

Leary & Kristiansen, *A Friendly Introduction to Mathematical Logic*

Danny Nygård Hansen

14th July 2022

1 • Structures and Languages

Languages: A first-order language \mathcal{L} is actually an (infinite) alphabet consisting of the usual symbols. Notice that we have a *countable* number of variables. These variables are common to all first-order languages, so we use the symbol *Vars*, which does not depend on which language we are considering, to denote the set of variables.

On the other hand we can have an arbitrary number of constant, function and relation symbols. I am not sure if the number of these symbols makes any difference to the theory. Instead of having a separate category for constants we may consider a constant as a function with arity 0. The arity of function symbols may, as in universal algebra, be formalised as a partial function $\rho: \mathcal{L} \rightarrow \mathbb{N}$, whose domain is the set of function and relation symbols.

Terms: Strings can be terms of a particular first-order language \mathcal{L} . We think of terms as picking out an element of the universe under consideration. These are obtained either as variables or as the images of functions in \mathcal{L} (or as constants if these are considered a class of their own).

Formulas: As with terms, strings can be formulas of a particular first-order language \mathcal{L} . Formulas are supposed to represent propositions about the objects in the universe. We may first construct formulas from \mathcal{L} -terms by applying relations (e.g. equality) to a list of terms. Afterwards we may apply either quantifiers or the usual connectives from propositional logic to the resulting formulas. Notice that the particular language \mathcal{L} only determines which terms are allowed and does not affect how these terms can be used to construct formulas.

A formula is said to be *atomic* if it is on the form $= t_1 t_2$ or $R t_1 \cdots t_n$ for an n -ary relation symbol R and terms t_1, \dots, t_n .

We say that a string ψ is a *subformula* of a formula φ if ψ itself is a formula, and if it is a substring of φ . By unique readability for formulas (Exercise 1.4.8) each formula φ that is not atomic can be broken up into smaller formulas in a unique way. These give rise to a (unique) construction tree for φ , and the subformulas are precisely the formulas that appear on this tree.

L&K call the symbol \forall itself a quantifier, but by *quantifier* we will mean $\forall v$ for a variable v . The *scope* of a quantifier $\forall v$ occurring in a formula $(\forall v)(\alpha)$ is the subformula α .

Free variables: We define recursively what it means for a variable v to be *free* in a formula φ . Alternatively we may simply say that v is free in φ if it is not in the scope of any quantifier $\forall v$ occurring in φ .

This also makes it easy to define what it means for an *occurrence* of a variable v in φ to be free: Namely that this occurrence does not lie in the scope of any quantifier $\forall v$.

Sentences: A *sentence* in \mathcal{L} is a formula of \mathcal{L} with no free variables.

Structures: An \mathcal{L} -*structure* \mathfrak{A} consists of a nonempty set A , the *universe* of \mathfrak{A} , along with functions $f^{\mathfrak{A}}: A^{\rho(f)} \rightarrow A$ and relations $R^{\mathfrak{A}} \in A^{\rho(R)}$ for each function symbol f and relation symbol R in \mathcal{L} . The particular language \mathcal{L} does not place any constraints on the universe A .

Compare this to universal algebra: Here we usually specify a collection \mathcal{F} of operation symbols and a type $\rho: \mathcal{F} \rightarrow \mathbb{N}$ specifying the arity of each symbol. A structure \mathfrak{A} of type ρ is then a nonempty set A along with $\rho(f)$ -ary operations $f^{\mathfrak{A}}$ on A for each $f \in \mathcal{F}$. In first-order logic we also allow relations, but otherwise the definition of a structure is basically the same.

Variable assignment functions: A (variable) *assignment function* into an \mathcal{L} -structure \mathfrak{A} is a function $\text{Vars} \rightarrow A$. This is the first step towards connecting the syntax and semantics of \mathcal{L} . Notice that regardless of the particular structure or language, it is always the variables in Vars that we need to assign values to. The codomain A however is determined by the structure \mathfrak{A} in question, which itself depends on the language \mathcal{L} (though A does not as such depend on \mathcal{L}).

Given $s: \text{Vars} \rightarrow A$, a variable x and a particular value $a \in A$, we define an *x-modification* of s denoted $s[x \mid a]$ with $s[x \mid a](x) = a$.

Term assignment function: Given $s: \text{Vars} \rightarrow A$ we define the *term assignment function* \bar{s} generated by s . This is defined by recursion on terms, and it is clearly the unique extension s' of s to all terms satisfying that

$$s'(f t_1 \cdots t_n) = f^{\mathfrak{A}}(s'(t_1), \dots, s'(t_n))$$

for all function symbols f and all terms t_1, \dots, t_n . Indeed, Lemma 1.7.6 shows that only the variables occurring in a particular term t determines the value of \bar{s} in t .

Satisfaction: We define what it means for an \mathcal{L} -structure \mathfrak{A} to *satisfy* an \mathcal{L} -formula φ with assignment $s: \text{Vars} \rightarrow A$ recursively. In the affirmative case we write $\mathfrak{A} \models_s \varphi$.

The base cases are the atomic formulas. This completes the connection between the syntax and semantics of \mathcal{L} : We have now effectively associated a truth value to each \mathcal{L} -formula, provided that we are given an interpretation of all the symbols in such a formula, namely function symbols (i.e. \mathfrak{A}) and variables (i.e. s).

Notice that the definition of satisfaction is *informal*, even if it is precise. Clearly we cannot hope to define the semantics of \mathcal{L} in terms of \mathcal{L} itself, so we must resort to a definition in natural language using terms like ‘not’, ‘or’ and ‘for all’.

Proposition 1.7.7 shows that satisfaction of a formula φ only depends on the free variables occurring in φ .

Models and validity: An \mathcal{L} -structure \mathfrak{A} is a *model* of an \mathcal{L} -formula φ , written $\mathfrak{A} \models \varphi$, if $\mathfrak{A} \models_s \varphi$ for all assignment functions s . In turn $\mathfrak{A} \models \varphi$ for all \mathcal{L} -structures \mathfrak{A} , then we say that φ is *valid* and write $\models \varphi$.

If σ is a sentence in \mathcal{L} , then $\mathfrak{A} \models \sigma$ if and only if $\mathfrak{A} \models_s \sigma$ for *any* s , in which case we say that σ is *true in* \mathfrak{A} .

Substitutions: We define recursively how to substitute a term in place of a variable, first in terms and then in formulas. L&K use the notation u_t^x for the term u with every occurrence of x replaced with the term t , but the notation $u[t/x]$ is also common. We similarly write φ_t^x for a formula φ .

If φ is a formula, t a term and x a variable, then we define when t is *substitutable* for x in φ . We do this by recursion, and the case when φ is on the form $(\forall y)(\alpha)$ is the interesting one. In this case t is substitutable for x if either of the following holds:

- (a) x is not free in φ , which includes the case when $y = x$. In this case every occurrence of x (as a variable and not in a quantifier) in φ lies in the scope of a quantifier $\forall x$. But when performing the substitution φ_t^x the relevant subformula $(\forall x)(\beta)$ is left untouched. Hence there is no danger of making a substitution in a problemating place, since we aren’t making any substitutions!
- (b) y does not occur in t , and t is substitutable for x in α . A problem could happen if the former is not satisfied, since any free occurrence of y would be bound by the quantifier, which is not desirable. The latter condition is of course necessary.

Logical implication: If Γ and Δ are sets of \mathcal{L} -formulas, then we say that Γ *logically implies* Δ , denoted $\Gamma \models \Delta$, if $\mathfrak{A} \models \Gamma$ implies $\mathfrak{A} \models \Delta$ for all \mathcal{L} -structures \mathfrak{A} .

If $\emptyset \models \varphi$, then we say that φ is *valid* and write $\models \varphi$. This agrees with the definition of validity above.

EXERCISE 1.9.3

Suppose that φ is an \mathcal{L} -formula and x is a variable. Prove that φ is valid if and only if $(\forall x)(\varphi)$ is valid. Thus, if φ has free variables x, y , and z , φ will be valid if and only if $\forall x \forall y \forall z \varphi$ is valid. The sentence $\forall x \forall y \forall z \varphi$ is called the *universal closure* of φ .

SOLUTION. Notice that $\models (\forall x)(\varphi)$ if and only if $\mathfrak{A} \models_s (\forall x)(\varphi)$ for all \mathcal{L} -structures \mathfrak{A} and variable assignment functions s into \mathfrak{A} . This is the case if and only if $\mathfrak{A} \models_{s[x|a]} \varphi$ for all \mathfrak{A}, s and elements a in the universe A . If φ is valid then this follows.

Conversely, notice that $s = s[x | s(x)]$, so every variable assignment function into \mathfrak{A} is on the form $s[x | a]$ for some s and a in A . Hence the above implies that $\mathfrak{A} \models_s \varphi$ for all \mathfrak{A} and s . Hence $\mathfrak{A} \models \varphi$ for all \mathfrak{A} , so $\models \varphi$. \square

EXERCISE 1.9.4

- (a) Assume that $\models (\varphi \rightarrow \psi)$. Show that $\varphi \models \psi$.
- (b) Suppose that φ is $x < y$ and ψ is $z < w$. Show that $\varphi \models \psi$ but $\not\models (\varphi \rightarrow \psi)$.

SOLUTION. (a) Let \mathfrak{A} be a structure such that $\mathfrak{A} \models \varphi$, and let s be any variable assignment function into \mathfrak{A} . In particular we thus have $\mathfrak{A} \models_s \varphi$. Furthermore, $\mathfrak{A} \models_s (\varphi \rightarrow \psi)$, which is the case if and only if either $\mathfrak{A} \not\models_s \varphi$ or $\mathfrak{A} \models_s \psi$. The former is impossible, so the latter holds. Since s was arbitrary we have $\mathfrak{A} \models \psi$, so $\varphi \models \psi$ as desired.

(b) We first show that $\varphi \models \psi$, so let \mathfrak{A} be a structure such that $\mathfrak{A} \models \varphi$. We must then show that $\mathfrak{A} \models \psi$. If $<^{\mathfrak{A}} = A \times A$ then this is obvious. Assume thus that there exist $a, b \in A$ such that $(a, b) \notin <^{\mathfrak{A}}$, i.e. such that $a <^{\mathfrak{A}} b$ is false. Let s be a variable assignment function into \mathfrak{A} with $s(x) = a$ and $s(y) = b$. Then $\mathfrak{A} \models_s \varphi$, so nothing has to be proved. This shows that $\varphi \models \psi$.

We next show that $\not\models (\varphi \rightarrow \psi)$. Let r be a variable assignment function into \mathfrak{A} with $r(x) = 0, r(y) = 1, r(z) = 1$, and $r(w) = 0$. Then $\mathfrak{A} \models_r \varphi$ but $\mathfrak{A} \not\models_r \psi$, so $\mathfrak{A} \not\models_r (\varphi \rightarrow \psi)$ as desired. \square

2 • Deductions

Deductions: We fix a language \mathcal{L} , a set Λ of *logical axioms*, and a set of ordered pairs (Γ, φ) called *rules of inference*. We further have a set Σ of *nonlogical*

axioms, whose content depends on the application. A *deduction from* Σ is then a finite sequence of \mathcal{L} -formulas such that each formula is either an axiom or follows from the previous formulas by the rules of inference. If there exists a deduction from Σ whose last element is φ , then this deduction is called a *deduction from* Σ of φ , and we write $\Sigma \vdash \varphi$.

We denote by Thm_Σ the set of formulas φ such that $\Sigma \vdash \varphi$. This is clearly the smallest set of formulas containing the axioms that is closed under application of the rules of inference.

Logical axioms: The logical axioms include:

- (E₁) Reflexivity of variables, i.e. $x = x$ for every variable x .
- (E₂) Substitution in function symbols: If two variables are the same, we can substitute them as arguments to any function symbol.
- (E₃) Substitution in relation symbols: Same as above, but for relation symbols.

For instance, for each n -ary function symbol f in \mathcal{L} and each collection $x_1, \dots, x_n, y_1, \dots, y_n$ of variables we get an axiom of the second kind above. Hence even if there is only one function symbol in \mathcal{L} , as long as this has positive arity there are infinitely many axioms of this kind: If f is unary, then we have an axiom

$$v_i = v_j \rightarrow f(v_i) = f(v_j)$$

for all $i, j \in \mathbb{Z}_+$. Hence we already have a potentially massive collection of axioms.

We also have axioms for quantifiers. For every variable x , term t and formula φ such that t is substitutable for x in φ we have the following:

- (Q₁) Universal instantiation: $(\forall x\varphi) \rightarrow \varphi_t^x$.
- (Q₂) Existential generalisation: $\varphi_t^x \rightarrow (\exists x\varphi)$.

Propositional consequence: Consider the restricted language \mathcal{P} only containing a set of propositional variables and the logical connectives. Formulas of \mathcal{P} are given recursively in the usual way, and the semantics of \mathcal{P} is that of standard propositional logic. In particular, a formula of \mathcal{P} is a *tautology* if every assignment of truth values to the propositional variables makes it true.

Given an \mathcal{L} -formula φ , we convert it to a \mathcal{P} -formula φ_P as follows:

- (a) Systematically replace with propositional variables every occurrence of subformulas on the form $(\forall x)(\alpha)$ that are not in the scope of another quantifier.
- (b) Systematically replace in a systematic way with propositional variables every remaining atomic subformula.

By ‘systematically’ we mean that multiple occurrences of the same subformula is replaced with the same propositional variable.

If Γ_P is a set of \mathcal{P} -formulas and φ_P is a \mathcal{P} -formula, then φ_P is a *propositional consequence* of Γ_P if φ_P is true under every truth assignment making every formula in Γ_P true.

If Γ is a *finite* set of \mathcal{L} -formulas and φ is an \mathcal{L} -formula, then φ is a *propositional consequence* of Γ if φ_P is a propositional consequence of Γ_P . I'm not sure exactly how significant the finiteness condition is. Perhaps it has to do with us having to go through every formula in Γ and convert each to a propositional formula? Perhaps the corresponding rules of inference will not be decidable.

Rules of inference: We first have the following:

(PC) If an \mathcal{L} -formula φ is a propositional consequence of a finite set Γ of \mathcal{L} -formulas, then (Γ, φ) is a rule of inference.

We also have the rules

$$\left(\{\psi \rightarrow \varphi\}, \psi \rightarrow (\forall x)(\varphi) \right) \quad \text{and} \quad \left(\{\varphi \rightarrow \psi\}, (\exists x)(\varphi) \rightarrow \psi \right)$$

of type (QR), where x is not free in ψ . This assumption is supposed to formalise that we make no assumptions about x . Intuitively, if x occurs in ψ then we may simply make the substitution ψ_y^x , where $y \neq x$ is a variable that does not occur in ψ . If we can then somehow prove a formula φ in which x is free, then this x cannot have come from ψ . Hence φ must hold whatever x is.

The second rule says that if x somehow 'disappears' when proving ψ , then it didn't matter what x was. Hence it's enough that there exists some x .

Soundness: If $\Sigma \vdash \varphi$, then $\Sigma \models \varphi$. [TODO proof]