concepts such as existence, being, becoming, and reality.

- In AI: part of knowledge engineering. A formal specification of a shared conceptualization (Gruber 1993), a formalism to define concepts, individuals, relationships and constraints (functions, attributes) within a domain.
Usefulness of ontologies (Charlet, 2002)

- Representation power (separate declarative & procedural knowledge)
  - Concepts: define aggregation of things
  - Individuals: instances of concepts
  - Properties (relationships): link concepts/individuals

- Logical reasoning capabilities: deduction, abduction, and subsumption. Most used language: OWL (web ontology language), based on description logics.

- Explainability: to extract a minimal set of covering models of interpretation from a knowledge base (KB) based on a set of observed actions, which could explain the observations.

- To represent and share knowledge by using a common vocabulary.
- To promote interoperability and knowledge reuse.
Description logics (DL)

- A family of formal logic-based knowledge representation formalisms tailored towards representing terminological knowledge of a domain in a structured and well-understood way.

- Notions (classes, relations, objects) of the domain are modelled using (atomic) concepts -unary predicates-, (atomic) roles -binary predicates-, and individuals:
  - to state constraints so that these notions can be interpreted
  - to deduce consequences (such as subclass and instance relationships from definitions and constraints).

- DLs differ from their predecessors (such as semantic networks and frames): they are equipped with a formal, logic-based semantics.
Why using DL in Knowledge Representation (KR)...

...rather than general first-order predicate logic (FOL)?

- Because it is a decidable fragment of FOL, therefore, amenable for automated reasoning\(^1\).

---

\(^1\)Decidability: Logics are decidable if computations/algorithms based on the logic will terminate in a finite time
- **TBox** (Terminological box): The vocabulary used to describe concept hierarchies and roles in the KB.
- **ABox** (Assertional box): States properties of individuals it correspond to in the KB (the data)
- Statements in TBox and ABox can be interpreted with DL rules and axioms to enable reasoning and inference (including satisfiability, subsumption, equivalence, disjointness, and consistency).
- **DL reasoning** supports decidability, completeness, and soundness.

\[
\text{Knowledge Base} = \text{TBox} + \text{ABox}
\]
TBox concept definition examples:

- **Men that are married to a doctor and all of whose children are either doctors or professors:** \( \text{HappyMan} \equiv \text{Human} \sqcap \neg \text{Female} \sqcap (\exists \text{married.Doctor}) \sqcap (\forall \text{hasChild.}(\text{Doctor} \sqcup \text{Professor})) \).

- **Only humans can have human children:** \( \exists \text{hasChild.}\text{Human} \sqsubseteq \text{Human} \)

ABox examples:

- \( \text{HappyMan(BOB)}, \text{hasChild(BOB, MARY)}, \neg \text{Doctor(MARY)} \)
Knowledge Base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, where $\mathcal{T}$ is a TBox and $\mathcal{A}$ is an ABox.

**Syntax:** atomic concepts and concept descriptions, atomic roles, constructors to build complex concepts and roles from atomic ones.

- **Concepts** correspond to classes.
- **Roles** are binary relations between objects.

**Semantics:** An interpretation $\mathcal{I}$ is a model of a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ ($\mathcal{I} \models \mathcal{K}$) if $\mathcal{I}$ is a model of $\mathcal{T}$ and $\mathcal{I}$ is a model of $\mathcal{A}$.

$\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$, where

- $\Delta^\mathcal{I}$ is a non empty set (domain of the interpretation)
- $\cdot^\mathcal{I}$ is an interpretation function that maps
  - each concept $C$ to a subset $C^\mathcal{I}$ of $\Delta^\mathcal{I}$
  - each role $r$ to a subset $R^\mathcal{I}$ of $\Delta^\mathcal{I} \times \Delta^\mathcal{I}$
Description logics syntax and interpretation:

<table>
<thead>
<tr>
<th>Constructor</th>
<th>Syntax</th>
<th>Example</th>
<th>Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>atomic concept</td>
<td>$A$</td>
<td>Human</td>
<td>$A^I \subseteq \Delta^I$</td>
</tr>
<tr>
<td>individual</td>
<td>$a$</td>
<td>Lea</td>
<td>$a^I \in \Delta^I$</td>
</tr>
<tr>
<td>Top</td>
<td>$\top$</td>
<td>Thing</td>
<td>$\top^I = \Delta^I$</td>
</tr>
<tr>
<td>Bottom</td>
<td>$\bot$</td>
<td>Nothing</td>
<td>$\bot^I = \emptyset^I$</td>
</tr>
<tr>
<td>atomic role</td>
<td>$r$</td>
<td>has-age</td>
<td>$\Delta^I \setminus C^I$</td>
</tr>
<tr>
<td>conjunction</td>
<td>$C \cap D$</td>
<td>Human $\cap$ Male</td>
<td>$C^I \cap D^I$</td>
</tr>
<tr>
<td>disjunction</td>
<td>$C \cup D$</td>
<td>Male $\cup$ Female</td>
<td>$C^I \cup D^I$</td>
</tr>
<tr>
<td>negation</td>
<td>$\neg C$</td>
<td>$\neg$ Human</td>
<td>$\Delta^I \setminus C^I$</td>
</tr>
<tr>
<td>existential restriction</td>
<td>$\exists r.C$</td>
<td>$\exists$ has-child.Girl</td>
<td>${x \in \Delta^I</td>
</tr>
<tr>
<td>universal restriction</td>
<td>$\forall r.C$</td>
<td>$\forall$ has-child.Human</td>
<td>${x \in \Delta^I</td>
</tr>
<tr>
<td>value restriction</td>
<td>$\exists r.{a}$</td>
<td>$\exists$ has-child.{Lea}</td>
<td>${x \in \Delta^I</td>
</tr>
<tr>
<td>number restriction</td>
<td>$(\geq nR)$</td>
<td>$(\geq 3$ has-child)</td>
<td>${x \in \Delta^I</td>
</tr>
<tr>
<td></td>
<td>$(\leq nR)$</td>
<td>$(\leq 1$ has-mother)</td>
<td>${x \in \Delta^I</td>
</tr>
<tr>
<td>Subsumption</td>
<td>$C \sqsubseteq D$</td>
<td>Man $\sqsubseteq$ Human</td>
<td>$C^I \subseteq D^I$</td>
</tr>
<tr>
<td>Concept definition</td>
<td>$C \equiv D$</td>
<td>Father $\equiv$ Man $\sqcap$ $\exists$ has-child.Human</td>
<td>$C^I = D^I$</td>
</tr>
<tr>
<td>Concept assertion</td>
<td>$a : C$</td>
<td>John:Man</td>
<td>$a^I \in C^I$</td>
</tr>
<tr>
<td>Role assertion</td>
<td>$(a, b) : R$</td>
<td>(John, Helen):has-child</td>
<td>$(a^I, b^I) \in R^I$</td>
</tr>
</tbody>
</table>
Example

Father $\equiv \neg\text{Female} \sqcap \exists\text{hasChild.\text{Human}}$

Interpretation $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$, with $\Delta^\mathcal{I} = \{\text{John, Mary}\}$

- $\text{Father}^\mathcal{I} = \{\text{John}\} \subseteq \Delta^\mathcal{I}$
- $\text{Human}^\mathcal{I} = \{\text{John, Mary}\}$
- $\text{hasChild}^\mathcal{I} = \{(\text{John, Mary})\}$
- $\exists\text{hasChild.\text{Human}}^\mathcal{I} = \{\text{John}\}$
Reasoning tasks

- Classification
- Retrieval
- Consistency checking
- Subsumption checking
- Satisfiability
- ...
Subsumption

• Superclass/subclass relationship, “isa”
• All members of a subclass can be inferred to be members of its superclasses

owl:Thing: superclass of all OWL Classes

• A subsumes B
• A is a superclass of B
• B is a subclass of A
• All members of B are also members of A

Defined explicitly or inferred by a reasoner
Subsumption $\mathcal{K} \models C_1 \subseteq C_2$: for all interpretations $\mathcal{I}$ such that $\mathcal{I} \models \mathcal{K}$, check $C_1^\mathcal{I} \subseteq C_2^\mathcal{I}$

Consistency

- of a concept: for all interpretations $\mathcal{I}$ such that $\mathcal{I} \models \mathcal{K}$, check $C^\mathcal{I} \neq \emptyset$
- of $\mathcal{K}$: there exists $\mathcal{I}$ such that $\mathcal{I} \models \mathcal{K}$

Instance checking $\mathcal{K} \models (a : C)$: $\forall \mathcal{I} s.t. \mathcal{I} \models \mathcal{K}$, $a^\mathcal{I} \in C^\mathcal{I}$

Relation checking $\mathcal{K} \models ((a, b) : R)$: $\forall \mathcal{I} s.t. \mathcal{I} \models \mathcal{K}$, $(a^\mathcal{I}, b^\mathcal{I}) \in R^\mathcal{I}$
Example:

Female ⊑ Human 
Child ⊑ Human 
Works ⊑ Human 
StudiesAtUni ⊑ Human 
SuccessfullMan ≡ ¬Female ⊓ InBusiness ⊓ ∃married.Lawyer ⊓ ∃child.(StudiesAtUni ⊔ Works) 
Pedro : ¬Female 
Pedro : InBusiness 
Mary : Lawyer 
John : Works 
(Pedro, Mary) : married 
(Pedro, John) : child

Is Pedro a successful man?
Relation with predicate logic

Translation function $\tau_x$ introducing a variable $x$:

- $\tau_x(C) = C(x)$
- $\tau_x(C \cap D) = \tau_x(C) \land \tau_x(D)$
- $\tau_x(C \cup D) = \tau_x(C) \lor \tau_x(D)$
- $\tau_x(\exists r.C) = \exists y, r(x, y) \land \tau_y(C)$
- $\tau_x(\forall r.C) = \forall y, r(x, y) \rightarrow \tau_y(C)$
- for all concept inclusions in the TBox:

$$\bigwedge_{C \sqsubseteq D \in TBox} \forall x (\tau_x(C) \rightarrow \tau_x(D))$$

($\sqsubseteq$ becomes logical implication)

- ABox: $(a : C)$ becomes $C(a)$, and $(a, b) : r$ becomes $r(a, b)$
Example: Prove that

$$\forall r. (A \cap B) \subseteq \forall r. A \cap \forall r. B$$

- using interpretations
- using translation into first order predicate logic
Applications:
- information retrieval,
- search, question answering,
- reasoning and decision support
- ...

Extensions
- fuzzy description logics
- knowledge graph (ontology as the underlying vocabulary)
- ...

Outlook
Some references

- J. Charlet. L’ingénierie des connaissances : développements, résultats et perspectives pour la gestion des connaissances médicales. HDR Université Pierre et Marie Curie - Paris VI, 2002. tel-00006920