Knowledge Representation and Learning

9. First Order Logic - Theories

Luciano Serafini

Fondazione Bruno Kessler

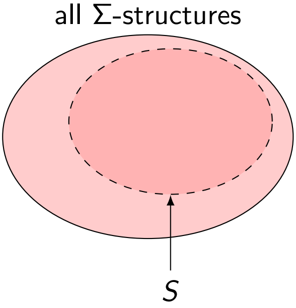
May 4, 2023

First order theories

- Mathematics focuses on the study of properties of certain structures. E.g. Natural/Rational/Real/Complex numbers, Algebras, Monoids, Lattices, Partially-ordered sets, Topological spaces, fields, ...
- In knowledge representation, mathematical structures can be used as a reference abstract model for a real world feature. e.g.,
  - natural/rational/real numbers can be used to represent linear time;
  - trees can be used to represent possible future evolutions;
  - graphs can be used to represent maps;
  - ...
- Logics provides a rigorous way to describe certain classes of mathematical structures.

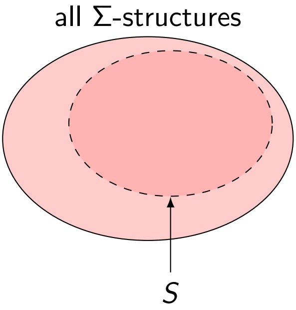
Focusing on a subclass of structures

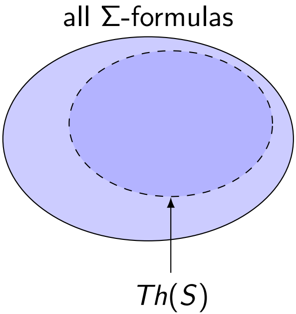
We are interested in considering only a subset of Σ-structures.



Focusing on a subclass of structures

We are interested in considering only a subset of Σ-structures.

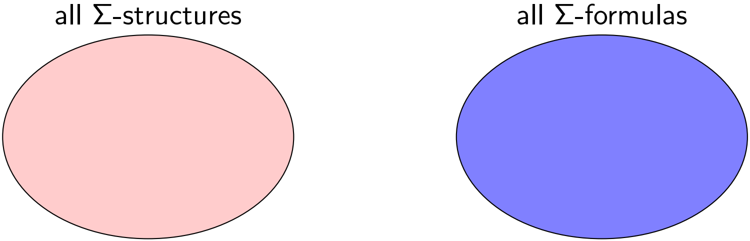




Theory of S

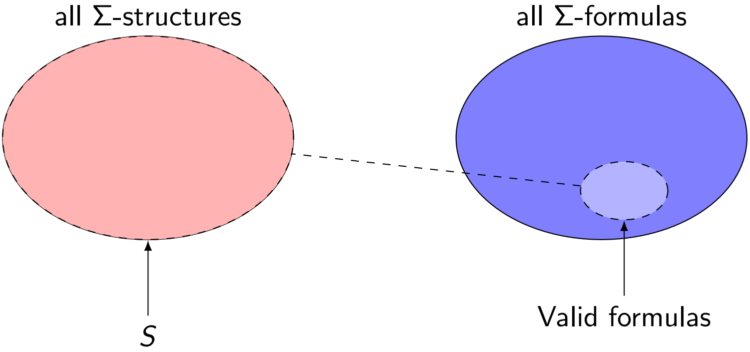
The theory of S is the set of formulas $Th(S)$ that are satisfied by all the Σ structures S.

Focusing on a subclass of structures



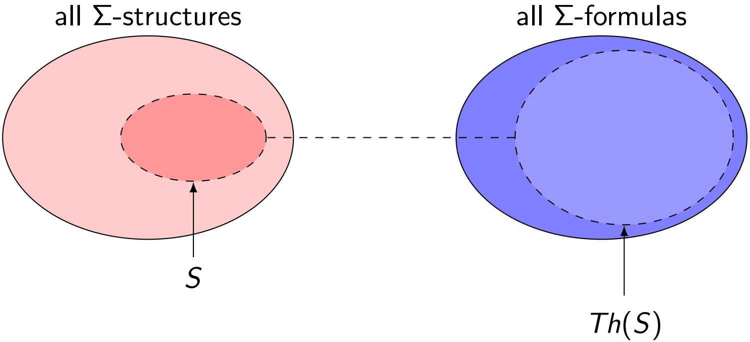
Focusing on a subclass of structures

- If S is the class of all Σ-structure then $Th(S)$ is the set of valid formulas;



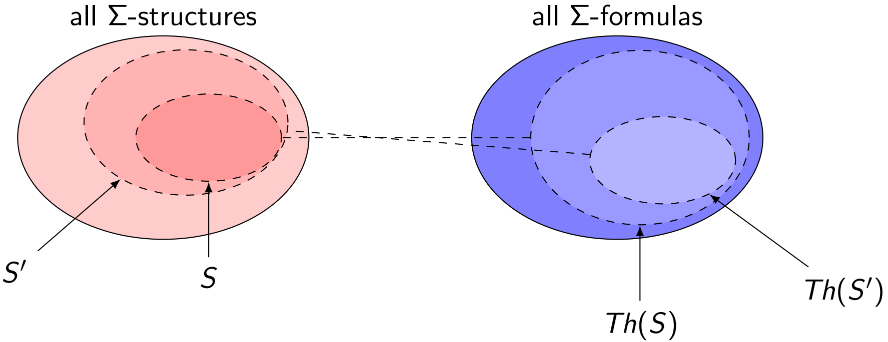
Focusing on a subclass of structures

- If S is the class of all Σ-structure then $Th(S)$ is the set of valid formulas;
- If $S \subseteq S'$ then $Th(S) \supseteq Th(S')$;



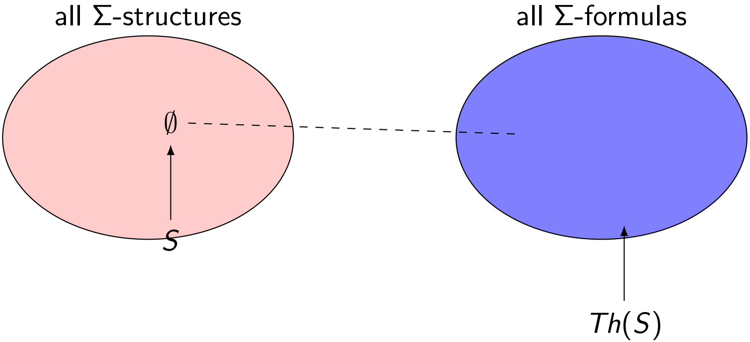
Focusing on a subclass of structures

- If S is the class of all Σ-structure then $Th(S)$ is the set of valid formulas;
- If $S \subseteq S'$ then $Th(S) \supseteq Th(S')$;



Focusing on a subclass of structures

- If S is the class of all Σ-structure then $Th(S)$ is the set of valid formulas;
- If $S \subseteq S'$ then $Th(S) \supseteq Th(S')$;
- If $S = \emptyset$ then $Th(S)$ is the set of all formulas.



First order theory

Definition (First order theory for a class of Σ structures)

A first order theory for a class of Σ-structures S, denoted with $Th(S)$ is the set of sentences in Σ that are satisfied by all the structures in S.

$$Th(S) = \{\phi \mid \text{for all } \mathcal{I} \in S.\ \mathcal{I} \models \phi\}$$

Definition (First order theory)

A first order theory is a set of formulas of the FOL language closed under the logical consequence relation. That is, T is a theory iff $T \models A$ implies that $A \in T$

Remark

A FOL theory always contains an infinite set of formulas. Indeed any theory T contains at least all the valid formulas (which are infinite).

Definition (Set of axioms for a theory)

A set of formulas Ω is a set of axioms for a theory T if for all $\phi \in T$, $\Omega \models \phi$.

First order theory (cont'd)

Definition (Finitely axiomatizable theory)

A theory T is finitely axiomatizable if it has a finite set of axioms.

Definition (Recursively axiomatizable theory)

A theory T is recursively axiomatizable if it is axiomatizable by a recursive subset Ω of formulas[1].

Definition (Axiomatizable class of Σ-structures)

A class of Σ structures S is (finitely/recursively) aximatizable if $Th(S)$ is (finitely/recursively) axiomatizable.

[1] S is a recursive subset of a set A is there is a decision procedure that decides in a finite time if $a \in S$ for every $a \in A$.

Defining theories

A theory may be expressed in two ways.

1. By giving a set Γ of formulas (axioms);
2. By giving a class S of Σ-structure.

There are theories that can not be expressed by one of the above two ways.

Example

- Number theory can only be defined using the model. There is no complete axiomatization. (Due to Gödel's incompleteness theorem)
- Set theory has no "natural model". We understand set theory via its axioms

Examples of first order theories

Number theory (or Peano Arithmetic) PA Ł contains the constant symbol 0, the 1-nary function symbol s, (for successor) and two 2-nary function symbol $+$ and $*$

1. $0 \neq s(x)$
2. $s(x) = s(y) \rightarrow x = y$
3. $x + 0 = x$
4. $x + s(y) = s(x + y)$
5. $x * 0 = 0$
6. $x * s(y) = (x * y) + x$
7. the Induction axiom schema: $\phi(0) \wedge \forall x.(\phi(x) \rightarrow \phi(s(x))) \rightarrow \forall x.\phi(x)$, for every formula $\phi(x)$ with at least one free variable

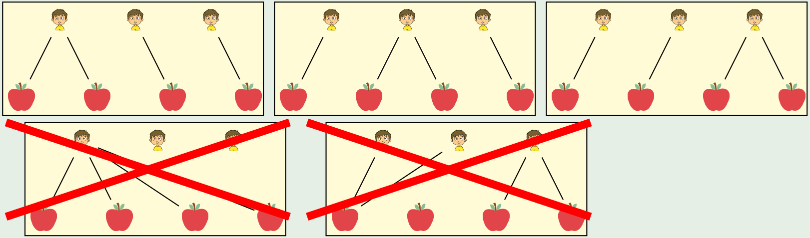
K. Gödel 1931 It's false that $\mathcal{I} \models PA$ if and only if $\mathcal{I}$ is isomorphic to the standard models for natural numbers.

Knowledge representation in FOL

- Knowledge representation in FOL is the task of listing the axioms of a theory that contains all and only the formulas that are true the structures which represents the configurations of the world that we believe are possible.

Example

Suppose that the world is composed of n apples and m childs, with $n \geq m$. We know that all the possible state of the world are thos that apples are equally distributed among child.



KR tricks

Closed world assumption

The world contains only the objects a, b, c, d, e

$$\forall x. x = a \vee x = b \vee x = c \vee x = d \vee x = e$$

Unique name assumption

Constants are interpreted in different objects:

$$a \neq b \wedge a \neq c \wedge \dots d \neq e$$

KR tricks

Grounding of quantifiers

An important consequence of CWA and UNA, is that FOL theories can be reduced (grounded) to propositional theories, i.e, quantifiers can be rewrittend in therms of finite conjunciton and disjunction

$$\begin{aligned} \forall x \phi(x) &\implies \phi(a) \wedge \phi(b) \wedge \phi(c) \wedge \phi(d) \wedge \phi(e) \\ \exists x \phi(x) &\implies \phi(a) \vee \phi(b) \vee \phi(c) \vee \phi(d) \vee \phi(e) \end{aligned}$$

Grounding generate very large formulas. For instance:

$$\forall x \exists y P(x,y) \implies \begin{array}{l} P(a,a) \vee P(a,b) \vee P(a,c) \vee P(a,d) \vee P(a,e) \wedge \\ P(b,a) \vee P(a,b) \vee P(b,c) \vee P(b,d) \vee P(b,e) \wedge \\ P(c,a) \vee P(a,b) \vee P(c,c) \vee P(c,d) \vee P(c,e) \wedge \\ P(d,a) \vee P(a,b) \vee P(d,c) \vee P(d,d) \vee P(d,e) \wedge \\ P(e,a) \vee P(a,b) \vee P(e,c) \vee P(e,d) \vee P(e,e) \end{array}$$

The advantage of grounding is that we can reuse methods designed for propositional formulas for first order formulas. For instance we can decide if a formula is satisfiable in a domain with k element by grounding with k constants and then apply sat algorithms.

KR tricks

Predicate enumeration

The binary predicate P contains only the pairs (a, b) and (c, d)

$$\forall xy.P(x, y) \rightarrow (x = a \wedge y = b) \vee (x = c \wedge y = d)$$

Domain with a given cardinality

The world conains exactly n elements

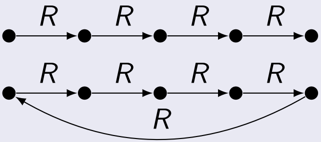
$$\exists x_1, \dots x_n \left(\bigwedge_{i<j=1}^{n} x_i \neq x_j \wedge \forall x \bigvee_{i=1}^{n} x = x_i \right)$$

Exercizes

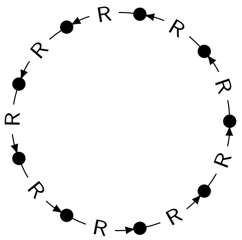
Problem

Provide a set of axioms in the signature $(0, R)$ *(0 is a constant and R a binary relation) for the class of structures that contains n elements which are isomorphic to the following structures (we provide examples with* $n = 5$*)*

1.
2.



Axiomatizing n-cycles



Since the R is a function, we can use the FOL signature Σ that contains only the funcitonal symbol R with arity equal to 1.

$$\exists x_1 \dots x_n \bigwedge_{i \leq 1 < j \leq n} x_l \neq x_j \qquad (1)$$

$$\forall xy \bigvee_{i=0}^{n-1} x = \overbrace{R(R \dots R(y) \dots)}^{i} \qquad (2)$$

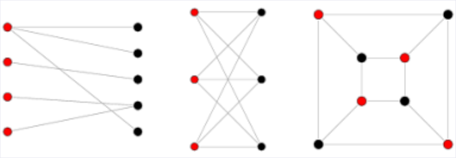
Clearly every Σ-structure $(\Delta, \mathcal{I})$ that is an n-cycle satisfies (1) and (2)
Viceversa, let us show that a Σ structure $(\Delta, \mathcal{I})$ that satisties (1) and (2) is an n-cycle.

1. $\mathcal{I} \models$ (1) implies that Δ contains at most n elements.
2. $\mathcal{I} \models$ (2) implies that every element can be reached from another element in at most $n - 1$ steps. This implies that you cannot have cycles smaller than n, Otherwise the elements outside the cycle will not be reachable from thos inside the ricle.
3. Furthermore there cannot be more than one cycle, since two cycles cannot share any elements, i.e., they are dsjoint, but this would falsify axiom (2) since two elements of two disjoint cycles are not reachable via R.
4. Therefore there is a unique cycle that contains all the elements of Δ, i,e, $(\Delta, \mathcal{I})$ is an n-cycle.

Exercizes

Problem

Bipartite graph (or bigraph) is a graph whose vertices can be divided into two disjoint and independent sets U and V such that every edge connects a vertex in U to one in V or viceversa



Consider a language that contains one unary predicates A a binary predicate R. Write a set of axioms for the theory of bipartite graphs where A is interpreted in one of the partition (either U or V).

Solution Consider the axiom:

$$\forall xy(R(x, y) \rightarrow (A(x) \equiv \neg A(y))$$

If a structure satisfies it, then R is interpreted in a binary relation that connects elements in $\mathcal{I}(A)$ with elements in $\Delta \setminus \mathcal{I}(A)$, or viceversa. Therefore $\mathcal{I}(A)$ and $\mathcal{I}(\neg A)$ correspond to the partition U and V. $\square$