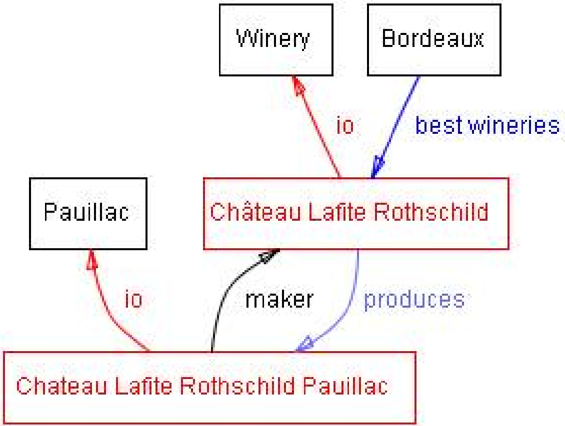
Ontologies and Description Logic

Isabelle Bloch (with contributions of Natalia Diaz)

LIP6, Sorbonne Université - LTCI, Télécom Paris

isabelle.bloch@sorbonne-universite.fr, isabelle.bloch@telecom-paris.fr


Winery
Bordeaux
io
best wineries
Pauillac
Château Lafite Rothschild
io
maker
produces
Chateau Lafite Rothschild Pauillac

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.

$$\boxed{\text{Knowledge Base} = \text{TBox} + \text{ABox}}$$

TBox concept definition examples:

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

ABox examples:

- $\texttt{HappyMan(BOB), hasChild(BOB, MARY),} \neg\ \texttt{Doctor(MARY)}$

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:

Constructor	Syntax	Example	Semantics
atomic concept	A	Human	$A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$
individual	a	Lea	$a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$
Top	$\top$	Thing	$\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$
Bottom	$\bot$	Nothing	$\bot^{\mathcal{I}} = \emptyset^{\mathcal{I}}$
atomic role	r	has-age	$R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
conjunction	$C \sqcap D$	Human $\sqcap$ Male	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
disjunction	$C \sqcup D$	Male $\sqcup$ Female	$C^{\mathcal{I}} \cup D^{\mathcal{I}}$
negation	$\neg C$	$\neg$ Human	$\Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$
existential restriction	$\exists r.C$	$\exists$has-child.Girl	$\{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}} : (x,y) \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$
universal restriction	$\forall r.C$	$\forall$has-child.Human	$\{x \in \Delta^{\mathcal{I}} \mid \forall y \in \Delta^{\mathcal{I}} : (x,y) \in R^{\mathcal{I}} \Rightarrow y \in C^{\mathcal{I}}\}$
value restriction	$\ni r.\{a\}$	$\ni$has-child.{Lea}	$\{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}} : (x,y) \in R^{\mathcal{I}} \Rightarrow y = a^{\mathcal{I}}\}$
number restriction	$(\geq nR)$	($\geq$ 3 has-child)	$\{x \in \Delta^{\mathcal{I}} \mid \ \lvert\{y \mid (x,y) \in R^{\mathcal{I}}\}\rvert \geq n\}$
	$(\leq nR)$	($\leq$ 1 has-mother)	$\{x \in \Delta^{\mathcal{I}} \mid \ \lvert\{y \mid (x,y) \in R^{\mathcal{I}}\}\rvert \leq n\}$
Subsumption	$C \sqsubseteq D$	Man $\sqsubseteq$ Human	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
Concept definition	$C \equiv D$	Father $\equiv$ Man $\sqcap$ $\exists$ has-child.Human	$C^{\mathcal{I}} = D^{\mathcal{I}}$
Concept assertion	$a : C$	John:Man	$a^{\mathcal{I}} \in C^{\mathcal{I}}$
Role assertion	$(a,b) : R$	(John,Helen):has-child	$(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$

Example

Father $\equiv \neg$Female $\sqcap \exists$hasChild.Human

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

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

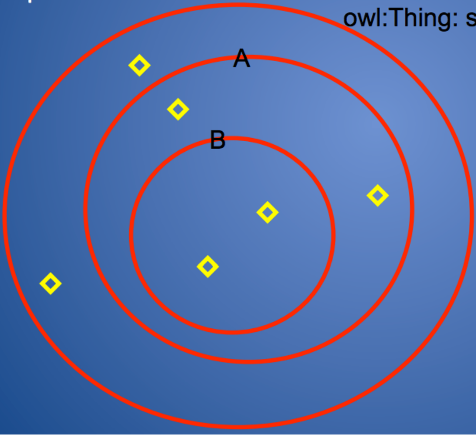
Reasoning tasks

- 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 $\mathcal{K} \models C_1 \sqsubseteq C_2$: for all interpretations $\mathcal{I}$ such that $\mathcal{I} \models \mathcal{K}$, check $C_1^{\mathcal{I}} \sqsubseteq C_2^{\mathcal{I}}$
- 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:

$\text{Female} \sqsubseteq \text{Human}$
$\text{Child} \sqsubseteq \text{Human}$
$\text{Works} \sqsubseteq \text{Human}$
$\text{StudiesAtUni} \sqsubseteq \text{Human}$
$\text{SuccessfullMan} \equiv \neg\text{Female} \sqcap \text{InBusiness} \sqcap \exists\text{married.Lawyer} \sqcap \exists\text{child.}(\text{StudiesAtUni} \sqcup \text{Works})$
$\text{Pedro} : \neg\text{Female}$
$\text{Pedro} : \text{InBusiness}$
$\text{Mary} : \text{Lawyer}$
$\text{John} : \text{Works}$
$(\text{Pedro}, \text{Mary}) : \text{married}$
$(\text{Pedro}, \text{John}) : \text{child}$

Is Pedro a successful man?

Relation with predicate logic

Translation function τ_x introducing a variable x:

- $\tau_x(C) = C(x)$
- $\tau_x(C \sqcap D) = \tau_x(C) \wedge \tau_x(D)$
- $\tau_x(C \sqcup D) = \tau_x(C) \vee \tau_x(D)$
- $\tau_x(\exists r.C) = \exists y, r(x, y) \wedge \tau_y(C)$
- $\tau_x(\forall r.C) = \forall y, r(x, y) \rightarrow \tau_y(C)$
- 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

$$\forall r.(A \sqcap B) \sqsubseteq \forall r.A \sqcap \forall r.B$$

- using interpretations
- using translation into first order predicate logic

Outlook

Applications:

- information retrieval,
- search, question answering,
- reasoning and decision support
- ...

Extensions

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

Some references

- F. Baader. The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press, 2003.
- F. Baader, I. Horrocks, and U. Sattler. Description logics. In F. van Harmelen, V. Lifschitz, and B. Porter, editors, Handbook of Knowledge Representation, pages 135-179. Elsevier, 2007.
- W. N. Borst. Construction of Engineering Ontologies for Knowledge Sharing and Reuse. PhD thesis, Institute for Telematica and Information Technology, University of Twente, Enschede, The Netherlands, 1997.
- 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
- T. R. Gruber. A translation approach to portable ontology specifications. Knowl. Acquis., 5(2):199-220, June 1993.
- U. Straccia. Foundations of Fuzzy Logic and Semantic Web Languages. CRC Studies in Informatics Series. Chapman & Hall, 2013.