first order language

Canonical name	FirstOrderLanguage
Date of creation	2013-03-22 12:42:46
Last modified on	2013-03-22 12:42:46
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	28
Author	CWoo (3771)
Entry type	Definition
Classification	msc 03C07
Classification	msc 03B10
Synonym	auxiliary symbol
Synonym	first-order language
Related topic	Type2
Related topic	Language
Related topic	AtomicFormula
Defines	first order language
Defines	term
Defines	formula

Let Σ be a signature. The *first order language* $\mathrm{FO}(\Sigma)$ on Σ contains the following:

1. the set $S(\Sigma)$ of *symbols* of $\mathrm{FO}(\Sigma)$, which is the disjoint union of the following sets:
   (a) Σ (the *non-logical symbols*),
   (b) a countably infinite set V of variables,
   (c) the set of logical symbols $\{\wedge, \vee, \neg, \Rightarrow, \Leftrightarrow, \forall, \exists\}$,
   (d) the singleton consisting of the equality symbol $\{=\}$, and
   (e) the set of parentheses (left and right) $\{(,)\}$;
2. the set $T(\Sigma)$ of *terms* of $\mathrm{FO}(\Sigma)$, which is built inductively from $S(\Sigma)$, as follows:
   (a) Any variable $v \in V$ is a term;
   (b) Any constant symbol in Σ is a term;
   (c) If f is an n-ary function symbol in Σ, and $t_1, ..., t_n$ are terms, then $f(t_1, ..., t_n)$ is a term.
3. the set $F(\Sigma)$ of *formulas* of $\mathrm{FO}(\Sigma)$, which is built inductively from $T(\Sigma)$, as follows:
   (a) If t_1 and t_2 are terms, then $(t_1 = t_2)$ is a formula;
   (b) If R is an n-ary relation symbol and $t_1, ..., t_n$ are terms, then $(R(t_1, ..., t_n))$ is a formula;
   (c) If φ is a formula, then so is $(\neg\varphi)$;
   (d) If φ and ψ are formulas, then so is $(\varphi \vee \psi)$;
   (e) If φ is a formula, and x is a variable, then $(\exists x(\varphi))$ is a formula.

In other words, $T(\Sigma)$ and $F(\Sigma)$ are the smallest sets, among all sets satisfying the conditions given for terms and formulas, respectively.

Formulas in 3(a) and 3(b), which do not contain any logical connectives, are called the *atomic formulas*.

For example, in the first order language of partially ordered rings, expressions such as

$$0, \quad x^2, \quad \text{and} \quad y + zx$$

are terms, while

$$(x = xy), \quad (x + y \leq yz), \quad \text{and} \quad (\exists x((x \leq 0) \vee (0 \leq x)))$$

are formulas, and the first two of which are atomic.

Remarks.

1. Generally, one omits parentheses in formulas, when there is no ambiguity. For example, a formula (φ) can be simply written φ. As such, the parentheses are also called the *auxiliary symbols*.

2. The other logical symbols are obtained in the following way :

$$\varphi \wedge \psi \overset{\text{def}}{:=} \neg(\neg\varphi \vee \neg\psi) \qquad \varphi \Rightarrow \psi \overset{\text{def}}{:=} \neg\varphi \vee \psi$$
$$\varphi \Leftrightarrow \psi \overset{\text{def}}{:=} (\varphi \Rightarrow \psi) \wedge (\psi \Rightarrow \varphi) \qquad \forall x(\varphi) \overset{\text{def}}{:=} \neg(\exists x(\neg\varphi))$$

   where φ and ψ are formulas. All logical symbols are used when building formulas.

3. In the literature, it is a common practice to write $\Sigma_{\omega\omega}$ for FO(Σ). The first subscript means that every formula in FO(Σ) contains a finite number of $\vee$'s (less than ω), while the second subscript signifies that every formula has a finite number of $\exists$'s. In general, $\Sigma_{\alpha\beta}$ denotes a language built from Σ such that, in any given formula, the number of occurrences of $\vee$ is less than α and the number of occurrences of $\exists$ is less than β. When the number of occurrences of $\vee$ (or $\exists$) in a formula is not limited, we use the symbol ∞ in place of α (or β). Clearly, if α and β are not ω, we get a language that is not first-order.

First Order Languages as Formal Languages

If the signature Σ and the set V of variables are countable, then $S(\Sigma), T(\Sigma)$, and $F(\Sigma)$ can be viewed as formal languages over a certain (finite) alphabet Γ. The set Γ should include all of the logical connectives, the equality symbol, and the parentheses, as well as the following symbols

$$R, F, V, I, \#,$$

where they are used to form words for relation, formula, and variable symbols. More precisely,

- $VI^n\#$ stands for the variable v_n,
- $RI^n\#I^m\#$ stands for the m-th relation symbol of arity n, and
- $FI^n\#I^m\#$ stands for the m-th function symbol of arity n,

where $m, n \geq 0$ are integers. The symbol # is used as a delimiter or separator. Note that the constant symbols are then words of the form $F\#I^m\#$. It can shown that $S(\Sigma), T(\Sigma)$ and $F(\Sigma)$ are context-free over Γ, and in fact unambiguous.

References

[1] W. Hodges, *A Shorter Model Theory*, Cambridge University Press, (1997).

[2] D. Marker, *Model Theory, An Introduction*, Springer, (2002).