KRR: Propositional logic and First Order Logic

Artificial Intelligence Challenge / Introduction to Artificial Intelligence

ICM 2A + M1 CPS²

8th March 2023

antoine.zimmermann@emse.fr

What is *logic*?

Logic…

- …is the study of the principles of correct reasoning
- …does not study truth *per se*
- …but may formalises the rules of valid inferences
- …may define what statements can *possibly* mean, independently of what they are *intended* to mean

What is ***a*** *logic*?

A logic…

- …is a **formal language**:
  - It has symbols (“words”, “glyphs”) and rules (a “grammar”) to assemble them into well-formed utterances (“formulas” or “statements”)
- …defines one of 2 things:

  Either:

  1. a system that explains how to arrange the symbols of formulas to derive other formulas that are assumed to “logically follow” from the initial formulas (a **proof theory**)

  Or:

  2. formal structures that model what symbols can mean, and the conditions when the structure representing the meaning makes a formula “true” (a **model theory**)

A logic in my course

- In my course, I will define a logic with two components:

1. The syntax, having:
   i. Symbols, grouped in different categories
   ii. Formulas, defined by a formal grammar
2. The semantics, having:
   i. A notion of interpretation (mathematical structure that maps symbols of some categories into some sets)
   ii. A relation of satisfaction that relates interpretations to formulas

Example of a simple logic: **propositional logic** (1)

1. Syntax:
   i. Symbols:
      a) Logical operators: {(,), $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\neg$}
      b) An infinite set $\mathcal{P}$ (the *propositions*), written with letters or between quotes “like this”
   ii. Formulas are sequences of symbols such that:
      a) every propositions are formulas
      b) given a formula f, $\neg f$ and (f) are formulas
      c) given formulas f_1 and f_2, $f_1 \wedge f_2$, $f_1 \vee f_2$, $f_1 \rightarrow f_2$, and $f_1 \leftrightarrow f_2$ are formulas

Examples:

(“It rains” $\wedge$ “I have my coat”) $\rightarrow$ “I am rich”

$(\neg A \vee B) \wedge (A \vee C) \rightarrow \neg\neg D$

Example of a simple logic: **propositional logic** (2)

2. Semantics:
   i. Interpretation:
      An interpretation in propositional logic is a mapping $\mathfrak{I}: \mathcal{P} \rightarrow \{\bot, \top\}$
      For a proposition $A \in \mathcal{P}$, $\mathfrak{I}(A) = \bot$ means that the proposition is interpreted as **false**, and $\mathfrak{I}(A) = \top$ means that the proposition is interpreted as **true**
   ii. Satisfaction relation:
      We define the satisfaction relation $\models$ (read "satisfies") between interpretations and formulas as follows:
      If A is a proposition, $\mathfrak{I} \models A$ if and only if $\mathfrak{I}(A) = \top$
      If f is a formula: a) $\mathfrak{I} \models (f)$ if and only if $\mathfrak{I} \models f$ b) $\mathfrak{I} \models \neg f$ if and only if $\mathfrak{I} \not\models f$
      If f_1 and f_2 are formulas:
      a) $\mathfrak{I} \models f_1 \wedge f_2$ if and only if $\mathfrak{I} \models f_1$ and $\mathfrak{I} \models f_2$
      b) $\mathfrak{I} \models f_1 \vee f_2$ if and only if $\mathfrak{I} \models f_1$ or $\mathfrak{I} \models f_2$ or both
      c) $\mathfrak{I} \models f_1 \rightarrow f_2$ if and only if $\mathfrak{I} \models \neg f_1$ or $\mathfrak{I} \models f_2$ or both
      d) $\mathfrak{I} \models f_1 \leftrightarrow f_2$ if and only if $\mathfrak{I} \models f_1 \rightarrow f_2$ and $\mathfrak{I} \models f_2 \rightarrow f_1$

Formalising all logics

A logic L is a 4-tuple (Sign_L,Form_L,Int_L,$\models_L$) where:

- Sign_L is a set of *signatures* (the symbols of the language)
- Form_L is a set of *formulas* (that can be written from symbols)
- Int_L is a set of *interpretations* (that interpret the signatures)
- $\models_L$ is the *satisfaction relation*, such that $\models_L \subseteq \text{Int}_L \times \text{Form}_L$

Entailment

- In any logic L defined as before, we say that:

a set of formula K entails a formula f (in L)

if and only if

all interpretations I in Int_L that satisfy all formulas in K also satisfy f

i.e.:

$$K \; L\text{-entails } f \text{ iff } \forall \mathfrak{I} \in \text{Int}_L, (\forall \alpha \in K, \mathfrak{I} \models_L \alpha) \Rightarrow \mathfrak{I} \models_L f$$

- In this case, we write $K \models_L f$

Exercise in propositional logic

- Define the rules of sudoku in propositional logic:
  - A sudoku is a 9x9 grid of numbers
  - Each row must contain all the numbers from 1 to 9
  - Each column must contain all the numbers from 1 to 9
  - Each of the nine 3x3 blocks that together makes the whole sudoku must contain all the numbers from 1 to 9

First order logic syntax (1)

1. Syntax:
   i. Symbols:
      a) Logical operators: {(,), $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\neg$} plus the comma character " , "
      b) Quantifiers: {$\exists$, $\forall$}
      c) An infinite $\mathcal{V}$ (the *variables*)
      d) An infinite $\mathcal{C}$ (the *constants*)
      e) An infinite set $\mathcal{P}$ (the *propositions*)
      f) For each natural number $n > 0$, an infinite set $\mathcal{F}_n$ (the *n-ary functions*)
      g) For each natural number $n > 0$, an infinite set $\mathcal{P}_n$ (the *n-ary predicates*)

   ii. Terms:

      To simplify the definition of formulas, we define the notion of *terms*

      a) A variable or a constant is a term
      b) For any $n > 0$, if f is an n-ary function and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term

First order logic syntax (2)

i. Formulas are sequences of symbols such that:
   a) every propositions are formulas
   b) given a formula f , $\neg f$ and (f) are formulas
   c) given formulas f_1 and f_2, $f_1 \wedge f_2$, $f_1 \vee f_2$, $f_1 \rightarrow f_2$, and $f_1 \leftrightarrow f_2$ are formulas
   d) For any $n > 0$, if P is an n-ary predicate and $t_1, \ldots, t_n$ are terms, then $P(t_1, \ldots, t_n)$ is a formula
   e) given a formula f and a variable x, $\exists x\, f$ and $\forall x\, f$ are formulas [*in this case, f is said to be the* scope *of x for the quantifier*]

If a variable in a formula appears outside its scope for a quantifier, it is a *free* variable for the formula. A formula without free variables is a *closed* formula.

Examples:

$$\forall x\ \exists y\ P(f(x)) \rightarrow R(x, y)$$

$$\exists z\ \forall x\ \exists x\ \neg(A \rightarrow R(t, y))$$

First order logic semantics (1)

i. Interpretation:

An interpretation $\mathfrak{I}$ in FOL is defined over a non empty set *U* (the *universe* or *domain* of interpretation or of discourse) such that:

a) For all constant $c \in \boldsymbol{C}$, $\mathfrak{I}(c) \in U$

b) For all *n*-ary function *f*, $\mathfrak{I}(f) : U^n \rightarrow U$ *

c) For all proposition $A \in \boldsymbol{\mathcal{P}}$, $\mathfrak{I}(A) \in \{\bot, \top\}$

d) For all *n*-ary function *f*, $\mathfrak{I}(f) : U^n \rightarrow \{\bot, \top\}$ *

- *NB: constants could be considered as* 0-*ary functions and propositions as* 0-*ary predicates*

ii. Variable assignment:

An interpretation does not interpret variables, but we need something to take them into account for the satisfaction relation.

A *variable assignment* for an interpretation $\mathfrak{I}$ is a function a : $\boldsymbol{\mathcal{V}} \rightarrow U$

First order logic semantics (2)

iii. Interpretation modulo a variable assignment:

If a formula is not closed, satisfaction must be defined according to a variable assignment *a*
We then extend $\mathfrak{I}$ to $\mathfrak{I}_a$ in order to interpret variables and terms:

For a variable x, $\mathfrak{I}_a(x) = a(x)$

For an n-ary function f and terms $t_1, \ldots, t_n$, $\mathfrak{I}_a(f(t_1, \ldots, t_n)) = \mathfrak{I}(f)(\mathfrak{I}_a(t_1), \ldots, \mathfrak{I}_a(t_n))$

First order logic semantics (3)

iv. Satisfaction modulo an assignment:

The satisfaction relation $\models$ between interpretations $\mathfrak{I}$ modulo assignment *a* is as follows:

If *P* is an *n-ary* predicate and $t_1, \ldots, t_n$, are terms:

$$\mathfrak{I}_a \models P(t_1, \ldots, t_n) \text{ iff } \mathfrak{I}(P)(\mathfrak{I}_a(t_1), \ldots, \mathfrak{I}_a(t_n)) = \top$$

If *f* is a formula: a) $\mathfrak{I}_a \models (f)$ if and only if $\mathfrak{I}_a \models f$ b) $\mathfrak{I}_a \models \neg f$ if and only if $\mathfrak{I}_a \not\models f$

c) $\mathfrak{I}_a \models \exists x\, f$ if and only if there exists *e* in *U* such that, if we change *a* into *a'* by simply having *a'*(*x*) = *e* and *a'* identical to *a* otherwise, then $\mathfrak{I}_{a'} \models f$

d) $\mathfrak{I}_a \models \forall x\, f$ if and only if for all *e* in *U*, if we change *a* into *a'* by simply having *a'*(*x*) = *e* and *a'* identical to *a* otherwise, then $\mathfrak{I}_{a'} \models f$

If f_1 and f_2 are formulas:

a) $\mathfrak{I}_a \models f_1 \wedge f_2$ if and only if $\mathfrak{I}_a \models f_1$ and $\mathfrak{I}_a \models f_2$

b) $\mathfrak{I}_a \models f_1 \vee f_2$ if and only if $\mathfrak{I}_a \models f_1$ or $\mathfrak{I}_a \models f_2$ or both

c) $\mathfrak{I}_a \models f_1 \rightarrow f_2$ if and only if $\mathfrak{I}_a \models \neg f_1$ or $\mathfrak{I}_a \models f_2$ or both

d) $\mathfrak{I}_a \models f_1 \leftrightarrow f_2$ if and only if $\mathfrak{I}_a \models f_1 \rightarrow f_2$ and $\mathfrak{I}_a \models f_2 \rightarrow f_1$

v. Satisfaction:

$\mathfrak{I} \models f$ if and only if for all assignment *a* $\mathfrak{I}_a \models f$

Modelling in FOL

- Gustav buys a second hand Tesla Model S from Roberto in July 6th, 2012 for $35,000
- Gustav sells the Tesla Model S to someone in October 10th, 2020 for an undisclosed price
- The same Tesla Model S is bought by Gustav for $12,000 at a forgotten date
- What are the predicates that we need?
- What are the constants?
- Is there general knowledge about transactions that we can assert independently of Gustav, the Tesla car owned by him, Roberto, and any date?