First-Order Logic

First-Order logic is a generalization of propositional logic. Propositional logic can represent propositions, whereas first-order logic can represent individuals and propositions about individuals.

For example, in propositional logic from “Socrates is a man” and “If Socrates is a man then Socrates is mortal” the conclusion “Socrates is mortal” can be drawn.

In first-order logic this can be represented much more fine-grained. From “Socrates is a man” and “All man are mortal” the conclusion “Socrates is mortal” can be drawn.

3.1.1 Definition (Many-Sorted Signature)

A *many-sorted signature* $\Sigma = (\mathcal{S}, \Omega, \Pi)$ is a triple consisting of
a finite non-empty set $\mathcal{S}$ of *sort symbols*,
a non-empty set Ω of *operator symbols* (also called *function symbols*) over $\mathcal{S}$ and
a set Π of *predicate symbols*.

3.1.1 Definition (Many-Sorted Signature Ctd)

Every operator symbol $f \in \Omega$ has a unique sort declaration $f : S_1 \times \ldots \times S_n \to S$, indicating the sorts of arguments (also called *domain sorts*) and the *range sort* of f, respectively, for some $S_1, \ldots, S_n, S \in \mathcal{S}$ where $n \geq 0$ is called the *arity* of f, also denoted with $\text{arity}(f)$. An operator symbol $f \in \Omega$ with arity 0 is called a *constant*.

Every predicate symbol $P \in \Pi$ has a unique sort declaration $P \subseteq S_1 \times \ldots \times S_n$. A predicate symbol $P \in \Pi$ with arity 0 is called a *propositional variable*. For every sort $S \in \mathcal{S}$ there must be at least one constant $a \in \Omega$ with range sort S.

3.1.1 Definition (Many-Sorted Signature Ctd)

In addition to the signature Σ, a variable set $\mathcal{X}$, disjoint from Ω is assumed, so that for every sort $S \in \mathcal{S}$ there exists a countably infinite subset of $\mathcal{X}$ consisting of variables of the sort S. A variable x of sort S is denoted by x_S.

3.1.2 Definition (Term)

Given a signature $\Sigma = (\mathcal{S}, \Omega, \Pi)$, a sort $S \in \mathcal{S}$ and a variable set $\mathcal{X}$, the set $T_S(\Sigma, \mathcal{X})$ of all *terms* of sort S is recursively defined by (i) $x_S \in T_S(\Sigma, \mathcal{X})$ if $x_S \in \mathcal{X}$, (ii) $f(t_1, \ldots, t_n) \in T_S(\Sigma, \mathcal{X})$ if $f \in \Omega$ and $f : S_1 \times \ldots \times S_n \to S$ and $t_i \in T_{S_i}(\Sigma, \mathcal{X})$ for every $i \in \{1, \ldots, n\}$.

The sort of a term t is denoted by $\text{sort}(t)$, i.e., if $t \in T_S(\Sigma, \mathcal{X})$ then $\text{sort}(t) = S$. A term not containing a variable is called *ground*.

For the sake of simplicity it is often written: $T(\Sigma, \mathcal{X})$ for $\bigcup_{S \in \mathcal{S}} T_S(\Sigma, \mathcal{X})$, the set of all terms, $T_S(\Sigma)$ for the set of all ground terms of sort $S \in \mathcal{S}$, and $T(\Sigma)$ for $\bigcup_{S \in \mathcal{S}} T_S(\Sigma)$, the set of all ground terms over Σ.

Note that the sets $T_S(\Sigma)$ are all non-empty, because there is at least one constant for each sort S in Σ. The sets $T_S(\Sigma, \mathcal{X})$ include infinitely many variables of sort S.

3.1.3 Definition (Equation, Atom, Literal)

If $s, t \in T_S(\Sigma, \mathcal{X})$ then $s \approx t$ is an *equation* over the signature Σ. Any equation is an *atom* (also called *atomic formula*) as well as every $P(t_1, \ldots, t_n)$ where $t_i \in T_{S_i}(\Sigma, \mathcal{X})$ for every $i \in \{1, \ldots, n\}$ and $P \in \Pi$, $\text{arity}(P) = n$, $P \subseteq S_1 \times \ldots \times S_n$.

An atom or its negation of an atom is called a *literal*.

Definition (Formulas)

The set $\text{FOL}(\Sigma, \mathcal{X})$ of *many-sorted first-order formulas with equality* over the signature Σ is defined as follows for formulas $\phi, \psi \in F_\Sigma(\mathcal{X})$ and a variable $x \in \mathcal{X}$:

$\text{FOL}(\Sigma, \mathcal{X})$	Comment
$\bot$	false
$\top$	true
$P(t_1, \ldots, t_n), s \approx t$	atom
$(\neg\phi)$	negation
$(\phi \circ \psi)$	$\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$
$\forall x.\phi$	universal quantification
$\exists x.\phi$	existential quantification

?? Definition (Positions)

The set of positions of a term, formula is inductively defined by:

$$
\begin{array}{rl}
\mathrm{pos}(x) & := \{\epsilon\} \text{ if } x \in \mathcal{X} \\
\mathrm{pos}(\phi) & := \{\epsilon\} \text{ if } \phi \in \{\top, \bot\} \\
\mathrm{pos}(\neg\phi) & := \{\epsilon\} \cup \{1p \mid p \in \mathrm{pos}(\phi)\} \\
\mathrm{pos}(\phi \circ \psi) & := \{\epsilon\} \cup \{1p \mid p \in \mathrm{pos}(\phi)\} \cup \{2p \mid p \in \mathrm{pos}(\psi)\} \\
\mathrm{pos}(s \approx t) & := \{\epsilon\} \cup \{1p \mid p \in \mathrm{pos}(s)\} \cup \{2p \mid p \in \mathrm{pos}(t)\} \\
\mathrm{pos}(f(t_1, \ldots, t_n)) & := \{\epsilon\} \cup \bigcup_{i=1}^{n} \{ip \mid p \in pos(t_i)\} \\
\mathrm{pos}(P(t_1, \ldots, t_n)) & := \{\epsilon\} \cup \bigcup_{i=1}^{n} \{ip \mid p \in pos(t_i)\} \\
\mathrm{pos}(\forall x.\phi) & := \{\epsilon\} \cup \{1p \mid p \in \mathrm{pos}(\phi)\} \\
\mathrm{pos}(\exists x.\phi) & := \{\epsilon\} \cup \{1p \mid p \in \mathrm{pos}(\phi)\}
\end{array}
$$

where $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$ and $t_i \in T(\Sigma, \mathcal{X})$ for all $i \in \{1, \ldots, n\}$.

An term t (formula ϕ) is said to *contain* another term s (formula ψ) if $t|_p = s$ ($\phi|_p = \psi$). It is called a *strict subexpression* if $p \neq \epsilon$. The term t (formula ϕ) is called an *immediate subexpression* of s (formula ψ) if $|p| = 1$. For terms a subexpression is called a *subterm* and for formulas a *subformula*, respectively.

The *size* of a term t (formula ϕ), written $|t|$ ($|\phi|$), is the cardinality of $\text{pos}(t)$, i.e., $|t| := |\text{pos}(t)|$ ($|\phi| := |\text{pos}(\phi)|$). The *depth* of a term, formula is the maximal length of a position in the term, formula: $\text{depth}(t) := max\{|p| \mid p \in \text{pos}(t)\}$ ($\text{depth}(\phi) := max\{|p| \mid p \in \text{pos}(\phi)\}$).

The set of *all* variables occurring in a term t (formula ϕ) is denoted by $\text{vars}(t)$ ($\text{vars}(\phi)$) and formally defined as
$\text{vars}(t) := \{x \in \mathcal{X} \mid x = t|_p, p \in \text{pos}(t)\}$
$(\text{vars}(\phi) := \{x \in \mathcal{X} \mid x = \phi|_p, p \in \text{pos}(\phi)\})$.
A term t (formula ϕ) is *ground* if $vars(t) = \emptyset$ ($vars(\phi) = \emptyset$). Note that $\text{vars}(\forall x.a \approx b) = \emptyset$ where a, b are constants. This is justified by the fact that the formula does not depend on the quantifier, see the semantics below. The set of *free* variables of a formula ϕ (term t) is given by $\text{fvars}(\phi, \emptyset)$ ($\text{fvars}(t, \emptyset)$) and recursively defined by $\text{fvars}(\psi_1 \circ \psi_2, B) := \text{fvars}(\psi_1, B) \cup \text{fvars}(\psi_2, B)$ where $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$, $\text{fvars}(\forall x.\psi, B) := \text{fvars}(\psi, B \cup \{x\})$, $\text{fvars}(\exists x.\psi, B) := \text{fvars}(\psi, B \cup \{x\})$, $\text{fvars}(\neg\psi, B) := \text{fvars}(\psi, B)$, $\text{fvars}(L, B) := \text{vars}(L) \setminus B$ ($\text{fvars}(t, B) := \text{vars}(t) \setminus B$.
For $\text{fvars}(\phi, \emptyset)$ I also write $\text{fvars}(\phi)$.

In $\forall x.\phi$ ($\exists x.\phi$) the formula ϕ is called the *scope* of the quantifier. An occurrence q of a variable x in a formula ϕ ($\phi|_q = x$) is called *bound* if there is some $p < q$ with $\phi|_p = \forall x.\phi'$ or $\phi|_p = \exists x.\phi'$. Any other occurrence of a variable is called *free*.

A formula not containing a free occurrence of a variable is called *closed*. If $\{x_1, \ldots, x_n\}$ are the variables freely occurring in a formula ϕ then $\forall x_1, \ldots, x_n.\phi$ and $\exists x_1, \ldots, x_n.\phi$ (abbreviations for $\forall x_1.\forall x_2 \ldots \forall x_n.\phi$, $\exists x_1.\exists x_2 \ldots \exists x_n.\phi$, respectively) are the *universal* and the *existential closure* of ϕ, respectively.

3.1.7 Definition (Polarity)

The *polarity* of a subformula $\psi = \phi|_p$ at position p is $pol(\phi, p)$ where *pol* is recursively defined by

$$\begin{aligned}
\text{pol}(\phi, \epsilon) &:= 1 \\
\text{pol}(\neg\phi, 1p) &:= -\text{pol}(\phi, p) \\
\text{pol}(\phi_1 \circ \phi_2, ip) &:= \text{pol}(\phi_i, p) \text{ if } \circ \in \{\wedge, \vee\} \\
\text{pol}(\phi_1 \rightarrow \phi_2, 1p) &:= -\text{pol}(\phi_1, p) \\
\text{pol}(\phi_1 \rightarrow \phi_2, 2p) &:= \text{pol}(\phi_2, p) \\
\text{pol}(\phi_1 \leftrightarrow \phi_2, ip) &:= 0 \\
\text{pol}(P(t_1, \ldots, t_n), p) &:= 1 \\
\text{pol}(t \approx s, p) &:= 1 \\
\text{pol}(\forall x.\phi, 1p) &:= \text{pol}(\phi, p) \\
\text{pol}(\exists x.\phi, 1p) &:= \text{pol}(\phi, p)
\end{aligned}$$

Semantics

3.2.1 Definition (Σ-algebra)

Let $\Sigma = (\mathcal{S}, \Omega, \Pi)$ be a signature with set of sorts $\mathcal{S}$, operator set Ω and predicate set Π. A Σ-*algebra* $\mathcal{A}$, also called Σ-*interpretation*, is a mapping that assigns (i) a non-empty carrier set $S^{\mathcal{A}}$ to every sort $S \in \mathcal{S}$, so that $(S_1)^{\mathcal{A}} \cap (S_2)^{\mathcal{A}} = \emptyset$ for any distinct sorts $S_1, S_2 \in \mathcal{S}$, (ii) a total function $f^{\mathcal{A}} : (S_1)^{\mathcal{A}} \times \ldots \times (S_n)^{\mathcal{A}} \to (S)^{\mathcal{A}}$ to every operator $f \in \Omega$, $\text{arity}(f) = n$ where $f : S_1 \times \ldots \times S_n \to S$, (iii) a relation $P^{\mathcal{A}} \subseteq ((S_1)^{\mathcal{A}} \times \ldots \times (S_m)^{\mathcal{A}})$ to every predicate symbol $P \in \Pi$, $\text{arity}(P) = m$. (iv) the equality relation becomes $\approx^{\mathcal{A}} = \{(e, e) \mid e \in \mathcal{U}^{\mathcal{A}}\}$ where the set $\mathcal{U}^{\mathcal{A}} := \bigcup_{S \in \mathcal{S}} (S)^{\mathcal{A}}$ is called the *universe* of $\mathcal{A}$.

A (variable) *assignment*, also called a *valuation* for an algebra $\mathcal{A}$ is a function $\beta : \mathcal{X} \to \mathcal{U}_{\mathcal{A}}$ so that $\beta(x) \in S_{\mathcal{A}}$ for every variable $x \in \mathcal{X}$, where $S = \mathrm{sort}(x)$. A *modification* $\beta[x \mapsto e]$ of an assignment β at a variable $x \in \mathcal{X}$, where $e \in S_{\mathcal{A}}$ and $S = \mathrm{sort}(x)$, is the assignment defined as follows:

$$\beta[x \mapsto e](y) = \begin{cases} e & \text{if } x = y \\ \beta(y) & \text{otherwise.} \end{cases}$$

The homomorphic extension $\mathcal{A}(\beta)$ of β onto terms is a mapping $T(\Sigma, \mathcal{X}) \to \mathcal{U}_{\mathcal{A}}$ defined as (i) $\mathcal{A}(\beta)(x) = \beta(x)$, where $x \in \mathcal{X}$ and (ii) $\mathcal{A}(\beta)(f(t_1, \ldots, t_n)) = f_{\mathcal{A}}(\mathcal{A}(\beta)(t_1), \ldots, \mathcal{A}(\beta)(t_n))$, where $f \in \Omega$, arity$(f) = n$.

Given a term $t \in T(\Sigma, \mathcal{X})$, the value $\mathcal{A}(\beta)(t)$ is called the *interpretation* of t under $\mathcal{A}$ and β. If the term t is ground, the value $\mathcal{A}(\beta)(t)$ does not depend on a particular choice of β, for which reason the interpretation of t under $\mathcal{A}$ is denoted by $\mathcal{A}(t)$.

An algebra $\mathcal{A}$ is called *term-generated*, if every element e of the universe $\mathcal{U}_{\mathcal{A}}$ of $\mathcal{A}$ is the image of some ground term t, i.e., $\mathcal{A}(t) = e$.

3.2.2 Definition (Semantics)

An algebra $\mathcal{A}$ and an assignment β are extended to formulas $\phi \in \mathrm{FOL}(\Sigma, \mathcal{X})$ by

$$\mathcal{A}(\beta)(\bot) := 0 \qquad \mathcal{A}(\beta)(\top) := 1$$

$$\begin{array}{rcl}
\mathcal{A}(\beta)(s \approx t) & := & 1 \text{ if } \mathcal{A}(\beta)(s) = \mathcal{A}(\beta)(t) \text{ else } 0 \\
\mathcal{A}(\beta)(P(t_1, \ldots, t_n)) & := & 1 \text{ if } (\mathcal{A}(\beta)(t_1), \ldots, \mathcal{A}(\beta)(t_n)) \in P_{\mathcal{A}} \text{ else } 0 \\
\mathcal{A}(\beta)(\neg\phi) & := & 1 - \mathcal{A}(\beta)(\phi) \\
\mathcal{A}(\beta)(\phi \land \psi) & := & \min(\{\mathcal{A}(\beta)(\phi), \mathcal{A}(\beta)(\psi)\}) \\
\mathcal{A}(\beta)(\phi \lor \psi) & := & \max(\{\mathcal{A}(\beta)(\phi), \mathcal{A}(\beta)(\psi)\}) \\
\mathcal{A}(\beta)(\phi \rightarrow \psi) & := & \max(\{(1 - \mathcal{A}(\beta)(\phi)), \mathcal{A}(\beta)(\psi)\}) \\
\mathcal{A}(\beta)(\phi \leftrightarrow \psi) & := & \text{if } \mathcal{A}(\beta)(\phi) = \mathcal{A}(\beta)(\psi) \text{ then 1 else } 0 \\
\mathcal{A}(\beta)(\exists x_S.\phi) & := & 1 \text{ if } \mathcal{A}(\beta[x \mapsto e])(\phi) = 1 \\
& & \text{for some } e \in S_{\mathcal{A}} \text{ and 0 otherwise} \\
\mathcal{A}(\beta)(\forall x_S.\phi) & := & 1 \text{ if } \mathcal{A}(\beta[x \mapsto e])(\phi) = 1 \\
& & \text{for all } e \in S_{\mathcal{A}} \text{ and 0 otherwise}
\end{array}$$

A formula ϕ is called *satisfiable by $\mathcal{A}$ under β* (or *valid in $\mathcal{A}$ under β*) if $\mathcal{A}, \beta \models \phi$; in this case, ϕ is also called *consistent*;

satisfiable by $\mathcal{A}$ if $\mathcal{A}, \beta \models \phi$ for some assignment β;

satisfiable if $\mathcal{A}, \beta \models \phi$ for some algebra $\mathcal{A}$ and some assignment β;

valid in $\mathcal{A}$, written $\mathcal{A} \models \phi$, if $\mathcal{A}, \beta \models \phi$ for any assignment β; in this case, $\mathcal{A}$ is called a *model* of ϕ;

valid, written $\models \phi$, if $\mathcal{A}, \beta \models \phi$ for any algebra $\mathcal{A}$ and any assignment β; in this case, ϕ is also called a *tautology*;

unsatisfiable if $\mathcal{A}, \beta \not\models \phi$ for any algebra $\mathcal{A}$ and any assignment β; in this case ϕ is also called *inconsistent*.

Given two formulas ϕ and ψ, ϕ *entails* ψ, or ψ is a *consequence* of ϕ, written $\phi \models \psi$, if for any algebra $\mathcal{A}$ and assignment β, if $\mathcal{A}, \beta \models \phi$ then $\mathcal{A}, \beta \models \psi$.

The formulas ϕ and ψ are called *equivalent*, written $\phi \models\!\!\!| \psi$, if $\phi \models \psi$ and $\psi \models \phi$.

Two formulas ϕ and ψ are called *equisatisfiable*, if ϕ is satisfiable iff ψ is satisfiable (not necessarily in the same models).

The notions of “entailment”, “equivalence” and “equisatisfiability” are naturally extended to sets of formulas, that are treated as conjunctions of single formulas.

Clauses are implicitly universally quantified disjunctions of literals. A clause C is satisfiable by an algebra $\mathcal{A}$ if for every assignment β there is a literal $L \in C$ with $\mathcal{A}, \beta \models L$.

Note that if $C = \{L_1, \ldots, L_k\}$ is a ground clause, i.e., every L_i is a ground literal, then $\mathcal{A} \models C$ if and only if there is a literal L_j in C so that $\mathcal{A} \models L_j$. A clause set N is satisfiable iff all clauses $C \in N$ are satisfiable by the same algebra $\mathcal{A}$. Accordingly, if N and M are two clause sets, $N \models M$ iff every model $\mathcal{A}$ of N is also a model of M.

3.3.1 Definition (Substitution (well-sorted))

A *well-sorted substitution* is a mapping $\sigma : \mathcal{X} \to T(\Sigma, \mathcal{X})$ so that

1. $\sigma(x) \neq x$ for only finitely many variables x and
2. $\text{sort}(x) = \text{sort}(\sigma(x))$ for every variable $x \in \mathcal{X}$.

The application $\sigma(x)$ of a substitution σ to a variable x is often written in postfix notation as $x\sigma$. The variable set $\text{dom}(\sigma) := \{x \in \mathcal{X} \mid x\sigma \neq x\}$ is called the *domain* of σ.

The term set $\text{codom}(\sigma) := \{x\sigma \mid x \in \text{dom}(\sigma)\}$ is called the *codomain* of σ. From the above definition it follows that $\text{dom}(\sigma)$ is finite for any substitution σ. The composition of two substitutions σ and τ is written as a juxtaposition $\sigma\tau$, i.e., $t\sigma\tau = (t\sigma)\tau$.

A substitution σ is called *idempotent* if $\sigma\sigma = \sigma$. A substitution σ is idempotent iff $\text{dom}(\sigma) \cap \text{vars}(\text{codom}(\sigma)) = \emptyset$.

Substitutions are often written as sets of pairs $\{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$ if $dom(\sigma) = \{x_1, \ldots, x_n\}$ and $x_i\sigma = t_i$ for every $i \in \{1, \ldots, n\}$.

The *modification* of a substitution σ at a variable x is defined as follows:

$$\sigma[x \mapsto t](y) = \begin{cases} t & \text{if } y = x \\ \sigma(y) & \text{otherwise} \end{cases}$$

A substitution σ is identified with its extension to formulas and defined as follows:

1. $\bot\sigma = \bot$,
2. $\top\sigma = \top$,
3. $(f(t_1, \ldots, t_n))\sigma = f(t_1\sigma, \ldots, t_n\sigma)$,
4. $(P(t_1, \ldots, t_n))\sigma = P(t_1\sigma, \ldots, t_n\sigma)$,
5. $(s \approx t)\sigma = (s\sigma \approx t\sigma)$,
6. $(\neg\phi)\sigma = \neg(\phi\sigma)$,
7. $(\phi \circ \psi)\sigma = \phi\sigma \circ \psi\sigma$ where $\circ \in \{\vee, \wedge\}$,
8. $(Qx\phi)\sigma = Qz(\phi\sigma[x \mapsto z])$ where $Q \in \{\forall, \exists\}$, z and x are of the same sort and z is a fresh variable.

The result $t\sigma$ ($\phi\sigma$) of applying a substitution σ to a term t (formula ϕ) is called an *instance* of t (ϕ).

The substitution σ is called *ground* if it maps every domain variable to a ground term, i.e., the codomain of σ consists of ground terms only.

If the application of a substitution σ to a term t (formula ϕ) produces a ground term $t\sigma$ (a variable-free formula, $\text{vars}(\phi\sigma) = \emptyset$), then $t\sigma$ ($\phi\sigma$) is called *ground instance* of t (ϕ) and σ is called *grounding* for t (ϕ). The set of ground instances of a clause set N is given by $\text{ground}(\Sigma, N) = \{C\sigma \mid C \in N, \sigma \text{ is grounding for } C\}$ is the set of *ground instances* of N.

A substitution σ is called a *variable renaming* if $\text{codom}(\sigma) \subseteq \mathcal{X}$ and for any $x, y \in \mathcal{X}$, if $x \neq y$ then $x\sigma \neq y\sigma$.

3.3.2 Lemma (Substitutions and Assignments)

Let β be an assignment of some interpretation $\mathcal{A}$ of a term t and σ a substitution. Then

$$\beta(t\sigma) = \beta[x_1 \mapsto \beta(x_1\sigma), \ldots, x_n \mapsto \beta(x_n\sigma)](t)$$

where $\mathrm{dom}(\sigma) = \{x_1, \ldots, x_n\}$.