Crash course in First Order Logic

Raphaël Carroy

Università di Torino

Marseille, France
Octobre 2013

What is this all about?

We want to formalize a framework in which all mathematics could be done.

What does maths do? For instance, study provable properties of a given structure, as the integers or the reals. Ideally a structure is **exactly caracterized** by a simple list of **definable properties** of its elements. Here,

- **exactly caracterized** means up to isomorphism, and
- **definable properties** means formulae in a specified language

Problem.

Is it possible to find a list of axioms concerning the reals that are valid up to isomorphism only in the real line?

What is this all about?

To address this question we need to define properly what we mean by **isomorphism, formulae** and **valid**. This is what modern logic does, and the starting point is the separation between *syntax* and *semantics*.

Syntax handles:

- What a language is,
- What a sentence is,
- What a proof is.

Semantics handles:

- What a structure is,
- What a morphism is,
- When does a sentence hold in a structure.

The very nice feature of FOL is the exact adequacy between these two sides.

Syntax: the language

A first-order language consists in two parts, a logical part which is the same in every language, and a symbolic part which can vary.

The logical part

- Boolean connectives: $\wedge, \vee, \neg, \rightarrow, \leftrightarrow$
- Quantifiers: $\forall, \exists$
- An infinite (countable) set of variables $x, y, z, x_0, x_1, \cdots$
- Parentheses: ()
- The symbol for the equality relation $=$.

The symbolic part

- Constant symbols, ex: $0, \pi, e, 1, \cdots$
- Function symbols with different arities, ex: $+, \cos, \cdots$
- Relation symbols with different arities, ex: $<, \in, \cong, \cdots$.

We often give only the symbolic part of a first-order language.

Syntax: example of the language of groups

The symbolic part of language of groups consists in:

- A binary function symbol $\star$,
- A constant symbol e.

The binary function symbol stands for the group operation and the constant stands for the neutral element.

Remarks

1. We gave the symbolic part, but the logical part is also there.
2. We could have added the unary operation inverse.
3. Everything can be done only with the binary operation $\star$!
4. The only predicate symbol of this language is $=$.

Syntax: example of the axioms of groups

Here are three formulae in the language of groups.

1. $\varphi_0 : \forall x\, \forall y\, \forall z\, ((x \star y) \star z = x \star (y \star z))$,
2. $\varphi_1 : \forall x\, \exists y\, (x \star y = e)$,
3. $\varphi_2 : \forall x\, (x \star e = x)$.

The formula φ_0 expresses the associativity of the group operation, φ_1 the existence of a right inverse and φ_2 the existence of the neutral element. Together they form the **axioms of group theory**.

Warning

We do not give here the formal definition of a formula! It is made by induction and uses trees. Instead, we walk through it to show that the sequences of symbols above are indeed formulae.

Syntax: formula and terms, general idea

- Which are the strings of symbols that form a formula? Could we run a computer program to sort them out?
- We need to decompose a formula. But what are the simplest possible formulae? Where do we stop?
- The simplest formulae are the **atomic** ones, without quantifiers or connectives. What do they apply to?
- Variables, constants, images of variables by a function... These strings of symbols are **terms**. They bear no meaning, and are just names for elements. These are the first we need to define.

Syntax: formula and terms, on an example

We first have to check that the terms are correctly built.

- variables and constant are terms.
- if t, t' are terms, so is $(t \star t')$

So here we have step one:

1. $\varphi_0 : \forall x \, \forall y \, \forall z \, ((x \star y) \star z = x \star (y \star z))$,
2. $\varphi_1 : \forall x \, \exists y \, (x \star y = e)$,
3. $\varphi_2 : \forall x \, (x \star e = x)$.

Syntax: formula and terms, on an example

We first have to check that the terms are correctly built.

- variables and constant are terms.
- if t, t' are terms, so is $(t \star t')$

So here we have step two:

1. $\varphi_0 : \forall x \, \forall y \, \forall z \, ((x \star y) \star z = x \star (y \star z))$,
2. $\varphi_1 : \forall x \, \exists y \, (x \star y = e)$,
3. $\varphi_2 : \forall x \, (x \star e = x)$.

Syntax: formula and terms, on an example

We first have to check that the terms are correctly built.

- variables and constant are terms.
- if t, t' are terms, so is $(t \star t')$

And finally step three:

1. $\varphi_0 : \forall x\, \forall y\, \forall z\, ((x \star y) \star z = x \star (y \star z))$,
2. $\varphi_1 : \forall x\, \exists y\, (x \star y = e)$,
3. $\varphi_2 : \forall x\, (x \star e = x)$.

Syntax: formula and terms, on an example

We then check that the formulae are correctly built.

- Relation symbols involving terms are **atomic formulae**.
- When φ and ψ are formulae, so are $\varphi \wedge \psi$, $\varphi \vee \psi$, $\cdots$
- When φ is a formula so is $\forall x\,(\varphi)$

So here we have step one:

1. $\varphi_0 : \forall x\,\forall y\,\forall z\,((x \star y) \star z = x \star (y \star z))$,
2. $\varphi_1 : \forall x\,\exists y\,(x \star y = e)$,
3. $\varphi_2 : \forall x\,(x \star e = x)$.

Syntax: formula and terms, on an example

We then check that the formulae are correctly built.

- Relation symbols involving terms are **atomic formulae**.
- When φ and ψ are formulae, so are $\varphi \wedge \psi$, $\varphi \vee \psi$, $\cdots$
- When φ is a formula so is $\forall x \, (\varphi)$

So here we have step two:

1. $\varphi_0 : \forall x \, \forall y \, \forall z \, ((x \star y) \star z = x \star (y \star z))$,
2. $\varphi_1 : \forall x \, \exists y \, (x \star y = e)$,
3. $\varphi_2 : \forall x \, (x \star e = x)$.

Syntax: formula and terms, on an example

We then check that the formulae are correctly built.

- Relation symbols involving terms are **atomic formulae**.
- When φ and ψ are formulae, so are $\varphi \wedge \psi$, $\varphi \vee \psi$, $\cdots$
- When φ is a formula so is $\forall x\, (\varphi)$

So here we have step three:

1. $\varphi_0 : \forall x\, \forall y\, \forall z\, ((x \star y) \star z = x \star (y \star z))$,
2. $\varphi_1 : \forall x\, \exists y\, (x \star y = e)$,
3. $\varphi_2 : \forall x\, (x \star e = x)$.

Syntax: formula and terms, on an example

We then check that the formulae are correctly built.

- Relation symbols involving terms are **atomic formulae**.
- When φ and ψ are formulae, so are $\varphi \wedge \psi$, $\varphi \vee \psi$, $\cdots$
- When φ is a formula so is $\forall x\,(\varphi)$

And finaly step four:

1. $\varphi_0 : \forall x\, \forall y\, \forall z\, ((x \star y) \star z = x \star (y \star z))$,
2. $\varphi_1 : \forall x\, \exists y\, (x \star y = e)$,
3. $\varphi_2 : \forall x\, (x \star e = x)$.

Syntax: Proofs, simplified version

Proofs are defined inductively as well. There are several ways to do it, we give here a general idea of the definition.

A **proof** of a formula φ from a set of formulae (or **theory**) Γ is a *finite* sequence of formulae that are

- either in Γ
- deduced from the beginning of the sequence using a finite set of rules.

Example of rule, $R_\wedge$.

If $\varphi \wedge \psi$ figures in the beginning of the sequence then both φ and ψ can figure afterwards. Conversely if both φ and ψ figure in the beginning, $\varphi \wedge \psi$ can figure afterwards.

Definition

When a proof of φ from Γ exists, we write $\Gamma \vdash \varphi$.

Syntax: Proofs, an example in the theory of groups

Instead of giving the rules for all connectives, quantifers and for the equality, let us now give an example of their use in a proof:

$$\varphi_0, \varphi_1, \varphi_2 \vdash \forall x \, \exists y (x \star y = e \wedge y \star x = e).$$

In words, group theory proves the existence of a two-sided inverse.

Syntax: Proofs, an example in the theory of groups

(φ_1)	$\forall x\, \exists y (x \star y = e)$	(1)
$(R_\exists, R_\forall, (1))$	$x \star y = e$	(2)
$(R_=, (2))$	$y \star (x \star y) = y \star e$	(3)
$(\varphi_0, \varphi_2, (3))$	$(y \star x) \star y = y$	(4)
$(\varphi_1, R_\forall)$	$\exists y' \, (y \star y' = e)$	(5)
$(R_\exists, (5))$	$y \star y' = e$	(6)
$(R_=, (4), (6))$	$((y \star x) \star y) \star y' = e$	(7)
$(\varphi_0, (7))$	$(y \star x) \star (y \star y') = e$	(8)
$(R_=, (5), (8))$	$(y \star x) \star e = e$	(9)
$(\varphi_2, (9))$	$y \star x = e$	(10)
$(R_\wedge, (2), (10))$	$x \star y = e \wedge y \star x = e$	(11)
$(R_\forall, R_\exists, (11))$	$\forall x\, \exists y (x \star y = e \wedge y \star x = e)$	(12)

Transition

What about semantics?

Semantics: $\mathcal{L}$-structures

Given a first-order language $\mathcal{L}$, an $\mathcal{L}$-**structure** $\mathfrak{M}$ is given by:

- A set M which we call the **domain** of $\mathfrak{M}$,
- for each constant symbol c in $\mathcal{L}$ an element $c^{\mathfrak{M}}$ of M,
- for each n-ary function symbol f in $\mathcal{L}$ an n-ary function $f^{\mathfrak{M}} : M^n \to M$,
- for each n-ary relation symbol R in $\mathcal{L}$ a subset $R^{\mathfrak{M}}$ of M^n.

If $\mathcal{L} = \{e, \star\}$ is the language of groups, here are a few examples of $\mathcal{L}$-structures:

- $\mathfrak{N} = (\mathbb{N}, 0, +)$
- $\mathfrak{R} = (\mathbb{R}, 0, +)$
- $\mathfrak{Z} = (\mathbb{Z}, 0, +)$
- $(\mathbb{Z}, 1, +)$
- $(\mathbb{R}, 1, \times)$
- $(\mathcal{C}(\mathbb{R}), id, \circ)$

Semantics: Morphisms between $\mathcal{L}$-structures

Given two $\mathcal{L}$-structures $\mathfrak{M}$ and $\mathfrak{N}$, a function $\sigma : M \to N$ is a $\mathcal{L}$-morphism from $\mathfrak{M}$ to $\mathfrak{N}$ if:

- for each constant symbol c in $\mathcal{L}$, $\sigma(c^{\mathfrak{M}}) = c^{\mathfrak{N}}$,
- for each n-ary function symbol f in $\mathcal{L}$, $\sigma(f^{\mathfrak{M}}(\bar{x})) = f^{\mathfrak{N}}(\bar{x})$,
- for each n-ary relation symbol R in $\mathcal{L}$, $R^{\mathfrak{M}}(\bar{x})$ implies $R^{\mathfrak{N}}(\sigma(\bar{x}))$.

If $\mathcal{L} = \{e, \star\}$ is the language of groups, here are a few examples of $\mathcal{L}$-morphisms:

$$\begin{aligned}(\mathbb{Z}, 0, +) &\longrightarrow (\mathbb{R}, 1, \times) \\ n &\longmapsto 2^n\end{aligned} \qquad \begin{aligned}(\mathbb{N}, 0, +) &\longrightarrow (\mathcal{C}(\mathbb{R}), id, \circ) \\ n &\longmapsto (x \mapsto x + n)\end{aligned}$$

An **isomorphism** is a bijective morphism which inverse is also a morphism.

Semantics: Validity in $\mathcal{L}$-structures

Validity of a formula φ in a $\mathcal{L}$-structure $\mathfrak{M}$ is defined (once again) by induction. We give now a general idea of how it is done.

- validity for atomic formulae in $\mathfrak{M}$ is given with $\mathfrak{M}$ itself.
- we say $\varphi \wedge \psi$ is valid in $\mathfrak{M}$ when both φ and ψ are valid in $\mathfrak{M}$,
- we say that $\forall x\,(\varphi)$ is valid in $\mathfrak{M}$ if for every possible value $a \in M$ that can be assigned to the variable x, the formula $\varphi(a)$ is valid in $\mathfrak{M}$.
- $\cdots$

Definition

When φ is valid in $\mathfrak{M}$ we say that $\mathfrak{M}$ is **a model** of φ and we write

$$\mathfrak{M} \models \varphi.$$

Semantics: Examples on groups.

Fix $\mathcal{L}$ the language of groups, recall the axioms of group theory:

1. $\varphi_0 : \forall x \, \forall y \, \forall z \, ((x \star y) \star z = x \star (y \star z))$,
2. $\varphi_1 : \forall x \, \exists y \, (x \star y = e)$,
3. $\varphi_2 : \forall x \, (x \star e = x)$.

and check whether or not they are valid in each of our examples:

- $\mathfrak{N} = (\mathbb{N}, 0, +)$
- $\mathfrak{R} = (\mathbb{R}, 0, +)$
- $\mathfrak{Z} = (\mathbb{Z}, 0, +)$
- $(\mathbb{Z}, 1, +)$
- $(\mathbb{R}, 1, \times)$
- $(\mathcal{C}(\mathbb{R}), id, \circ)$

Semantics: Examples on groups.

Star with φ_0, associativity of the operation.

$$\varphi_0 : \forall x \, \forall y \, \forall z \, ((x \star y) \star z = x \star (y \star z)).$$

Addition is associative in $\mathbb{N}, \mathbb{Z}$ and $\mathbb{R}$, multiplication is associative in $\mathbb{R}$ and composition of functions is associative, so:

- $\mathfrak{N} = (\mathbb{N}, 0, +) \models \varphi_0$
- $\mathfrak{R} = (\mathbb{R}, 0, +) \models \varphi_0$
- $\mathfrak{Z} = (\mathbb{Z}, 0, +) \models \varphi_0$
- $(\mathbb{Z}, 1, +) \models \varphi_0$
- $(\mathbb{R}, 1, \times) \models \varphi_0$
- $(\mathcal{C}(\mathbb{R}), id, \circ) \models \varphi_0$

Semantics: Examples on groups.

Continue with φ_1, the existence of an inverse

$$\varphi_1 : \forall x \, \exists y \, (x \star y = e).$$

We have:

- $\mathfrak{N} \models \neg(1 + n = 0)$ for any integer $n \in \mathbb{N}$
- $\mathfrak{R} \models x + (-x) = 0$ for any real $x \in \mathbb{R}$
- $\mathfrak{Z} \models n + (-n) = 0$ for any $n \in \mathbb{Z}$
- $(\mathbb{Z}, 1, +) \models n + (-n + 1) = 1$ for any $n \in \mathbb{Z}$
- $(\mathbb{R}, 1, \times) \models \forall x \, (0 \times x = 0)$
- $(\mathcal{C}(\mathbb{R}), id, \circ) \models \forall x \, (\neg(\mathbf{1} \circ x = id))$

Semantics: Examples on groups.

Continue with φ_1, the existence of an inverse

$$\varphi_1 : \forall x \, \exists y \, (x \star y = e).$$

So:

- $\mathfrak{N} = (\mathbb{N}, 0, +) \models \varphi_0 \land \neg\varphi_1$
- $\mathfrak{R} = (\mathbb{R}, 0, +) \models \varphi_0 \land \varphi_1$
- $\mathfrak{Z} = (\mathbb{Z}, 0, +) \models \varphi_0 \land \varphi_1$
- $(\mathbb{Z}, 1, +) \models \varphi_0 \land \varphi_1$
- $(\mathbb{R}, 1, \times) \models \varphi_0 \land \neg\varphi_1$
- $(\mathcal{C}(\mathbb{R}), id, \circ) \models \varphi_0 \land \neg\varphi_1$

Semantics: Examples on groups.

Finish with the neutrality of the distinguished element for the operation.

$$\varphi_2 : \forall x\,(x \star e = x).$$

By definition 0 is neutral for $+$ in $\mathbb{N}, \mathbb{Z}$ and $\mathbb{R}$, 1 is neutral for $\times$ and the identity is neutral for $\circ$, but

$$(\mathbb{Z}, 1, +) \models \neg(n + 1 = n) for any n \in \mathbb{Z}.$$

So:

- $\mathfrak{N} \models \varphi_0 \wedge \neg\varphi_1 \wedge \varphi_2$
- $\mathfrak{R} \models \varphi_0 \wedge \varphi_1 \wedge \varphi_2$
- $\mathfrak{Z} \models \varphi_0 \wedge \varphi_1 \wedge \varphi_2$
- $(\mathbb{Z}, 1, +) \models \varphi_0 \wedge \varphi_1 \wedge \neg\varphi_2$
- $(\mathbb{R}, 1, \times) \models \varphi_0 \wedge \neg\varphi_1 \wedge \varphi_2$
- $(\mathcal{C}(\mathbb{R}), id, \circ) \models \varphi_0 \wedge \neg\varphi_1 \wedge \varphi_2$

Semantics: Consequence

We now have a notion of **semantic consequence**:

Definition

Given a first-order language $\mathcal{L}$, a theory Γ and a formula φ in $\mathcal{L}$, we say that φ is the **semantic consequence** of Γ if every model of Γ is a model of φ, and we write $\Gamma \models \varphi$.

Every rule for syntactic proof correspond to a point in the very definition of validity so we have indeed

Fact

$\Gamma \vdash \varphi$ *implies* $\Gamma \models \varphi$.

Hence every element of any group has a two-sided inverse.

Semantics: Consequence, example in groups

In any group, the neutral element is two-sided neutral:

Proposition

$\varphi_0, \varphi_1, \varphi_2 \models \forall x \, (e \star x = x)$

Proof.

Take $\mathfrak{G} = (G, e_G, \star_G)$ any group, and g any element of G. Call h its two-sided inverse. Then we have

$$e_G \star_G g = (g \star_G h) \star_G g = g \star_G (h \star_G g) = g \star_G e_G = g.$$

□

Gödel's completeness Theorem.

Given a language $\mathcal{L}$ and a theory Γ, we say that Γ is **consistent** if it has a model. It is **incoherent** if $\Gamma \vdash \varphi \wedge \neg\varphi$ for some $\mathcal{L}$-formula φ.

Theorem

Γ is consistent if and only if it is coherent.
Equivalently, for any formula φ we have

$$\Gamma \vdash \varphi \text{ iff } \Gamma \models \varphi.$$

As consequence, we have for groups

$$\varphi_0, \varphi_1, \varphi_2 \vdash \forall x\,(e \star x = x).$$

Applications: Complete theories.

Definition

Given a language $\mathcal{L}$ and a theory Γ in it, we say that Γ is **complete** if for any formula φ either $\Gamma \vdash \varphi$ or $\Gamma \vdash \neg\varphi$.

A complete coherent theory Γ is maximal among coherent theories, since for any $\varphi \notin \Gamma$ we have $\neg\varphi \in \Gamma$ so $\Gamma \cup \{\varphi\} \vdash \varphi \wedge \neg\varphi$ so it is incoherent.

Fact

Group theory is incomplete.

Proof.

There are commutative groups and non-commutative groups, so by the Completeness Theorem we have both

$$\varphi_0, \varphi_1, \varphi_2 \nvdash \forall x \, \forall y \, (y \star x = x \star y) \text{ and}$$
$$\varphi_0, \varphi_1, \varphi_2 \nvdash \neg(\forall x \, \forall y \, (y \star x = x \star y)). \quad \square$$

Applications: Elementary equivalence.

When two structures have the exact same properties definable in $\mathcal{L}$,

Definition

We say that two $\mathcal{L}$-structures $\mathfrak{M}$ and $\mathfrak{N}$ are **elementary equivalent** and we write $\mathfrak{M} \equiv \mathfrak{N}$ if for any $\mathcal{L}$-formula φ we have

$$\mathfrak{M} \models \varphi \text{ iff } \mathfrak{N} \models \varphi.$$

Using the Completeness Theorem, we get elementary equivalence as the semantic counterpart of complete theories:

Proposition

A coherent theory Γ is complete iff all models of Γ are elementary equivalent.

Applications: Rephrasing the original problem.

We can now rephrase our original problem concerning the reals. Let $\mathcal{L}$ be the language of ordered fields, so $\mathcal{L} = \{e_0, e_1, \star_+, \star_\times, S\}$ where e_0, e_1 are constant symbols, $\star_+$ and $\star_\times$ are binary function symbols and S is a binary relation symbol.
Set $\mathfrak{R} = (\mathbb{R}, 0, 1, +, \times, \leq)$, then

Problem.

Given a $\mathcal{L}$-structure $\mathfrak{M}$ do we have

$$\mathfrak{R} \equiv \mathfrak{M} \text{ implies } \mathfrak{R} \cong \mathfrak{M}?$$

Short answer, using Completeness: NO!