Introduction to Logic and Model Theory

Greg Oman

University of Colorado, Colorado Springs

goman@uccs.edu

September 13, 2017

First-order languages

A **first-order language with equality** consists of a set L whose members are arranged as follows:

I *Logical symbols*

(i) Parentheses: (and).
(ii) Logical operators: $\neg$, $\vee$, $\wedge$, $\rightarrow$, and $\leftrightarrow$.
(iii) Variables: a variable $\mathbf{v}_n$ for every positive integer n.
(iv) Equality symbol: $\approx$.

II *Parameters*

(i) Quantifier symbols: $\forall$ and $\exists$.
(ii) Predicate symbols: for each positive integer n, some set (maybe empty) of symbols, called n-**place predicate symbols**.
(iii) Constant symbols: some set (possibly empty) of symbols, called **constant symbols**.
(iv) Function symbols: for each positive integer n, some set (maybe empty) of symbols, called n-**place function symbols**.

First-order languages

Example

The language of set theory (usually) consists of a single 2-place (or *binary*) predicate symbol $\in$, no constant symbols, and no function symbols.

Example

The language of (unital) ring theory consists of no predicate symbols, constant symbols $\mathbf{0}$ and $\mathbf{1}$, a two-place function symbol $+$, a two-place function symbol $\cdot$, and a unary function symbol $\mathbf{I}$ (whose interpretation in a ring is the function $I(x) := -x$).

Formulas

Our next goal is to give a rigorous definition of "formula" relative to the languages we just defined. Toward this end, let n be a positive integer, S a set, and $f\colon S^n \to S$ a function. Recall that a set $X \subseteq S$ is *closed under* f provided that for any $x_1, \ldots, x_n \in X$, also $f(x_1, \ldots, x_n) \in X$. We call n the *arity* of the function f. Suppose now that $\mathcal{F}$ is a collection of functions on S, each of finite arity (we do *not* assume that all functions are of the same arity). Then $X \subseteq S$ is *closed under the functions in* $\mathcal{F}$ provided that whenever $f \in \mathcal{F}$ has arity k and $x_1, \ldots, x_k \in X$, also $f(x_1, \ldots, x_k) \in X$. Next, suppose that U is a set, $\mathcal{F}$ is a collection of operations on U, each of finite arity, and that $B \subseteq U$. Then *the subset of* U *generated from* B *by the functions in* $\mathcal{F}$ is simply the intersection of all subsets of U containing B which are closed under the functions in $\mathcal{F}$, which we denote by $\overline{B}$. Two important properties of $\overline{B}$ are that it is closed under the functions in $\mathcal{F}$ and also satisfies the following induction principle: if $B \subseteq X \subseteq \overline{B}$ and X is closed under the functions in $\mathcal{F}$, then $X = \overline{B}$.

Formulas

Next, let us suppose that we are given a first-order language L. Let us define the set of *L-expressions* to be the set of all finite sequences of elements of the language L, which we denote by $\mathrm{seq}(L)$ (we identify the finite sequences of length one with elements of L).

Example

If L is the language of ring theory, then $(\cdot, +, \forall, \forall, \rightarrow, \mathbf{1}) \in \mathrm{seq}(L)$.

Our next goal is to distinguish those expressions which tell us something meaningful from those which don't. First, if $\alpha := (x_1, \ldots, x_n)$ and $\beta := (y_1, \ldots, y_m)$ are members of $\mathrm{seq}(L)$, then we let $\alpha\beta$ denote the concatenated sequence $(x_1, \ldots, x_n, y_1, \ldots, y_m)$.

Formulas

Definition

Suppose that $\mathbf{f}$ is an n-place function symbol, and define an operation $\varphi_{\mathbf{f}} \colon \mathrm{seq}(L)^n \to \mathrm{seq}(L)$ by $\varphi_{\mathbf{f}}(\epsilon_1, \ldots, \epsilon_n) := \mathbf{f}\epsilon_1\epsilon_2 \cdots \epsilon_n$. Now set $\mathcal{F} := \{\varphi_{\mathbf{f}} \colon \ \mathbf{f} \text{ a function symbol}\}$. Then the subset of $\mathrm{seq}(L)$ generated from the constant symbols and the variables by the functions in $\mathcal{F}$ is called the set of **terms** of a first-order language L.

Example

Let L be the language of ring theory. Then $\mathbf{0}$ is a term because it is a constant. Next, $+\mathbf{00}$ is a term (think of this as $\mathbf{0} + \mathbf{0}$), and thus $+ + \mathbf{000}$ is also a term (think of this as $(\mathbf{0} + \mathbf{0}) + \mathbf{0}$).

Formulas

Definition

An **atomic formula** is an expression of the form $Pt_1 t_2 \cdots t_n$, where P is an n-place predicate and $t_1, \ldots, t_n$ are terms.

Observe that *some* atomic formulas always exist since by definition, the two-place equality predicate $\approx$ is present in every language. Next, fix a first-order language L and define the following operations on $\mathrm{seq}(L)$:

1. $\varphi_{\neg}(\epsilon) := (\neg\epsilon)$,
2. $\varphi_{*}(\epsilon, \beta) := (\epsilon * \beta)$ for $* \in \{\vee, \wedge, \rightarrow, \leftrightarrow\}$,
3. for $n \in \mathbb{Z}^{+}$, $\varphi_{\forall_n}(\epsilon) := \forall \mathbf{v}_n \epsilon$, and
4. for $n \in \mathbb{Z}^{+}$, $\varphi_{\exists_n}(\epsilon) := \exists \mathbf{v}_n \epsilon$.

Formulas

Definition

Let L be a first-order language. Then the collection of L-**formulas** (or simply *formulas* when the language is clear) is the subset of $\mathrm{seq}(L)$ generated from the atomic formulas by the functions in groups (1)–(4) on the previous slide.

L-structures

Consider the language consisting of a single predicate symbol $<$, and let $\mathbf{x}$ and $\mathbf{y}$ be variables. Then $\forall \mathbf{x} \exists \mathbf{y} < \mathbf{xy}$ is a formula. The intended translation of this formula is, "For all x, there exists y such that $x < y$." Now, it makes no sense to ask whether the above formula is *true*. It depends on the intended interpretation of the formula inside of some structure. For example, the formula is true in the context of the reals with their usual order. On the other hand, the assertion is false if instead we consider the set $\{0, 1, 2\}$ with the usual order. The moral: in general, there is no notion of a formula being "true" or "false" in a vacuum; we need some interpretation of the parameters.

L-structures

Definition

Let L be a first-order language. An L-**structure** is a function $\mathcal{U}$ defined on a subset of L as follows:

1. $\mathcal{U}$ assigns to $\forall$ some nonempty set $|\mathcal{U}|$, called the *universe* of $\mathcal{U}$.
2. $\mathcal{U}$ assigns to the equality symbol $\approx$ the equality relation on $|\mathcal{U}|$ (this is why $\approx$ is a logical symbol and not a parameter: it is not open to interpretation).
3. $\mathcal{U}$ assigns to each n-place predicate **P** an n-ary relation $P^{\mathcal{U}}$ on $|\mathcal{U}|$.
4. $\mathcal{U}$ assigns to each constant symbol **c** an element $c^{\mathcal{U}} \in |\mathcal{U}|$.
5. $\mathcal{U}$ assigns to each n-place function symbol **f** a function $f^{\mathcal{U}} \colon |\mathcal{U}|^n \to |\mathcal{U}|$.

Satisfiability

Suppose that L is a first-order language and that $\mathcal{U}$ is an L-structure. Consider the formula $\approx \mathbf{v}_1\mathbf{v}_2$ (more readably, $\mathbf{v}_1 \approx \mathbf{v}_2$). We have no way to determine if this formula is true or false, even relative to an explicit L-structure $\mathcal{U}$ (such that $|\mathcal{U}|$ has more than one element). The issue is simply that we don't know which elements of $|\mathcal{U}|$ that $\mathbf{v}_1$ and $\mathbf{v}_2$ denote. Once we specify what values the variables assume, then we can determine the truth/falsity of any formula (relative to this assignment).

Satisfiability

Definition

Let L be a first-order language and let $\mathcal{U}$ be an L-structure. A **variable assignment** is a function $s\colon V \to |\mathcal{U}|$ (here V is the set of variables). If $s\colon V \to |\mathcal{U}|$ is a variable assignment, $\mathbf{x}$ is a variable, and $c \in |\mathcal{U}|$, then the notation $s(\mathbf{x}|c)$ denote the variable assignment which is the same as s except $\mathbf{x}$ is mapped to c.

Satisfiability

Definition

Let L be a first-order language, $\mathcal{U}$ an L-structure, and s a variable assignment. We shall define what it means for $\mathcal{U}$ to **satisfy** an L-formula φ with s (intuitively, this means that the formula is *true* relative to the variable assignment s), which we shall denote by $\models_{\mathcal{U}} \varphi[s]$.

Satisfiability

Fix a language L and an L-structure $\mathcal{U}$. Now let $s\colon V \to |\mathcal{U}|$ be a variable assignment. We begin by extending s (via recursion) to a function $\overline{s}\colon T \to |\mathcal{U}|$, where T is the set of terms of L. Begin by setting $\overline{s}(\mathbf{x}) := s(\mathbf{x})$ for a variable $\mathbf{x}$ and $\overline{s}(\mathbf{c}) = c^{\mathcal{U}}$. Now suppose that $\overline{s}(t_1), \ldots, \overline{s}(t_k)$ have been defined, and let $\mathbf{f}$ be a k-place function symbol. Then set $\overline{s}(\mathbf{f}\,t_1 \cdots t_k) := f^{\mathcal{U}}(\overline{s}(t_1), \ldots, \overline{s}(t_k))$.

Example

Consider the language L of abelian group theory; this language has $\approx$, a constant symbol $\mathbf{0}$, a two-place function symbol $+$, and a unary function symbol $\mathbf{I}$ (intented to denote the inversion map). Consider the structure with universe $\mathbb{R}$, and interpret $\mathbf{0}$ as the real number 0 and $+$ as the usual addition on the reals. If $s\colon V \to \mathbb{R}$ is a variable assignment, then the terms of L interpret as finite sums of elements of $\{0, \pm s(\mathbf{v}_1), \pm s(\mathbf{v}_2), \ldots\}$.

Satisfiability

Continuing, we now define the expression "$\models_{\mathcal{U}} \varphi[s]$" (read "$\mathcal{U}$ satisfies φ with s") for every L-formula φ. Again, we proceed by recursion as follows:

1. $\models_{\mathcal{U}} \mathbf{P} t_1 \cdots t_n[s]$ iff $(\overline{s}(t_1), \ldots, \overline{s}(t_n)) \in P^{\mathcal{U}}$ for an n-place predicate $\mathbf{P}$.
2. $\models_{\mathcal{U}} (\neg\alpha)[s]$ iff $\not\models_{\mathcal{U}} \alpha[s]$.
3. $\models_{\mathcal{U}} (\alpha \wedge \beta)[s]$ iff $\models_{\mathcal{U}} \alpha[s]$ and $\models_{\mathcal{U}} \beta[s]$.
4. $\models_{\mathcal{U}} (\alpha \vee \beta)[s]$ iff $\models_{\mathcal{U}} (\alpha)[s]$ or $\models_{\mathcal{U}} \beta[s]$.
5. $\models_{\mathcal{U}} (\alpha \to \beta)[s]$ iff either $\not\models_{\mathcal{U}} \alpha[s]$ or $\models_{\mathcal{U}} \beta[s]$.
6. $\models_{\mathcal{U}} (\alpha \leftrightarrow \beta)[s]$ iff either both $\models_{\mathcal{U}} \alpha[s]$ and $\models_{\mathcal{U}} \beta[s]$ or both $\not\models_{\mathcal{U}} \alpha[s]$ and $\not\models_{\mathcal{U}} \beta[s]$.
7. $\models_{\mathcal{U}} \exists \mathbf{x} \alpha[s]$ if and only if there is some $c \in |\mathcal{U}|$ such that $\models_{\mathcal{U}} \alpha[s(\mathbf{x}|c)]$.
8. $\models_{\mathcal{U}} \forall \mathbf{x} \alpha[s]$ if and only if $\models_{\mathcal{U}} \alpha[s(\mathbf{x}|c)]$ for *every* $c \in |\mathcal{U}|$.

Sentences

Recall from basic logic that, roughly, a variable $\mathbf{x}$ occurs **free** in a formula φ if it is not quantified.

Example

1. $\mathbf{x}$ occurs free in the formula $\mathbf{x} \approx \mathbf{x}$.
2. $\mathbf{x}$ in not free (i.e. it is **bound**) in the formula $\forall \mathbf{x}(\mathbf{x} \approx \mathbf{x})$.
3. $\mathbf{x}$ occurs free in the formula $(\forall \mathbf{x}(\mathbf{x} \approx \mathbf{x})) \vee (\mathbf{x} \approx \mathbf{x})$.

Sentences

An appealing attribute of sentences is that their satisfiability is independent of variable assignments:

Theorem

Let L be a language, $\mathcal{U}$ an L-structure, and suppose that φ is a sentence. If $s, t \colon V \to |\mathcal{U}|$ are variable assignments, then $\models_{\mathcal{U}} \varphi[s]$ if and only if $\models_{\mathcal{U}} \varphi[t]$.

If φ is a sentence such that there is some variable assignment such that $\mathcal{U}$ satisfies φ with s, then we say that $\mathcal{U}$ is a **model** of φ, and we write $\models_{\mathcal{U}} \varphi$. Suppose now that $\sum$ is a collection of L-sentences. Then we say that an L-structure $\mathcal{U}$ is a model of $\sum$ if $\mathcal{U}$ is a model of every sentence in $\sum$.

Example

Consider the language of groups, which is the language with equality, a constant symbol $\mathbf{e}$, a two-place function symbol $\times$, and a unary function symbol $\mathbf{I}$. Observe that we may express the group axioms as sentences in this language. For example, the inverse axiom is: $\forall \mathbf{x} \exists \mathbf{y}((\mathbf{x} \times \mathbf{y} \approx \mathbf{e}) \wedge (\mathbf{y} \times \mathbf{x} \approx \mathbf{e}))$.

Compactness

Theorem (Compactness Theorem)

Let $\sum$ be a collection of sentences in a language L. If every finite subset of $\sum$ has a model, then $\sum$ has a model.

This theorem is a more or less immediate consequence of Kurt Gödel's Completeness Theorem for first order logic (1930). Certainly compactness is one of the most important features of first-order logic, and has some very far-reaching consequences. For example, if G is a graph with the property that every finite subgraph of G can be colored with k colors, then the entire graph can be colored with k colors. This result is "really" a result of logic, not graph theory.

Lowenheim-Skolem Theorems

Definition

Let L be a language, and let $\mathcal{U}$ and $\mathcal{B}$ be L-structures. We say that $\mathcal{U}$ and $\mathcal{B}$ are *elementarily equivalent* provided $\mathcal{U}$ and $\mathcal{B}$ satisfy the same set of L-sentences.

Example

The following hold:

1. Isomorphic structures are always elementarily equivalent.
2. The converse fails (more on this soon, if time): $(\mathbb{Q}, +)$ is elementarily equivalent to $(\mathbb{R}, +)$ (as L-structures, where L is the language of group theory), but the two groups are not isomorphic.

Lowenheim-Skolem Theorems

Theorem (Lowenheim-Skolem Theorem)

Let L be a language of cardinality κ, and let $\sum$ be a collection of L-sentences. If $\sum$ has an infinite model, then $\sum$ has a model of every cardinality $\alpha \geq \kappa$. In particular, if one takes any infinite L-structure, then there is an elementarily equivalent L-structure of any cardinality κ or larger.

Example

Let κ be an infinite cardinal. One can prove the existence of a field of cardinality κ using just ring theory and basic set theory. Indeed, simply consider the polynomial ring $D := \mathbb{Q}[X_i : i \in \kappa]$ in κ many variables over $\mathbb{Q}$. Basic set theory yields that this ring has size κ. Thus the fraction field of D yields a field of cardinality κ. On the other hand, the axioms for a field can all be expressed in first-order logic in the language of ring theory, which is a countable language. Since $\mathbb{Q}$ is an infinite model of the field axioms, it follows by LST that there are fields of every infinite cardinality.

Elementary Submodels

Theorem (Existence of Elementary Submodels)

Let L be a countable first-order language, and let $\mathcal{U}$ be an L-structure. If $A \subseteq |\mathcal{U}|$ is infinite, then there exists a substructure $\mathcal{V}$ of $\mathcal{U}$ such that

1. *$|\mathcal{V}|$ contains A as a subset,*
2. *the cardinality of $|\mathcal{V}|$ is the same as the cardinality of A, and*
3. *$\mathcal{V} \equiv \mathcal{U}$ (that is, $\mathcal{V}$ is elementarily equivalent to $\mathcal{U}$).*

Elementary Submodels

We conclude this talk with an example of the utility of elementary submodels in ring theory. Consider the ring $V := \mathbb{Q}[[X]]$ of formal power series in the variable X with coefficients in $\mathbb{Q}$. The ground set of V is the set of all maps $f : \mathbb{N} \to \mathbb{Q}$, and thus $|V| = 2^{\aleph_0}$. It is well-known that V is a discrete valuation domain (DVR) – that is, a PID with a unique nonzero prime ideal. We can use elementary submodels to prove the existence of a countable subring of V which is also a DVR. Toward this end, augment the language by adding an additional constant $\mathbf{x}$ and interpret $\mathbf{x}$ as the variable X in the structure V. Observe that the polynomial ring $\mathbb{Q}[X]$ is a countable subring (substructure) of V. Thus there is a countable elementary substructure S of V such that $\mathbb{Q}[X] \subseteq S \subseteq \mathbb{Q}[[X]] = V$. Observe that the axioms for a commutative integral domain with identity are expressible in the language of ring theory. We conclude that S is an integral domain. Now consider that "For every a, b, either there is c such that $ac = b$ or $bc = a$" is clearly expressible in first-order logic. As this sentence is true in V, it is also true in S.

Elementary Submodels

Next, we can express "every non-unit is divisible by X" in first order logic (recall that we have a constant symbol which names X), and this sentence is true in $\mathbb{Q}[[X]]$, so it is also true in S. We claim that every nonzero nonunit of S has the form uX^n for some positive integer n and some unit u of S. This implies that S is a DVR. Toward this end, let $s \in S$ be an arbitrary nonzero nonunit. Then X divides s in S, so there is $t \in S$ such that $Xt = s$. If t is a unit of S, we're done. Otherwise, X divides t. So $X^2v = s$ for some $v \in S$. If v is a unit of S, we're done. Otherwise we continue. The process must terminate after finitely many steps, lest X^n divide s in S for every positive integer n. But then $s \in \mathbb{Q}[[X]]$ and $X^n | s$ in $\mathbb{Q}[[X]]$ for every positive integer n, and this can only happen if $s = 0$. As $s \neq 0$, the argument is concluded.

Supplementary Topics

Theorem (Los)

Suppose that $\sum$ is a collection of sentences in some language of cardinality κ. Suppose further that $\sum$ has the following properties:

1. *All models of $\sum$ are infinite, and*
2. *there is some cardinal $\lambda \geq \kappa$ such that $\sum$ is λ-categorical, that is, any two models of $\sum$ of size λ are isomorphic.*

Then any two models of $\sum$ are elementarily equivalent.

Proof.

Consider two models $\mathfrak{U}$ and $\mathfrak{V}$ of $\sum$. Since both models are infinite, we may apply Lowenheim-Skolem to obtain the existence of models $\mathfrak{U}'$ and $\mathfrak{V}'$, both of cardinality λ, such that $\mathfrak{U} \equiv \mathfrak{U}'$ and $\mathfrak{V} \equiv \mathfrak{V}'$. But now both $\mathfrak{U}'$ and $\mathfrak{V}'$ are models of $\sum$, and hence are isomorphic. Because isomorphic structures are elementarily equivalent, we see that all four structures are elementarily equivalent. □

Supplementary Topics

Next, consider the language of abelian group theory (which has a constant symbol $\mathbf{0}$, a binary function symbol $+$, and a unary function symbol $-$). One can express the notion of "nontrivial divisible torsion-free abelian group" in this language. For example, we can express "nontrivial" by "$\exists x(x \neq \mathbf{0})$". We can easily express the abelian group axioms in this language. To capture "torsion-free", we simply use a sentence φ_n, one for every n, which says $\forall x(x + x + x + x = 0 \rightarrow x = 0)$, where the $+$ occurs $n - 1$ times. Similarly, for "divisible", we simply need a sentence saying that everything has an nth-root (one sentence for every positive integer n): "$\forall x \exists y(ny = g)$".

Supplementary Topics

Next, let $\sum$ be the collection of axioms defining "nontrivial divisible torsion–free abelian group". Any model of these sentences is a non-trivial vector space over $\mathbb{Q}$, and so is infinite. Further, it follows that any two models of $\sum$ of the same uncountable cardinality are isomorphic. Thus by Los' Theorem, any two nontrivial divisible torsion-free abelian groups are elementarily equivalent.

Supplementary Topics

Now, it is not difficult to show that there is no field F such that the multplicative group of nonzero elements of F is isomorphic to $(\mathbb{Q}, +)$. It is a little more difficult, but one can show (using model theory!) that $\bigoplus_{\aleph_0} \mathbb{Q}$ (the direct sum of countably infinitely many copies of $\mathbb{Q}$ under addition) *is* the multiplicative group of nonzero elements of some field (necessarily of characteristic 2). The upshot:

Supplementary Topics

Corollary

There is no set $\sum$ of sentences in the language of group theory such that for all groups G, G is a model of $\sum$ if and only if G is the multiplicative group of nonzero elements of a field.

(in other words, we cannot "capture" being a multiplicative group of a field in first-order logic)

Proof.

Suppose some such set $\sum$ of sentences existed. Since $\bigoplus_{\aleph_0} \mathbb{Q}$ is the multiplicative group of some field, we deduce that $\bigoplus_{\aleph_0} \mathbb{Q}$ is a model of $\sum$. But recall that $\mathbb{Q}$ is a nontrivial divisible torsion-free abelian group, so $\mathbb{Q}$ is elementarily equivalent to $\bigoplus_{\aleph_0} \mathbb{Q}$, thus also a model of $\sum$. Therefore, $\mathbb{Q}$ is the multiplicative group of nonzero elements of a field, a contradiction. □

Thanks

THANK YOU!