CDM

First-Order Logic

Klaus Sutner

Carnegie Mellon University

40-fol-basics 2017/12/15 23:21

1. First-Order Logic

- Syntax
- Model Theory

There are three major parts in the design of a new logic:

- language (syntax)
- model theory (semantics)
- proof theory (deductions)

Describing a language is fairly straightforward, in particular if one is not interested in actually building a parser or implementing various algorithms.

Model theory brushes up agains set theory (at least for infinite models) and can be quite challenging.

Proof theory is arguably the most difficult part: while deductions are just finite data structures, understanding them in detail is hard.

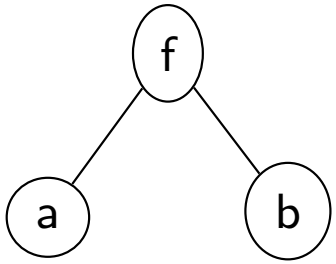
According to the Handbook of Mathematical Logic, the field is organized into 4 areas:

- model theory
- set theory
- recursion theory
- proof theory

Note that there is no special category for the syntax of logic languages, it is considered the easy part of model theory and proof theory.

Fair enough in math, but in the computational universe there are quite a few interesting algorithmic challenges.

Writing $f(a, b)$ is perhaps most convenient for a human user, but there are other options: fab, abf or even a labeled tree like



The other representations are arguably easier to handle algorithmically (in particular the tree is standard tool in parsers).

We will usually stick with the standard human-readable form and avoid higher levels of precision.

Neither propositional nor equational logic is powerful enough to express statements such as

- The function $x \mapsto x^3$ is continuous.
- There is a prime between n and $2n$ (for $n > 1$).
- Deadlock cannot occur (in some protocol).
- Upon completion, the stack is empty.
- The algorithm terminates on all inputs.

For example, the second assertion (known as the Bertrand-Chebyshev theorem) is of importance for algorithms that need to select primes of some suitable size: there is a prime between 2^k and 2^{k+1}, so we can select a prime with any given number of bits (this is just the tip of an iceberg).

How could we state this assertion clearly, as a formal expression that might be computer-understandable?

So how do we construct a language suitable for (most of) mathematics and computer science?

Let's start with a small fragment, say, arithmetic. A reasonable approach would be to use only

- basic arithmetic concepts (addition, multiplication, order) and
- purely logical constructs such as "and", "not" etc.

The logical constructs will include all of propositional logic so we can still combine assertions by connectives "and", "or" and the like. But there will also be quantifiers that allow one to make statements about the existence of objects with certain properties and about properties of all objects.

Terminology: This type of system is called first-order logic (FOL) or predicate logic. In the 20th century, FOL emerged as the de facto workhorse in all of math and TCS.

We could express the statement about primes like so:

- for all x - universal quantifier
- there exists a y - existential quantifier
- $x \leq y$ - binary relation
- and - logical connective
- $y \leq 2x$ - binary relation
- and - logical connective
- y is prime - unary relation

The assertion "is prime" in the last line still needs to be unraveled.

Clearly, we can express primality in terms of multiplication.

x is prime iff

- for all u - universal quantifier
- for all v - universal quantifier
- $x = u \cdot v$ - binary function, equation
- implies - logical connective
- $u = 1$ - constant, equation
- or - logical connective
- $v = 1$ - constant, equation

As it turns out, this is all we need: adding quantifiers increases the expressiveness of our language enormously.

Note that the last definition tacitly assumes that we interpret the variables as ranging over the integers.

If we were dealing with the rationals the definition would, of course, no longer make sense. We will pin down details in the model theory section.

We will see in a moment that the range of quantifiers is pinned down precisely when we address the question of semantics: what exactly is the meaning of a formula in FOL?

The language of first-order logic consists of the following pieces:

constants	that denote individual objects,
variables	that range over individual objects,
relation symbols	that denote relations,
function symbols	that denote functions,
logical connectives	"and," "or," "not," "implies,"
existential quantifiers	that express "there exists,"
universal quantifiers	that express "for all."

- a, b, c, ... for constants,
- x, y, z, ... for variables,
- $\vee$, $\wedge$, $\neg$ for the logical connectives,
- $\exists$ for the existential quantifier,
- $\forall$ for the universal quantifier,
- f, g, h, ... for function symbols,
- R, P, Q, ... for relation symbols.
- Always allow $=$ for equality.

These are just conventions, there are no sacred cows here. E.g., we might also use subscripted symbols and additional connectives such as $\rightarrow$ or $\leftrightarrow$.

One should distinguish more carefully between function and relation symbols of different arity: nullary, unary, binary functions and so on.

In most concrete structures only a finite number of function and relation symbols are needed.

Hence, we can convey the structure of the non-logical part of the language in a signature or similarity type: a list that indicates the arities of all the function and relation symbols.

Example

For group theory one often uses a signature $(2, 1, 0)$: there is one binary function symbol for the group multiplication, one unary function symbol for the inverse operation and a nullary function symbol (i.e. a constant) for the neutral element.

We could also use a language of signature (2) but at the cost of slightly more complicated axioms.

There are occasions when one wants to use countably many function or relation symbols. It is easy to describe languages of this form. For example, we could allow for constants c_n, $n \in \mathbb{N}$ (together with axioms $c_i \neq c_j$ for $i < j$ this would make sure that all models are infinite).

In mathematical logic one also considers languages of some higher cardinality κ, but this will be of no interest to us. Note that some amount of set theory is needed just to define such a language.

For the classical algebraic theory of fields one minimally uses a language

$$\mathcal{L}(+, \cdot, 0, 1) \text{ of signature } (2, 2, 0, 0)$$

- binary function symbols $+$ and $\cdot$ for addition and multiplication,
- constants 0 and 1 for the respective neutral elements.

Of course, more functions could be added: additive inverse, subtraction, multiplicative inverse, division. Introducing only addition and multiplication is the minimalist approach here.

Now consider formulae such as

$$\forall\, x, y\, (x + y = y + x)$$
$$\forall\, x\, (x \cdot 0 = 0)$$
$$\forall\, x\, \exists\, z\, ((\neg x = 0) \to z \cdot x = 1)$$
$$1 + 1 = 0$$

It is intuitively clear what these formulae mean.

Except for the last one, they are all true in any field. In fact, the first two hold in any ring.

The third one is true in a field, but is false in an arbitrary ring.

And the last only holds in rings of characteristic 2. In fact, it defines rings of characteristic 2.

In Boolean algebra one uses the language

$$\mathcal{L}(\sqcup, \sqcap, \overline{\ }, 0, 1) \quad \text{of signature } (2, 2, 1, 0, 0)$$

So we have

- binary function symbols $\sqcup$ and $\sqcap$ (for join and meet)
- unary function symbol (for complement)
- constants 0 and 1

Some formulae:

$$\forall\, x, y\, \exists\, z\, (x \sqcup y = \overline{z})$$
$$\exists\, x\, \forall\, y\, (x \sqcap y = 0)$$
$$\forall\, x\, \exists\, y\, (x \sqcup y = 1 \wedge x \sqcap y = 0)$$

For arithmetic it is convenient to have a relation symbol for order:

$\mathcal{L}(+, \cdot, 0, 1; <)$ of signature $(2, 2, 0, 0; 2)$

- binary function symbols $+$ and $\cdot$ (for integer addition and multiplication)
- constants 0 and 1 (for integers 0 and 1)
- a binary relation symbol $<$ (for the less-than relation)

Assuming we are quantifying over the natural numbers we have assertions like

$$\forall\, x\, \exists\, y\, (x < y)$$
$$\exists\, x\, \forall\, y\, (x = y \vee x < y)$$
$$\forall\, x\, (x + 0 = x)$$

Intuitively, these are all true.

Writing arithmetic statements in FOL may seem overly terse, but consider the following classical statement:

> *There do not exist four numbers, the last being larger than two, such that the sum of the first two, both raised to the power of the fourth, are equal to the third, also raised to the power of the fourth.*

In slightly more modern parlance this turns into

> *There are no positive integers* x, y, z *and* n, *where* $n > 2$, *such that* $x^n + y^n = z^n$.

and is now recognizable as Fermat's Last theorem.

This combination of ordinary language and formal expressions is the de facto gold-standard in mathematics and computer science; essentially all the literature is written this way.

We can take one more step and translate into FOL. Starting from the last sentence, this translation is not hard:

$$\neg\exists\, x, y, z, n \in \mathbb{N}^{+}\, (x^n + y^n = z^n \wedge n > 2)$$

The variables here are assumed to range over the positive integers as indicated by the annotation at the quantifier.

Even if you prefer the standard notation over the more compact one (and most people do), the latter is clearly better suited as input to any kind of algorithm – it is much easier to parse. Always remember the Entscheidungsproblem.

As another example of the expressiveness of FOL consider the venerable Principle of Induction. We can write it down as a formula as follows.

$$R(0) \wedge \forall\, x \left(R(x) \rightarrow R(x+1)\right) \rightarrow \forall\, x\, R(x)$$

Here we have added another unary relation symbol R to our language. So $R(x)$ asserts that number x has some unspecified property.

The Principle of Induction then simply asserts that is formula is true over the natural numbers. Note, though, that in many applications $R(x)$ would actually be replaced by a formula of arithmetic such as

$$\sum_{i \leq x} i = x(x+1)/2$$

The formal expression reflects very nicely the various components of an induction argument.

$$
\begin{aligned}
& R(0) \wedge && \text{base case} \\
& \forall\, x \,(\; R(x) && \text{induction hypothesis} \\
\rightarrow\; & R(x+1)\;) && \text{induction step} \\
& \Longrightarrow \\
& \forall\, x\, R(x) && \text{conclusion}
\end{aligned}
$$

As usual, the hard part is to show that $R(x) \rightarrow R(x+1)$ without any further assumptions about x.

In the language of arithmetic we can write a formula prime(x) that expresses the assertion "x is prime".

$$1 < x \wedge \forall\, u, v\, (x = u \cdot v \rightarrow u = 1 \vee v = 1)$$

Note that x here is a free variable (not quantified over) and thus is not bound to any particular value or range of values.

If we replace x by $3 = 1 + 1 + 1$ we get a true statement. But if we replace x by $4 = 1 + 1 + 1 + 1$ we get a false statement.

So, we can use free variables to pick out elements with certain properties.

In the assertion that there are infinitely many primes x appears as a bound variable.

$$\forall\, z\, \exists\, x\, (x > z \wedge \mathsf{prime}(x))$$

- First-Order Logic

2. Syntax

- Model Theory

While we are not interested in writing a parser or theorem prover, we still need to be a bit more careful about the syntax of our language of FOL. The components of a formula can be organized into a taxonomy like so:

- variables, constants and terms,
- equations,
- atomic formulae,
- propositional connectives, and
- quantifiers.

Every programming language has a defining report (which no one ever reads, other than perhaps compiler writers, it's the document that uses the imperative "shall" a lot), so think of this as the defining report for FOL.

Only the quantification part is really new; propositional connectives will be the same as in propositional logic while terms and equations will be the same as in equational logic.

We need a supply of variables Var as well as function symbols and predicate symbols. These form a graded alphabet $\Sigma = \Sigma_0 \cup \Sigma_1$. where every function symbol and relation symbol has a fixed number of arguments, determined by a arity map:

$$\text{ar} : \Sigma \to \mathbb{N}$$

We write $\mathcal{L}(\Sigma)$ for the language constructed from signature Σ.

Function symbols of arity 0 are constants and relation symbols of arity 0 are Boolean values (true or false) and will always be written $\top$ and $\bot$.

In any concrete application, Σ will be finite. However, in the literature you will often find a big system approach that introduces a countable supply of function and relation symbols for each arity. For our purposes a signature custom-designed for a particular application is more helpful.

Definition

The set $\mathcal{T} = \mathcal{T}(\Sigma) = \mathcal{T}(\mathsf{Var}, \Sigma)$ of all terms is defined by

- Every variable is a term.
- If f is an n-ary function symbol, and $t_1, \ldots, t_n$ are terms, $n \geq 0$, then $f(t_1, \ldots, t_n)$ is also a term.

A ground term is a term that contains no variables.

Note that $f()$ is a term for each constant (0-ary function symbol) f. For clarity, we write a, b, c and the like for constants.

The idea is that every ground term corresponds to a specific element in the underlying structure. For arbitrary terms we first have to replace all variables by constants. For example, in arithmetic the term $(1 + 1 + 1) \cdot (1 + 1)$ corresponds to the natural number 6. We could introduce constants for all the natural numbers, but there is no need to do so: we can build a corresponding term from the constant 1 and the binary operation $+$.

Given a few terms, we can apply a predicate to get a basic assertion (like $x + 4 < y$).

Definition

An atomic formula is an expression of the form

$$R(t_1, \ldots, t_n)$$

where R is an n-ary relation symbol, and the $t_1, \ldots, t_n$ are terms.

These are basically the atomic assertions in propositional logic: once we have values for the variables that might appear in the terms, an atomic formula can be evaluated to true or false.

But note that we do need bindings for the variables, $R(x, y)$ per se has no truth value.

Equality plays a special role in our setup. If we wanted to be extra careful, we would select a special binary relation symbol $\approx$ in our language, with the intent that

$$s \approx t$$

means that terms s and t denote the same element.

Thus, unlike with all the other relation symbols, the meaning of $\approx$ is fixed once and for all: it is always interpreted as equality.

We will soon cave and write the standard equality symbol $=$ rather than $\approx$, it is always clear from context what is meant.

The system just described is first-order logic with equality.

One can also consider first-order logic without equality (there is no direct way of asserting equality of terms).

Lastly, one can consider the fragment of FOL that has no function symbols nor equality, only relations (the actual predicate logic).

In the presence of equality, one still fake functions to a degree by considering relations F such that

$$\forall\, x\, \exists\, y\, F(x, y) \wedge \forall\, x, u, v\, (F(x, u) \wedge F(x, v) \rightarrow u = v)$$

Definition

The set of formulae of FOL is defined by

- Every atomic formula is a is a formula.
- If φ and ψ are formulae, so are $(\neg\varphi)$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, and $(\varphi \to \psi)$.
- If φ is a formula and x a variable, then $(\exists\, x\, \varphi)$ and $(\forall\, x\, \varphi)$ are also formulae.

So these are compound formulae versus the atomic formulae from above.

The φ in $(\exists\, x\, \varphi)$ and $(\forall\, x\, \varphi)$ is called the matrix of the formula.

Note that a term by itself is not a formula: it denotes an element (if it has no free variables) rather than a truth value.

Since we are not compilers, we omit unnecessary parentheses and write formulae such as

$$\forall\, x\, (R(x) \to \exists\, y\ S(x, y))$$

The intended meaning is: "For any x, if x has property R, then there exists a y such that x and y are related by S."

Binary relation and function symbols are often written in infix notation, using standard mathematical symbols: As usual, one often uses infix notation for terms and atomic formulae: $x + y < z$ is easier to read than $<(+(x, y), z)$.

One often contracts quantifiers of the same kind into one block.

$$\forall\, x\, \forall\, y\, \exists\, z\, (\approx\, (+(x, z), y))$$

Sloppy, but eminently readable, style

$$\forall\, x, y\, \exists\, z\, (x + z \approx y)$$

It is entirely standard to write $\varphi \to \psi$.

Alas, this notation becomes less than ideal when we also deal with functions or relations $f : A \to B$ or in the context of diagrams.

We will write $\varphi \Rightarrow \psi$ if we want to make sure the arrow is understood as implication.

Other old-fashioned notation: $\mathbf{C}(\varphi, \psi)$ and $\varphi \supset \psi$.

We assume that implication associates to the right, so

$$\varphi \to \psi \to \chi \text{ means } \varphi \to (\psi \to \chi)$$

Definition

A variable that is not in the range of a quantifier is free or unbound (as opposed to bound). We write $\mathrm{FV}(\varphi)$ for the set of free variables in φ. A formula without free variables is closed, or a sentence.

One often indicates free variables like so:

$\varphi(x, y)$	x and y may be free
$\exists\, x\, \varphi(x, y)$	only y may be free
$\forall\, y \exists\, x\, \varphi(x, y)$	closed.

Substitutions of terms for free variables are indicated like so:

$\varphi(s, t)$	replace x by s, y by t
$\varphi[s/x, t/y]$	replace x by s, y by t

We will explain shortly how to compute the truth value of a sentence.

The notation

$$\varphi(x_1, \ldots, x_n)$$

only expresses the fact that $\mathsf{FV}(\varphi) \subseteq \{x_1, \ldots, x_n\}$.

Note: this does not mean that they all actually occur, some or even all of them may be missing. This turns out to be preferable over the alternative.

This is really no different from saying that $2x^2 - y$ is a polynomial in variables x, y and z.

Exercise

Construct an algorithm that computes free variables and determines the scope of quantifiers.

A priori, a formula $\varphi(x)$ with a free variable x has no truth value associated with it: we need to replace x by a ground term to be able to evaluate.

However, taking inspiration from equational logic, it is convenient to define the truth value a formula with free variables to be the same as its universal closure: put universal quantifiers in front, one for each free variable.

$$\forall\, x, y, z\, \big(x * (y * z) \approx (x * y) * z\big)$$

is somewhat less elegant and harder to read than

$$x * (y * z) \approx (x * y) * z$$

In algebra, the latter form is much preferred.

Note that according to our definition it is perfectly agreeable to quantify over an already bound variable. To make the formula legible one then has to rename variables. For practical reasons it is best to simply disallow clashes between free and bound variables.

$$\forall\, x \left(R(x) \wedge \exists\, x \forall\, y\ S(x, y)\right)$$

Better: rename the x inside

$$\forall\, x \left(R(x) \wedge \exists\, z \forall\, y\ S(z, y)\right)$$

These issues are very similar to problems that arise in programming languages (global and local variables, scoping issues). They need to be addressed but are not of central importance.

Our definition of a formula of FOL is sufficiently precise to reason about the logic, but it is still a bit vague if we try to actually implement an algorithm that operates on these formulae.

Exercise

Explain how to implement formulae in FOL. Describe the data structure and make sure that it can be manipulated in a reasonable way.

Exercise

Give a precise definition of free and bound variables by induction on the buildup of the formula.

Exercise

Implement a renaming algorithm that removes potential scoping clashes in the variables of a formula.

Exercise

Implement a substituting algorithm that replaces a free variable in a formula by a term.

- First-Order Logic
- Syntax

3. Model Theory

How can we explain precisely what it means for an assertion to be true, to be valid? We would like some sweeping, global definition of truth that handles all areas of discourse, at least in mathematics and computer science. While this broad approach corresponds nicely to one's philosophical assumptions about the world (at least for an unreconstructed Platonist like myself) it leads to some rather dicey problems: for example, what should we do with assertions like

> *"This sentence is false."*

We certainly would want to have the ability to declare certain assertions such as "100 is a prime number" as false.

How can we then avoid paradoxes as in the self-referential statement above?

In order to succeed one needs to distinguish carefully between an object language and a meta-language. It is also helpful to restrict one's attention to truth in a particular domain such as, say, group theory or number theory. In the 1930s Alfred Tarski was the first to seriously tackle to problem of explaining truth for a sentence in a formal language.

We already have a nice formal language, though so far a formula is just a syntactic object, think of it as a string or a parse tree. We can now use Tarski's ideas to attach meaning to a formula, to establish a relationship with the world of actual objects such as numbers, functions, hash tables, algorithms, and so on.

The key is to define the truth value of a formula relative to a given structure, a mini-universe that allows us to make sense out of the components of the formula.

Here is a simple example from basic arithmetic (over the natural numbers). Consider the formula

$$\forall\, x\, \exists\, y\, (x < y)$$

Its components are

x, y	variables, range over natural numbers
$<$	comparison, the less-than relation

The intended meaning of the formula, expressed in classical math-speak, is then

> *For any natural number x there exists a natural number y such that x is less than y.*

So this formula is clearly true, valid.

The key problem is that the formula $\forall\, x \,\exists\, y \,(x < y)$ itself does not express any of the auxiliary information:

- There is no indication that the variables range over natural numbers, and
- there is no indication that the binary relation symbol $<$ corresponds to the standard order relation on the naturals.

This would become more clear if we had written $\forall\, x \,\exists\, y \, R(x, y)$ instead, the use of a well-known symbol such as $<$ is just syntactic sugar.

Moreover, if we were to interpret the variables as ranging over the integers instead (a relatively minor change) the formula would be false, invalid.

To settle all these issues we have to consider so-called structures over which the formulae of FOL can be evaluated.

So what exactly are these structures that we need to interpret a formula in FOL? Fix some language $\mathcal{L} = \mathcal{L}(\Sigma)$ where Σ is a graded alphabet as above.

Definition

A (first order) structure is a set together with a collection of functions and relations on that set. The signature of a first order structure is the list of arities of its functions and relations.

In order to interpret formulae in $\mathcal{L}(\Sigma)$ the signatures have to match, which we will tacitly assume from now on. So a structure in general looks like so:

$$\mathcal{A} = \langle\, A; f_1, f_2, \ldots, R_1, R_2, \ldots \rangle$$

The set A is the carrier set of the structure. Unary and binary functions and relations are by far the most important in applications, but higher arities may occur.

Note that a first order structure is not all that different from a data type. To wit, we are dealing with a

- collection of objects,
- operations on these objects, and
- relations on these objects.

In the case where the carrier set is finite (actually, finite and small) we can in fact represent the whole FO structure by a suitable data structure (for example, explicit lookup tables). For infinite carrier sets, things are a bit more complicated.

Data types (or rather, their values) are manipulated in programs, we are here interested in describing properties of structures using the machinery of FOL.

Given a formula and a structure of the same signature we can associate the function and relation symbols in the formula with real functions and relations in a structure (of the same arity).

$$f \text{ function symbol} \quad \rightsquigarrow \quad f^{\mathcal{A}} \text{ a function in } \mathcal{A}$$

$$R \text{ function symbol} \quad \rightsquigarrow \quad R^{\mathcal{A}} \text{ a relation in } \mathcal{A}$$

Given this interpretation of function and relation symbols over $\mathcal{A}$, we can determine whether a formula holds true over $\mathcal{A}$.

This is very different from trying to establish universal truth. All we are doing here is to confirm that, in the context of a particular structure, a certain formula is valid. As it turns out, this is all that is really needed in the real world.

Of course, when the structure changes the formula may well become false.

Consider arithmetic: The language is $\mathcal{L}(+, \cdot, 0, 1; <)$ and has type $(2, 2, 0, 0; 2)$.

We can interpret a formula in this language over any structure of the same type. Of course, the most important structure for arithmetic is

$$\mathcal{N} = \langle \mathbb{N}; +, \cdot, 0, 1, < \rangle$$

the set of natural numbers together with the standard operations but we will see that there are others.

Note the slight abuse of notation here (as is standard practice). More precise would be to write: The function symbol $+$ is interpreted by $+^{\mathcal{N}}$, the standard operation of addition of natural numbers.

For our purposes there is no gain in being quite so careful.

With this interpretation over $\mathcal{N}$, the formula

- $0 < 1$ is true
- $\forall\, x \exists\, y\, (x < y)$ is true
- $\forall\, x, y\, (x < y \to \exists\, z\, (x < z \wedge z < y))$ is false

How about the primality formula

$$\varphi(x) = 1 < x \wedge \forall\, y, z\, (x \approx y \cdot z \to y \approx 1 \vee z \approx 1)$$

This formula has a free variable x, so we need to bind x (replace it by a term) before we can determine truth.

We use the numeral $\underline{n} = \underbrace{1 + 1 + \ldots + 1}_{n\text{times}}$ to represent the natural number n as a term. The set of primes is then

$$\{\, n \mid \varphi(\underline{n}) \text{ holds in } \mathcal{N} \,\}.$$

Strictly speaking, $\underline{n} = \underbrace{1 + 1 + \ldots + 1}_{n\text{times}}$ is also sloppy, we really should define

$$\underline{0} = 0 \quad \text{the term, not the number}$$
$$\underline{n+1} = +(\underline{n}, 1)$$

by induction, so that $\underline{n}$ is a bonified term of our system.

One often avoids this level of precision, but for any algorithmic treatment there is no way around it.

Suppose we interpret our formulae over the structure of the real numbers instead. This is possible since it has same signature $(2, 2, 0, 0; 2)$:

$$\mathcal{R} = \langle \mathbb{R}; +, \cdot, 0, 1, < \rangle$$

Now $+$ refers to addition of reals, 0 is the real number zero, and so on. These operations are much more complicated, but as far as our formula is concerned all that matters is that they have the right arity.

Over the reals the density formula

$$\forall\, x, y\, (x < y \to \exists\, z\, (x < z \wedge z < y)$$

holds.

On the other hand, the primality statement $\varphi(x)$ is not interesting over $\mathcal{R}$: there is no binding for x that makes it true.

We can now give a first informal definition of truth or validity.

Definition

A formula of FOL is valid if it holds over any structure of the appropriate signature.

For free variables this definition is motivated by formulae such as $x * y \approx y * x$: we want this to mean that the operation $*$ is commutative. Note, though, that we apply the same approach to function and relation symbols (other than equality): any possible interpretation has to be taken into account

Of course, we can pin down some of the properties of the corresponding functions and relations in the formula itself, but that's all we can do. There is no magic that says that the symbol $+$ must always be interpreted as some sort of addition over some algebraic structure.

It is clear that a formula like $\varphi \to \varphi$ is valid. In fact, if we replace the propositional variables in any tautology by arbitrary sentences we obtain a valid sentence of FOL. More precisely, let

$$\varphi(p_1, p_2, \ldots, p_n)$$

be a tautology with propositional variables $p_1, p_2, \ldots, p_n$. Let $\psi_1, \ldots, \psi_n$ be arbitrary sentences of FOL. Then

$$\varphi(\psi_1, \psi_2, \ldots, \psi_n)$$

is a valid sentence of FOL. True, but not too interesting.

Again in analogy to propositional logic we can define satisfiability.

Definition

A formula of FOL is satisfiable if it is true for some interpretation of the variables, functions and relations. It is a contradiction if it is true for no interpretation of the variables, functions and relations.

For example, in the language of binary relations the formula

$$x\ R\ x \wedge (x\ R\ y \wedge y\ R\ z \rightarrow x\ R\ z)$$

is satisfied by any structure $\mathcal{A}$ that carries an irreflexive transitive relation $R^{\mathcal{A}}$. On the other hand,

$$\forall\, x\, (x \not\approx c)$$

where c is a constant is a contradiction: we can interpret x as the element in the structure denoted by c in which case equality holds.

Slightly more complicated examples for valid formulae are

$$\forall x \forall y \varphi(x, y) \to \forall y \forall x \varphi(x, y)$$
$$\exists x \exists y \varphi(x, y) \to \exists y \exists x \varphi(x, y)$$
$$\exists x \forall y \varphi(x, y) \to \forall y \exists x \varphi(x, y)$$

Note, though, that the following is not valid:

$$\forall x \exists y \varphi(x, y) \to \exists y \forall x \varphi(x, y)$$

Exercise

Verify that the first three formulae are true and come up with an example that shows that the last one is not.

Another good source of valid formulae are assertions about the number of elements in the underlying structure. How do we say "there are exactly n elements in the ground set" in FOL? First, a formula which states that there are at most n elements.

$$\mathsf{EX}_{\leq n} = \exists\, x_1, \ldots, x_n \,\forall\, y\, (y \approx x_1 \vee \ldots \vee y \approx x_n)$$

Second, a formula which states that there are at least n elements.

$$\mathsf{EX}_{\geq n} = \exists\, x_1, \ldots, x_n\, (x_1 \not\approx x_2 \wedge x_1 \not\approx x_3 \wedge \ldots \wedge x_{n-1} \not\approx x_n)$$

All these formulae are clearly satisfiable. The conjunction $\mathsf{EX}_{\leq n} \wedge \mathsf{EX}_{\geq n}$ pins down the cardinality to exactly n. Also, a formula

$$\mathsf{EX}_{\geq n} \rightarrow \mathsf{EX}_{\geq m}$$

is valid whenever $m \leq n$.

How about a formula that states that there are infinitely many elements?

$$\mathsf{EX}_{\geq 1} \wedge \mathsf{EX}_{\geq 2} \wedge \ldots \wedge \mathsf{EX}_{\geq n} \wedge \ldots$$

does not work since it is not a finite formula. The attempt

$$\forall\, n\, \mathsf{EX}_{\geq n}$$

also fails; we cannot quantify over formulae in our logic.

Exercise

Explain precisely why these "formulae" are not admissible in FOL.

After some more fruitless attempts one might suspect that the statement "there are infinitely many thingies" cannot be expressed in FOL.

Wrong! Let f be a unary function symbol and c a constant. Consider

$$\varphi = \forall\, x\, (f(x) \not\approx c) \wedge \forall\, x, y\, (f(x) \approx f(y) \rightarrow x \approx y).$$

So φ states that f is not surjective but injective. Hence in any interpretation that makes φ true the carrier set must be infinite.

Exercise

Use a total order to produce another formulae in FOL that forces the ground set to be infinite.

Again, there are natural decision problems associated with this classification.

Problem:	**Validity**
Instance:	A FOL formula φ.
Question:	Is φ valid?

Problem:	**Satisfiability**
Instance:	A FOL formula φ.
Question:	Is φ satisfiable?

As usual, there is the search version of Satisfiability: we would like to construct a satisfying interpretation if one exists. Note that this may well entail the construction of an infinite structure.

As one might suspect from the few examples, these problems are much harder in FOL than in propositional logic and will turn out to be highly undecidable in general.

BTW, in CS this is called model checking.

Time to give a precise definition of validity and satisfiability. We begin by defining assignments in the context of FOL.

Definition

An assignment or valuation (over a structure $\mathcal{A}$) associates variables of the language with elements in the ground set A.

Given an assignment $\sigma : \text{Var} \to A$, we can associate an element $\sigma(t)$ in A with each term t.

- $t = x$: then $\sigma(t) = \sigma(x)$
- $t = f(r_1, \ldots, r_n)$: then $\sigma(t) = f^{\mathcal{A}}(\sigma(r_1), \ldots, \sigma(r_n))$

Example

Over $\mathcal{N}$ let $\sigma(x) = 3$. Then $\sigma(x \cdot (1 + 1))) = 6$ whereas $\sigma(x) = 0$ produces $\sigma(x \cdot (1 + 1))) = 0$.

Once we have an assignment for all the free variables in an atomic formula we can determine a truth value for it.

Definition

Let σ be an assignment over a structure $\mathcal{A}$ and $\varphi = R(t_1, \ldots, t_n)$ an atomic formula. Define the truth value of φ (under σ over $\mathcal{A}$) to be

$$\mathcal{A}_\sigma(\varphi) = \begin{cases} \text{tt} & \text{if } R^{\mathcal{A}}(\sigma(t_1), \ldots, \sigma(t_n)) \text{ holds,} \\ \text{ff} & \text{otherwise.} \end{cases}$$

Example

Over the natural numbers $\mathcal{N}$ suppose $\sigma(x) = 0$ and $\sigma(y) = 1$. Then

$$\mathcal{N}_\sigma(x + y < 1 + 1) = \mathcal{N}_\sigma(0 + 1 < 1 + 1) = \text{tt}$$

but for $\sigma(x) = \sigma(y) = 1$ we get

$$\mathcal{N}_\sigma(x + y < 1 + 1) = \mathcal{N}_\sigma(1 + 1 < 1 + 1) = \text{ff}$$

Once we have a truth value for atomic formulae, we can extend this evaluation to compound formulae without quantifiers.

Definition (Propositional Connectives)

- $\varphi = \psi \wedge \chi$: then $\mathcal{A}_\sigma(\varphi) = H_{and}(\mathcal{A}_\sigma(\psi), \mathcal{A}_\sigma(\chi))$
- $\varphi = \psi \vee \chi$: then $\mathcal{A}_\sigma(\varphi) = H_{or}(\mathcal{A}_\sigma(\psi), \mathcal{A}_\sigma(\chi))$
- $\varphi = \neg\psi$: then $\mathcal{A}_\sigma(\varphi) = H_{not}(\mathcal{A}_\sigma(\psi))$

Example

Suppose $\sigma(x) = 0$ and $\sigma(y) = 1$. Then

$$\begin{aligned}\mathcal{N}_\sigma(x < y \vee y < x) &= H_{or}(\mathcal{N}_\sigma(x < y), \mathcal{N}_\sigma(y < x)) \\ &= H_{or}(\mathcal{N}(0 < 1), \mathcal{N}(1 < 0)) \\ &= H_{or}(\mathsf{tt}, \mathsf{ff}) \\ &= \mathsf{tt}\end{aligned}$$

For an assignment σ, let us write $\sigma[a/x]$ for the assignment that is the same as σ everywhere, except that $\sigma[a/x](x) = a$. Think of this as substituting "a for x".

Definition (Quantifiers)

- $\varphi = \exists\, x\, \psi$:
  Then $\mathcal{A}_\sigma(\varphi) = \mathsf{tt}$ if there is an a in A such that $\mathcal{A}_{\sigma[a/x]}(\varphi) = \mathsf{tt}$.
- $\varphi = \forall\, x\, \psi$:
  Then $\mathcal{A}_\sigma(\varphi) = \mathsf{tt}$ if for all a in A $\mathcal{A}_{\sigma[a/x]}(\varphi) = \mathsf{tt}$.

Note that σ only needs to be defined on the free variables of φ to produce a truth value for φ, the values anywhere else do not matter. If φ is a sentence, σ can be totally undefined.

Definition (Formulae)

A formula φ is valid in $\mathcal{A}$ under assignment σ if $\mathcal{A}_\sigma(\varphi) = 1$.

A formula φ is valid in $\mathcal{A}$ if it is valid in $\mathcal{A}$ for all assignments σ. The structure $\mathcal{A}$ is then said to be a model for φ or to satisfy φ.

A sentence is valid (or true) if it is valid over any structure (of the appropriate signature).

Notation:

$$\mathcal{A} \models_\sigma \varphi, \qquad \mathcal{A} \models \varphi, \qquad \models \varphi$$

One uses the same notation for sets of formulae Γ. So $\mathcal{A} \models \Gamma$ means that $\mathcal{A} \models \varphi$ for all $\varphi \in \Gamma$.

Note the condition for validity: the formula has to hold in all structures.

Definition

A formula is satisfiable if there is some structure $\mathcal{A}$ and some assignment σ for all the free variables in φ such that $\mathcal{A} \models_\sigma \varphi$.

In other words, the existentially quantified formula

$$\exists\, x_1, \ldots, x_n \; \varphi(x_1, \ldots, x_n)$$

has a model, where $x_1, \ldots, x_n$ are all the free variables of φ.

In shorthand: $\exists\, \boldsymbol{x}\, \varphi(\boldsymbol{x})$.

This is analogous to validity where we insist that $\forall\, x_1, \ldots, x_n \; \varphi(x_1, \ldots, x_n)$ holds, or $\forall\, \boldsymbol{x}\, \varphi(\boldsymbol{x})$ in compact notation.

The definitions of truth given here is due to A. Tarski (two seminal papers, one in 1933 and a second one in 1956, with R. Vaught).

A frequent objection to this approach is that we are using "for all" to define what a universal quantifier means.

True, but the formulae in our logic are syntactic objects, and define their meaning in terms of structures, which are not syntactic, they are real (in the world of mathematics and TCS).

Think of this as a program: you can "compute" the truth value of a formula, as long as you can perform certain operations in the structure (evaluate $f^{\mathcal{A}}$, loop over all elements, search over all elements, ...). This is a bit problematic over infinite structures, but for finite ones we can actually perform the computation (at least if we ignore efficiency).

More precisely, suppose we wanted to construct an algorithm

$$\mathsf{ValidQ}(\mathcal{A}, \varphi)$$

that checks if formula φ is valid over structure $\mathcal{A}$.

What would be the appropriate input for such an algorithm? The easy part is the formula: any standard representation (string, parse tree, sequence number) will be fine.

The real problem is the structure $\mathcal{A}$.

For standard structures such as the natural numbers or reals we understand (more or less) how to interpret the operations. But in the general case we need some representation.

The formula is easily represented as a data structure (some kind of parse tree is a good choice).

But for data complexity and combined complexity there is a problem with the structure $\mathcal{A}$: we need to specify a first-order structure in some finitary way.

Next time we will take a look at so-called automatic structures where $\mathcal{A}$ is described by finite state machines.

Suppose $\mathcal{A}$ is a structure of some signature Σ. In order to describe $\mathcal{A}$ the first step is to augment the language $\mathcal{L}(\Sigma)$ by constant symbols c_a for each element a in the carrier set A of $\mathcal{A}$, obtaining a new signature $\Sigma_{\mathcal{A}}$.

$\mathcal{A}$ is naturally also a structure of signature $\Sigma_{\mathcal{A}}$.

This step is not necessary when there already are terms in the language for all the elements of the structure. E.g., in arithmetic we can denote every natural number by a ground term

$$\underline{n} = 1 + 1 + \ldots + 1$$

This works since we are dealing with the naturals, but in general there is no reason why every element in a structure should be denoted by a term.

For example, suppose we want to deal with the reals. The standard language has constants for 0 and 1 but nothing else.

$$\mathcal{R} = \langle\, \mathbb{R}; +, \cdot, 0, 1; < \,\rangle$$

Using the numeral trick, we can obtain terms for rationals (at least if we add division), but no more.

Just to write down all available facts about $\mathbb{R}$ we need to add uncountably many constants, one for each real other than the rationals. This is no problem at all in set theory.

But it wrecks the language: the new constants are not finitary data structures. We cannot even build a parser.

Definition

The (atomic) diagram of a structure of some fixed signature is the set of all atomic sentences and their negations in $\mathcal{L}(\Sigma_{\mathcal{A}})$ that are valid in $\mathcal{A}$.

In symbols: $\mathrm{diag}\mathcal{A}$.

The point is that the validity of any formula over $\mathcal{A}$ is completely determined by $\mathrm{diag}\mathcal{A}$: no other information is used in our definition of truth. Hence our algorithm should take as inputs the diagram and the formula and use recursion to obtain the answer:

$$\mathrm{ValidQ}(\mathrm{diag}\mathcal{A}, \varphi)$$

If the underlying structure $\mathcal{A}$ is finite (and thus the diagram is finite), then we can actually perform this computation: it is just recursion and copious table lookups. In fact, ValidQ will be primitive recursive given any reasonable coding.

How do we represent the atomic diagram diag$\mathcal{A}$? Given enough constants we can simply write down table.

E.g., for a unary function symbol f we can use a table with entries c_a and c_b provided that $b = f^{\mathcal{A}}(a)$. Each entry corresponds to an identity $f(c_a) \approx c_b$ in the diagram.

For binary function symbols we get a classical Cayley style "multiplication table", and higher dimensional tables for functions of higher arity.

Relations can be handled by similar tables with entries in $\mathbb{B}$. This corresponds to the familiar interpretation of a relation $R \subseteq A^k$ as a function $R : A^k \to \mathbb{B}$.

So, the whole diagram is just a bunch of tables using special constants for all elements in the structure and Boolean values. For small structures this is a perfectly good representation, though even for finite but large structures explicit tables are not feasible.

Pushing ahead into the realm of infinite structures things become much more complicated. If the carrier set is uncountable our machinery from classical computability theory simply does not apply – the individual elements are not finitary objects and we have no handle.

However, if the carrier set is countable computability theory does apply and we can use it to measure the complexity of the structure.

Definition

The (complete) diagram is the collection of all sentences in $\mathcal{L}(\Sigma_{\mathcal{A}})$ that are valid in $\mathcal{A}$. In symbols: $\mathrm{diag}^c\mathcal{A}$.

$\mathcal{A}$ is computable if its atomic diagram is decidable. $\mathcal{A}$ is decidable if its complete diagram is decidable.

All finite structures are trivially decidable and even primitive recursive, though things become more complicated if we consider lower levels of the computational hierarchy.

For example, satisfiability of a propositional formula is easily expressed as a validity problem of a formula over a two-element structure, yet no polynomial time algorithm is known for this problem.

Also, often finite structures are parametrized (for example by the number of elements) and one really would like a solution uniformly in terms of the parameter. This is usually much harder than staring at a single finite structure.

The standard structure of the natural numbers is a good example for a computable structure that fails to be decidable. The atomic sentences here come down to assertions of the type

$$(1+1) * \underline{3} \approx \underline{6}$$

and the like and are trivially decidable. Indeed, this all comes down to evaluating polynomials over the integers and can be handled very easily.

Alas, the addition of quantifiers destroys any hope for effective decision methods: even just a single existential quantifier renders the validity problem undecidable.

In fact, truth over the structure of natural numbers is separated from decidability by infinitely many levels of a hierarchy of complexity.

Note that the equality symbol $\approx$ is treated differently from other binary relation symbols: it is always interpreted as actual equality over $\mathcal{A}$. Hence the following formulae are all valid and are tyically used as axioms for equality.

$$\forall x\,(x \approx x)$$
$$\forall x \forall y\,(x \approx y \to y \approx x)$$
$$\forall x \forall y \forall z\,(x \approx y \wedge y \approx z \to x \approx z)$$

$$\forall \boldsymbol{x} \forall \boldsymbol{y}\,(\boldsymbol{x} \approx \boldsymbol{y} \to (R(\boldsymbol{x}) \leftrightarrow R(\boldsymbol{y})))$$
$$\forall \boldsymbol{x} \forall \boldsymbol{y}\,(\boldsymbol{x} \approx \boldsymbol{y} \to f(\boldsymbol{x}) \approx f(\boldsymbol{y}))$$

Exercise

Explain what would happen if we adopt these formulae as axioms for an otherwise undistinguished binary relation symbol $\approx$. Describe the structures that satisfy these axioms.