Model Theory

Aaron Anderson

Week 3, Mathcamp 2025

1 Basics

Let's start with a puzzle. How can we tell $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$ apart using only the symbols $\{0, 1, +, \times, \leq\}$, and the basic symbols of logic?

The logic symbols we allow, for reference, are:

$=$	equals
$\wedge$	and
$\vee$	or
$\neg$	not
$\rightarrow$	implies
$\forall$	for $\forall$ll
$\exists$	there $\exists$xists

in addition to any commas and parentheses we may want to be clear how to interpret the other symbols.

In $\mathbb{N}$, it's true that $\forall x, 0 \leq x$, while this is false for the other three. Only the latter two satisfy $\forall x, x = 0 \vee \exists y, x \times y = 1$, and only $\mathbb{R}$ satisfies $\exists x, x \times x = 1 + 1$.

All 4 things we were comparing, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$, are examples of *mathematical structures*, which we have chosen to study in a particular language, $\{0, 1, +, \times, \leq\}$, to highlight some of their properties.

In the first few weeks of Mathcamp, particularly if you took a bunch of classes with names like "Introduction to Widget Theory", you've probably been bombarded with definitions of various kinds of mathematical structures, constructed as "a set with ...". Algebra has groups, rings, fields, vector spaces and modules, combinatorics has like 8 definitions of a graph, and there are orderings: posets, linear orderings, and lattices. If that wasn't enough, you can combine these, getting ordered groups, ordered rings, and ordered fields.

Each of these kinds of structure can be studied in its own appropriate language, where we can compare structures by using formal *sentences*, like we wrote above to compare $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$, written only in symbols from the language, plus our logic symbols. This abstract approach lets us study all of these kinds of structure at once, from a birds-eye view. By writing down the entire axiom system of a kind of structure in formal sentences, we can study and compare those too. We will see how certain structures, like "dense linear orders", vector spaces, or the $\{0, 1, +, \times, \leq\}$ structure on $\mathbb{R}$, can be understood very well using this symbolic approach, while other familiar structures, like $\mathbb{N}$ in the language $\{0, 1, +, \times, \leq\}$, are the subject of famous open problems, and will *never* be fully understood.

1.1 Languages

Let's go over what kinds of languages we can work with:

Definition 1.1. A language L consists of these three collections of symbols:

- A set of constant symbols $\mathcal{C}$
- A set of function symbols $\mathcal{F}$, and a positive integer n_f for each symbol $f \in \mathcal{F}$
- A set of relation symbols $\mathcal{R}$, and a positive integer n_R for each symbol $R \in \mathcal{R}$.

The number n_f assigned to a function symbol f indicates how many variables that function takes as inputs, and the number n_R for a relation symbol R works the same way. These are the basic symbols that we'll use to build axioms, but we're also going to be allowed to use variable symbols, the "=" relation, and some logical symbols that I'll get to in a moment. Here's an example that uses each kind of symbol:

Example 1. In the language $\mathcal{L} := \{0, 1, +, \times, \leq\}$, 0 and 1 are constant symbols, + and $\times$ are binary function symbols (meaning $n_+ = n_\times = 2$) and $\leq$ is a binary relation symbol.

1.2 Structures

Now that we have all of these symbols, it's time to give them meaning. A symbol such as + can be interpreted a bunch of different ways, as addition of natural numbers, integers, rationals, reals, complexes, integers mod n, what have you. A structure will give us the context behind these symbols.

Definition 1.2. An $\mathcal{L}$-structure $\mathcal{M}$ consists of the following things:

- A set M, called the *universe* of $\mathcal{M}$
- For each constant symbol $c \in \mathcal{C}$, an element $c^{\mathcal{M}} \in M$
- For each function symbol $f \in \mathcal{F}$, a function $f^{\mathcal{M}} : M^{n_f} \to M$
- For each relation symbol $R \in \mathcal{R}$, a relation $R^{\mathcal{M}} \subseteq M^{n_R}$.

These specific constants, functions, and relations $c^{\mathcal{M}}$, $f^{\mathcal{M}}$, and $R^{\mathcal{M}}$ are called the *interpretations* of c, f, and R.

Example 2. Here's an example using the language $\mathcal{L} = \{0, 1, +, \times, \leq\}$ from earlier: $\mathbb{N}$.

What do we need to define to turn this into a structure?

We can build an $\mathcal{L}$-structure where $M = \mathbb{N}, 0^{\mathbb{N}} = 0, 1^{\mathbb{N}} = 1$, (the constants are elements of the universe), $+^{\mathbb{N}}$ and $\times^{\mathbb{N}}$ are the usual addition and multiplication operations (binary functions), and $\leq^{\mathbb{N}}$ is the usual order (a function that takes in two inputs from $\mathbb{N}$ and outputs true or false).

1.3 Terms and Formulas

Now that we have these basic symbols, let's start putting them together. The constant symbols and variable symbols are supposed to represent elements of our structure, but we can build more complicated expressions to represent elements of our structure, called *terms*.

Definition 1.3. If $\mathcal{L}$ is a language, the set of all $\mathcal{L}$-terms (or just "terms" for short) is the smallest set of strings of symbols such that:

- Constant symbols and variable symbols are terms.
- If $f \in \mathcal{F}$ is a function symbol, and $t_1, \ldots, t_{n_f}$ are terms, then $f(t_1, \ldots, t_{n_f})$ is also a term.

Going back to our example language, we can now construct some terms like $0, x, x + 1$, and $x \times x + y \times y$, but what do these terms actually mean? If $\mathcal{M}$ is an $\mathcal{L}$-structure, then we can give these terms interpretations in $\mathcal{M}$. We already know how to interpret individual constant and function symbols, and terms are built out of constant and function symbols, so there's only one good way to do this: if a term consists of a single constant or function symbol, we already know how to interpret it, great. If our term is $f(t_1, \ldots, t_{n_f})$, then we interpret it as $f^{\mathcal{M}}\left(t_1^{\mathcal{M}}, \ldots, t_{n_f}^{\mathcal{M}}\right)$. This'll end up being a function $M^k \to M$ for some k.

Terms are great, but they output more elements of our structure, and in order to talk about axioms, we need something that outputs "true" or "false." We call these *formulas*:

Definition 1.4. A *formula* is a string formed using the following kinds of step:

- If t_1, t_2 are terms, then $t_1 = t_2$ is a formula
- If $R \in \mathcal{R}$ is a relation symbol and $t_1, \ldots, t_{n_R}$ are terms, then $R(t_1, \ldots, t_{n_R})$ is a formula.
- If ϕ, ψ are formulas, then so are
  - $\neg\phi$ (not ϕ)
  - $\phi \vee \psi$ (ϕ or ψ)
  - $\phi \wedge \psi$ (ϕ and ψ)
  - $\phi \to \psi$ (ϕ implies ψ/if ϕ, then ψ).
- If x is a variable and ϕ is a formula, then so are
  - $\exists x\phi$ (there ∃xists x such that ϕ)
  - $\forall x\phi$ (for ∀ll x, ϕ).

  (The symbols $\exists, \forall$ are called *quantifiers*.)

With formulas, we can compare our terms, with examples like $x \times x + y \times y = 1$ or $0 \leq x$. The interpretations of these end up as functions $M^k \to \{\texttt{true}, \texttt{false}\}$, which we often prefer to think of as subsets of M^k. (How?)

Often we want to get just a single "true" or "false" out of a formula, and to do that, we need to assign values to all of the variables. There's notation for this. If for instance, $\phi(x_1, \ldots, x_n)$ is an atomic formula that uses n variables, $x_1, \ldots, x_n$, and $a_1, \ldots, a_n \in M$, then we can evaluate $\phi(x_1, \ldots, x_n)$ at $a_1, \ldots, a_n$. If it turns out to be true, then we write $\mathcal{M} \models \phi(a_1, \ldots, a_n)$, otherwise, we write $\mathcal{M} \not\models \phi(a_1, \ldots, a_n)$.

Formulas are built recursively, one step at a time, and to evaluate whether they are true or false, we retrace that construction, one step at a time. Let's see an example in the language $\{0, 1, +, \times, \leq\}$, which happens to be true in $\mathbb{R}$:

$$\mathbb{R} \models \forall x, \left(0 \leq x \to \exists y, (x = y^2)\right).$$

Unpacking this by one step, this means that

$$\text{for all } a \in \mathbb{R}, \mathbb{R} \vDash 0 \leq a \to \exists y, (x = y^2)$$
$$\text{for all } a \in \mathbb{R}, \text{if } \mathbb{R} \vDash 0 \leq a, \text{ then } \mathbb{R} \vDash \exists y, (a = y^2)$$
$$\text{for all } a \in \mathbb{R}, \text{if } \mathbb{R} \vDash 0 \leq a, \text{ then there exists } b \in \mathbb{R} \text{ such that } \mathbb{R} \vDash a = b^2$$
$$\text{for all } a \in \mathbb{R}, \text{if } 0 \leq a, \text{ then there exists } b \in \mathbb{R} \text{ such that } a = b^2.$$

The interpretation of a formula in a given structure will again be a boolean-valued function, or a subset of M^k for some k. These subsets are called the *definable* sets. But how can we quickly look through and find out what k is?

In a formula such as $\forall x, y \leq x \times y$, we label a variable which under a quantifier, such as x here, a bound variable, and the other variables, in this case y, free. The free variables should be thought of as the actual inputs to the formula, while the bound variables are just things we use to calculate whether the formula is true. When we use (usually Greek letter names) to refer to formulas, we'll put the free variables in parentheses, like this: $\phi(x, y)$ will be a function where x and y are the free variables.

For example, I might refer to $\forall x, y \leq x \times y$ as $\psi(y)$, because y is the only free variable. If I then want to interpret it in a particular structure M, I need to first pick a value for y, and then I can calculate a truth value, although calculating it will involve thinking about possible values for x. If $a \in M$, and when I set y to be a, this turns out to be true, I say $\mathcal{M} \vDash \psi(a)$, otherwise, I say $\mathcal{M} \nvDash \psi(a)$.

For reference, here are the formal rules for interpreting the logical symbols, where $a_1, \ldots, a_k \in M$:

- If ϕ is $t_1 = t_2$ where t_1, t_2 are terms, then $\mathcal{M} \vDash \phi(a_1, \ldots, a_k)$ means $t_1^{\mathcal{M}}(a_1, \ldots, a_k) = t_2^{\mathcal{M}}(a_1, \ldots, a_k)$.
- If ϕ is $R(t_1, \ldots, t_{n_R})$, a relation symbol applied to terms, then $\mathcal{M} \vDash \phi(a_1, \ldots, a_k)$ means $R^{\mathcal{M}}(a_1, \ldots, a_{n_R})$.
- If ϕ, ψ have free variables in $x_1, \ldots, x_k$, then we say that
  - $\mathcal{M} \vDash \neg\phi(a_1, \ldots, a_k)$ when $\mathcal{M} \nvDash \phi(a_1, \ldots, a_k)$.
  - $\mathcal{M} \vDash (\phi \vee \psi)(a_1, \ldots, a_k)$ when $\mathcal{M} \vDash \phi(a_1, \ldots, a_k)$ or $\mathcal{M} \vDash \psi(a_1, \ldots, a_k)$.
  - $\mathcal{M} \vDash (\phi \wedge \psi)(a_1, \ldots, a_k)$ when $\mathcal{M} \vDash \phi(a_1, \ldots, a_k)$ and $\mathcal{M} \vDash \psi(a_1, \ldots, a_k)$.
  - $\mathcal{M} \vDash (\phi \to \psi)(a_1, \ldots, a_k)$ when $\mathcal{M} \vDash \psi(a_1, \ldots, a_k) \vee \neg\phi(a_1, \ldots, a_k)$.
- If ϕ has free variables $x, x_1, \ldots, x_k$, then we say that
  - $\mathcal{M} \vDash \exists x \phi(x, x_1, \ldots, x_k)$ when there exists some $a \in M$ such that $\mathcal{M} \vDash \phi(a, a_1, \ldots, a_k)$.
  - $\mathcal{M} \vDash \forall x \phi(x, x_1, \ldots, x_k)$ when for all $a \in M$, $\mathcal{M} \vDash \phi(a, a_1, \ldots, a_k)$.

A formula with no free variables is a *sentence*, and these are what our axioms and theorems will be. In particular, because a sentence ϕ has no free variables, we can ask whether $\mathcal{M} \vDash \phi$ without specifying any more information.

Homework

The following homework problems are divided into sections depending on which prerequisites they have. The Arithmetic/Number Theory section and Order sections will generally have no difficult prerequisites. For the flow of the class, I strongly recommend attempting some problems from those sections.

The other sections have prerequisites and are optional, but still recommended if you have the prerequisites.

For several homework problems today, we will want the following definition:

Definition 1.5. Let $\mathcal{M}$ be an $\mathcal{L}$-structure.
A subset $A \subseteq \mathcal{M}^k$ is called $\emptyset$-*definable* when there is some $\mathcal{L}$-formula $\phi(x_1, \dots, x_k)$ (with k free variables) such that for all $a_1, \dots, a_k \in M$, $\mathcal{M} \vDash \phi(a_1, \dots, a_k)$ if and only if $(a_1, \dots, a_k) \in A$.

1.4 Arithmetic/Number Theory

Let $\mathcal{L}_{\mathsf{nt}} = \{0, 1, +, \times, \leq\}$.

Problem 1.6. Prove that the following sets are $\emptyset$-definable by finding formulas that define them:

- The set of even numbers
- The set of pairs (x, y) such that y is a multiple of x
- The set of prime numbers
- The set of twin primes

(Note: $\mathbb{N}^1$ is basically the same thing as $\mathbb{N}$.)

Problem 1.7. Find sentences in $\mathcal{L}$ that, if interpreted in $\mathbb{N}$, describe

- The twin primes conjecture
- The Goldbach conjecture

Feel free to look up what these conjectures mean, the challenge is to write the statements formally. You can use abbreviations for the formulas from the previous problem.

In the following problems, let $\mathcal{L}_{\mathsf{alg}} = \{0, 1, +, \times\}$. This is just $\mathcal{L}_{\mathsf{nt}}$ without $\leq$. Interpret $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ as $\mathcal{L}_{\mathsf{alg}}$-structures in this language using the usual definitions of $0, 1, +, *$.

Problem 1.8. Show that the set $\{(x, y) \in \mathbb{N}^2 : x \leq y\}$ is still $\emptyset$-definable in the structure $\mathbb{N}$ in the language $\mathcal{L}_{\mathsf{alg}}$.

Problem 1.9. Show that the set $\{(x, y) \in \mathbb{R}^2 : x \leq y\}$ is $\emptyset$-definable in the structure $\mathbb{R}$ in the language $\mathcal{L}_{\mathsf{alg}}$.

Problem 1.10. Show that the sets $\{(x, y) \in \mathbb{Z}^2 : x \leq y\}$, $\{(x, y) \in \mathbb{Q}^2 : x \leq y\}$ are *ALSO* $\emptyset$-definable in $\mathbb{Z}$ and $\mathbb{Q}$ respectively in the language $\mathcal{L}_{\mathsf{alg}}$. For this, you may need to look up "Lagrange's 4-Square Theorem".

1.5 Order

Problem 1.11. Explain how a poset can be thought of as a structure in the language $\mathcal{L} = \{\leq\}$ with one binary relation. Find a sentence ϕ such that for any $\mathcal{L}$-structure $\mathcal{M}$, $\mathcal{M} \vDash \phi$ if and only if $\mathcal{M}$ is a poset. Does such a sentence exist for linear orderings?

1.6 Graphs

Let $\mathcal{L}_{\mathsf{graph}} = \{E\}$ be the language with one binary relation.

Problem 1.12. Explain how a graph can be thought of as an $\mathcal{L}$-structure. Find a sentence ϕ such that for any $\mathcal{L}$-structure $\mathcal{M}$, $\mathcal{M} \vDash \phi$ if and only if M is a (simple, undirected) graph.

1.7 Groups

Let $\mathcal{L}_{\mathsf{grp}} = \{1, \times, ^{-1}\}$ be the language of multiplicative groups, where 1 is a constant, $\times$ is a binary function, and $^{-1}$ is a unary function.

Problem 1.13. Describe a sentence ϕ (maybe a very long one!) such that if $\mathcal{G}$ is a structure in this language, then $\mathcal{G} \vDash \phi$ if and only if G is a group, where $1^{\mathcal{G}}$ is the identity, $\times^{\mathcal{G}}$ is the group operation, and $(^{-1})^{\mathcal{G}}$ is the inverse function.

1.8 Rings

Let $\mathcal{L}_{\mathsf{alg}} = \{0, 1, +, \times\}$, as in the arithmetic section.

Problem 1.14. If R is a ring, interpreted as a $\mathcal{L}_{\mathsf{alg}}$-structure in the sensible way, show that $\{(a, b, c) \in R^3 : a - b = c\}$ is $\emptyset$-definable.

(This is what it means for the subtraction *function*, sending (x, y) to $x - y$, to be $\emptyset$-definable. When you have a definable function, you can basically pretend it's in the language already as a function symbol. To see what I mean, try using this idea to solve the next problem.)

Problem 1.15. Describe a sentence ϕ (maybe a very long one!) such that if $\mathcal{R}$ is a structure in this language, then $\mathcal{R} \vDash \phi$ if and only if R is a ring, interpreted in the sensible way.

2 Theories

Finally we can capture the idea of an axiom system:

Definition 2.1. An $\mathcal{L}$-theory is a set of $\mathcal{L}$-sentences.
If T is a $\mathcal{L}$-theory, and $\mathcal{M}$ is a $\mathcal{L}$-structure, then we say that $\mathcal{M} \vDash T$ when $\mathcal{M} \vDash \phi$ for EVERY $\phi \in T$. When this happens, we also say that $\mathcal{M}$ is a model of T.

Finally, a justification of the word "model" in "Model Theory"!

Each of the classes of objects we talked about at the beginning of class can, if we're careful, be defined as "the models of T" for some well-chosen language $\mathcal{L}$ and some well-chosen $\mathcal{L}$-theory T. There is, for instance, a theory of graphs and a theory of groups. In the homework, we will look at how to write down theories for a few different kinds of objects.

2.1 Proofs and Completeness

Now that we have a formal model of an axiom system, it's time to look at what happens when we actually use these axioms to prove theorems. There are many definitions of what it means for a theory (again, a set of sentences, which we think of as axioms) to "prove" a theorem, but all the *good* definitions that I'm aware of turn out to be equivalent, so we're only going to take a quick tour through proof theory on our way to how model theory really deals with proofs and implication.

What actually is a mathematical proof? In most contexts, the answer involves convincing text, written mostly in a natural, human language, but let's go fully abstract and symbolic here. How can we write a proof totally formally in first-order logic?

There are, as I said, a few good answers, but ultimately, a formal proof is a list of first-order formulas, where each formula is derived somehow from the allowed axioms and the sentences before it.

Definition 2.2. Let ϕ be a formula, and let T be a theory. Then a *formal proof* of ϕ from T is a list $\chi_1, \dots, \chi_n$ of formulas, such that for each $1 \leq i \leq n$, χ_i satisfies one of the following conditions:

- $\chi_i \in T$ (you're allowed to assume anything from T)
- χ_i is a logical axiom (these are some formulas that are true in every $\mathcal{L}$-structure)
- There are some $j, k < i$ such that χ_k is $\chi_j \to \chi_i$. (This rule is called *modus ponens*.)
- There is some $j < i$ and some variable x such that $\chi_i = \forall x, \chi_j$.

and $\phi \in \{\chi_1, \dots, \chi_n\}$.

Definition 2.3. If ϕ is a formula, and T a theory, we say that "T proves ϕ" ($T \vdash \phi$) when there is a formal proof of ϕ from T.

This definition captures the idea of actual mathematical proofs, and what it means for a statement to follow from a set of axioms. But these proofs aren't very useful if they don't actually tell us whether a statement is true for actual models! Luckily, these proofs are useful in just that way. Any statement that follows from the axioms is actually true in every model.

Theorem 2.4 (Soundness)**.** *If $T \vdash \phi$, then for every model $\mathcal{M} \models T$, then $\mathcal{M} \models \phi$.*

Proof. If $T \vdash \phi$, then we can write down a proof for ϕ, call it $\chi_1, \ldots, \chi_n$, and $\phi = \chi_j$ for some j. Fix a model $\mathcal{M}$, and we can show by induction and casework that $\mathcal{M} \models \chi_i$ for each i. That's clearly true if $\chi_i \in T$, or if χ_i is a logical axiom (ok, we actually should check that the logical axioms are true in every $\mathcal{L}$-structure, but they are.)
Then we have two inductive steps. If $\mathcal{M} \models \chi_j$ and $\mathcal{M} \models \chi_j \to \chi_i$, then $\mathcal{M} \models \chi_i$ (Exercise if you don't believe me!)
Lastly, if $\mathcal{M} \models \phi$, then $\mathcal{M} \models \forall x, \phi$. This is basically just how we defined $\mathcal{M} \models \phi$ when ϕ has free variables. □

Now that we've established that this idea of proof actually describes some real behavior of models, we can use it to describe basic properties of theories:

Definition 2.5.
- A theory T is *inconsistent* if there is some sentence ϕ such that $T \vdash \phi$ and $T \vdash \neg\phi$. (Otherwise it's consistent.)
- A theory T is *complete* if it is consistent and for every ϕ, either $T \vdash \phi$ or $T \vdash \neg\phi$ (but necessarily not both.)

Inconsistent theories are bad - their consequences are contradictory, and once we get that one contradiction, we can derive whatever we want from it. That is, if T is inconsistent, then for every ϕ, $T \vdash \phi$. (To see this, it's enough to check that $\{\phi, \neg\phi\} \vdash \psi$.) Also, no model actually models contradictory sentences, so by Soundness, we see that T cannot have any models at all.

Meanwhile, complete theories are great! They tell us everything we might possibly want to know about their models - well, everything that first-order sentences describe. For an example of a complete theory, pick your favorite structure $\mathcal{M}$. We define *the theory of* $\mathcal{M}$, $\mathrm{Th}(\mathcal{M})$, to be the set of all ϕ such that $\mathcal{M} \models \phi$. This will be a complete theory, and in general, if T is a complete theory, then for each model $\mathcal{M} \models T$, $\mathrm{Th}(\mathcal{M})$ will be exactly the *consequences* of T. That is, $\{\phi : T \vdash \phi\}$. (Soundness ensures that $\mathcal{M} \models \phi$ for every such ϕ, and any more would give a contradiction.)

Now let's introduce a theorem that's way more powerful than it looks at first.

Theorem 2.6. *A theory T is consistent if and only if every finite subtheory $T_0 \subseteq T$ is consistent.*

Proof. First - does anyone have any suggestions for how to prove this?
It is easier to prove the contrapositive - if a subtheory $T_0 \subseteq T$ is inconsistent, then T is definitely inconsistent, because $T_0 \vdash \phi$ implies $T \vdash \phi$.
Why though do we know that if T is inconsistent, then there is some *finite* T_0 that is also inconsistent? The fundamental idea is that *proofs are finite.* If $T \vdash \phi$ and $T \vdash \neg\phi$, then we can write down a finite list T_0 of just the axioms we need in those two proofs. □

This is useful enough for describing how $\vdash$ works, but when we combine it with the following theorem, it becomes massively useful for describing models.

Theorem 2.7 (Completeness)**.** *A theory T is consistent if and only if it has a model.*

One direction of this is just Soundness, but the other direction is much deeper. Essentially this tells us that the only thing stopping a theory from having a model is if it actually has a contradiction. The proof is too technical for me to want to lecture it, but we'll see another proof later this week that gets the basic idea across. If you're interested, I can have some optional homework problems that walk you through the completeness proof.

The completeness theorem can be phrased in another remarkable way: that $T \vdash \phi$ if and only if $T \models \phi$, where $T \models \phi$ means that if $\mathcal{M} \models T$, then $\mathcal{M} \models \phi$. To prove this version, observe that $T \vdash \phi$ if and only if $T \cup \{\neg\phi\}$ is inconsistent, and $T \models \phi$ if and only if there is no model of $T \cup \{\neg\phi\}$.

This brings us to perhaps my favorite theorem ever:

Theorem 2.8 (Compactness)**.** *A theory T has a model if and only if each finite subtheory T_0 has a model.*

This follows pretty quickly from what we've said earlier about the finiteness of proofs and the completeness theorem. There's also an awesome proof using ultrafilters. Ask me about it at TAU - it could be optional bonus homework (Exercises 2.5.18 through 2.5.20 in Marker's Model Theory). A word of warning about the proof - the completeness and compactness theorems both need a version of the AXIOM OF CHOICE to prove (specifically, a magic wand called the ultrafilter lemma).

Compactness can do several tricks, but in this class, we'll show off its usefulness for constructions. If you want to build a strange structure, it can be useful to design a custom theory (often in a new language with extra constants) and then show that each of its finite subtheories has a model. Then you get a model of your custom theory, which hopefully has the properties you want.

Homework

I strongly recommend attempting the General and Arithmetic/Number Theory section.

The other sections have prerequisites and are optional, but I still recommend trying some problems from those sections if you have the prerequisites.

2.2 General

Problem 2.9. Given a theory T, construct a theory T_∞ whose models are precisely the infinite models of T.

Problem 2.10. Show that if a theory T has arbitrarily large finite models, it has an infinite model.

Problem 2.11. Show that if there is a theory T_{fin} whose models are precisely the finite models of T, then there is some $n \in \mathbb{N}$ such that all finite models of T have size at most n.

Problem 2.12. Let T and T' be $\mathcal{L}$-theories such that every $\mathcal{L}$-structure models either T or T', but not both. Show that there is a finite theory T_0 such that $\mathcal{M} \models T_0$ if and only if $\mathcal{M} \models T$. We call such a theory *finitely axiomatizable.*

(In fact, we can choose $T_0 \subseteq T$.)

2.3 Arithmetic/Number Theory

Problem 2.13. Recall the language $\mathcal{L}_{\mathsf{alg}} = \{0, 1, +, \times, \leq\}$.

Let T be your favorite $\mathcal{L}_{\mathsf{alg}}$-theory such that $\mathbb{N} \vDash T$.

We call a model $\mathcal{M} \vDash T$ *Archimedean* if for every $m \in M$, there is some $n \in \mathbb{N}$ such that

$$\mathcal{M} \vDash m \leq \underbrace{1 + 1 + \cdots + 1}_{n \text{ times}}.$$

Show that there is a non-Archimedean model of T.

(Hint: Add a constant c to $\mathcal{L}_{\mathsf{alg}}$, and apply compactness to a theory in that language.)

2.4 Linear Algebra

Let K be a field. The language we will use for K-vector spaces consists of the symbols $0, +, -$ of additive groups, together with a set of unary function symbols $\{c\cdot : c \in K\}$ representing scalar multiplication.

Problem 2.14. Write axioms in this language that give a theory $T_{K-\mathsf{vs}}$ of K-vector spaces. (You may assume that T_{grp} is the theory of additive groups, and build from there.)

2.5 Groups

Problem 2.15. Show that there is no theory whose models are precisely the finite groups.

Problem 2.16. A group G is called a *torsion* group when for every $g \in G$, there is some natural number n such that $g^n = e$. Show that there is no theory T in the language $\mathcal{L} = \{e, *, ^{-1}\}$ such that for every $\mathcal{L}$-structure $\mathcal{M}$, $\mathcal{M} \models T$ if and only if $\mathcal{M}$ is a torsion group.

2.6 Graphs

Definition 2.17. Given a graph with vertex set V, define the *edge relation* E of the graph by

$$E(v, w) \iff \text{there is an edge from } v \text{ to } w.$$

Let $\mathcal{L} = \{E\}$ consist of single binary relation.

Problem 2.18. Convince yourself that the models of the theory

$$T_{\mathsf{graph}} = \left\{ \begin{array}{l} \forall x, \neg E(x, x), \\ \forall x, \forall y, E(x, y) \to E(y, x) \end{array} \right\}$$

are exactly the edge relations of simple, undirected graphs.

Recall from Problem 2.12 that a theory T is *finitely axiomatizable* if there is a finite theory T_0 such that $M \vDash T_0$ if and only if $M \vDash T$.

Problem 2.19. Show that there is a theory T_{bp} such that for any $\mathcal{L}$-structure $\mathcal{M}$, $\mathcal{M} \models T_{\mathsf{bp}}$ if and only if $\mathcal{M}$ is a bipartite (2-colorable) graph, but that this theory is not finitely axiomatizable.

Problem 2.20. Show that there is a theory T_{pl} such that for any *finite* $\mathcal{L}$-structure $\mathcal{M}$, $\mathcal{M} \vDash T_{\mathsf{pl}}$ if and only if $\mathcal{M}$ is a planar graph. (You may find it useful to look up Kuratowski's Theorem, but it is not necessary.)

Problem 2.21. Show that there is no $\mathcal{L}$-theory T such that for any $\mathcal{L}$-structure $\mathcal{M}$, $\mathcal{M} \models \phi$ if and only if $\mathcal{M}$ is a connected graph.

Problem 2.22. Let $k \in \mathbb{N}$. Show that a graph is k-colorable if and only if all of its finite subgraphs are k-colorable.

(Hint: Add k unary relation symbols to the language.)

2.7 Fields

Problem 2.23. A field K is *algebraically closed* if every nonconstant polynomial with coefficients in K has a zero in K. For example, $\mathbb{C}$ is algebraically closed.

Describe a theory T such that for any $\mathcal{L}$-structure $\mathcal{M}$, $\mathcal{M} \models T$ if and only if $\mathcal{M}$ is an algebraically closed field.

Problem 2.24. The *characteristic* of a field is either the smallest positive integer n such that in that field, $\underbrace{1+1+\cdots+1}_{n \text{ times}} = 0$, or 0 if there is no such number. (For instance, $\mathbb{R}$ has characteristic 0, while $\mathbb{Z}/5\mathbb{Z}$ has characteristic 5.)

Describe a theory T such that for any $\mathcal{L}$-structure $\mathcal{M}$, $\mathcal{M} \models T$ if and only if $\mathcal{M}$ is a field of characteristic 0.

Problem 2.25. Let $\mathcal{L} = \{0, 1, +, *\}$ and let ϕ be an $\mathcal{L}$-sentence. Show that if there are fields of arbitrarily high characteristic that model ϕ, then there is a field of characteristic 0 that models ϕ.

3 Comparing Structures

Let's go back to the beginning of this class. How do we compare two structures to see if they act similar? Also, if we have a theory, how similar do we know its models are?

Definition 3.1. If $\mathcal{M}, \mathcal{N}$ are $\mathcal{L}$-structures, say they are *elementarily equivalent*, and write $\mathcal{M} \equiv \mathcal{N}$, when for every $\mathcal{L}$-sentence ϕ,

$$\mathcal{M} \models \phi \iff \mathcal{N} \models \phi.$$

Informally, structures are elementarily equivalent when no sentence can tell them apart.

Our puzzle at the beginning of the first day told us that $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ are NOT elementarily equivalent in the language $\{0, 1, +, \times, \leq\}$ That said, if we lose multiplication and just look at the language $\{0, +, \leq\}$, then $\mathbb{Q}$ and $\mathbb{R}$ actually *are* elementarily equivalent. We will be able to prove this at least in the language $\{\leq\}$ later in the week.

The most informative kind of consistent theory is a *complete* theory, which is a theory T such that for every sentence ϕ, either $T \models \phi$ or $T \models \neg\phi$. This means that all models of T are elementarily equivalent: If $M, N \models T$, then for every sentence ϕ, either $T \models \phi$, in which case $M, N \models \phi$, or $T \models \neg\phi$, in which case $M, N \models \neg\phi$. Either way, $M \models \phi \iff N \models \phi$.

A stronger way of saying two structures are the same is to say they're isomorphic:

Definition 3.2. An *isomorphism* between two $\mathcal{L}$-structures $\mathcal{M}, \mathcal{N}$ is a function $g : M \to N$ such that

- g is a bijection
- If c is a constant symbol, then $g(c^{\mathcal{M}}) = c^{\mathcal{N}}$.
- If f is a function symbol, and $m_1, \ldots, m_{n_f} \in M$, then

$$f(g(m_1), \ldots, g(m_{n_f})) = g(f(m_1, \ldots, m_{n_f}))$$

- If R is a relation symbol, and $m_1, \ldots, m_{n_f} \in M$, then

$$\mathcal{N} \models R(g(m_1), \ldots, g(m_{n_f})) \iff \mathcal{M} \models R(m_1, \ldots, m_{n_f}).$$

We call $\mathcal{M}$ and $\mathcal{N}$ *isomorphic* if there is some isomorphism between them.

From this, we can show by induction on the length of formulas that ANY formula is preserved by an isomorphism:

Lemma 3.3. *If $g : M \to N$ is an isomorphism, $\phi(x_1, \ldots, x_k)$ is an $\mathcal{L}$-formula, and $m_1, \ldots, m_k \in M$, then*

$$\mathcal{N} \models \phi(g(m_1), \ldots, g(m_{n_f})) \iff \mathcal{M} \models \phi(m_1, \ldots, m_{n_f}).$$

From this, we see that if two structures are isomorphic, they are also elementarily equivalent.

Example 3. In the language $\{\leq\}$ of orders, any increasing bijection between linear orders is an isomorphism, because it preserves the only symbol, $\leq$.

We know that two structures aren't isomorphic if they aren't elementarily equivalent, but there's another easy way to check. If two structures are isomorphic, there is a bijection between them, so they must have the same size (cardinality).

3.1 Up and Down

Given a theory T, assuming it has models, how big can those models be?

We've seen in the homework yesterday that if T has arbitrarily large finite models, then it has an infinite model. (We will see a similar proof today.) The real interesting stuff happens when we try to measure infinite models.

We can now state an impressive theorem, which we will work towards proving. In this whole section, until the homework, we will assume that $\mathcal{L}$ is a countable theory, meaning that the set of symbols is countable.

Theorem 3.4 (Löwenheim-Skolem)**.** *If T is a $\mathcal{L}$-theory with an infinite model, and S is an infinite set then T has a model $\mathcal{M} \vDash T$ with $|M| = |S|$.*

The full statement is a bit more involved than this, and we can prove the stronger versions in optional homework. For now, let us focus on this: if a theory has ANY infinite model, then it has a model of EVERY infinite size. This means in particular that there is no theory with infinite models, all of which are isomorphic, because then they would all have to have the same size.

We will prove this theorem in two parts, based on growing or shrinking models. This always makes me think of Alice in Wonderland, where Alice eats little snacks that make her grow and shrink to paradoxical sizes.

Theorem 3.5 (Upward Löwenheim-Skolem)**.** *If T has an infinite model, and S is a set, then T has a model $\mathcal{M} \vDash T$ with $|M| \geq |S|$.*

Theorem 3.6 (Downward Löwenheim-Skolem)**.** *If T has an infinite model $\mathcal{M} \vDash T$, and S is a set with $|S| \leq \mathcal{M}$, then T has a model $\mathcal{N} \vDash T$ with $|N| = |S|$, (and $N \subseteq M$).*

To prove the whole Löwenheim-Skolem theorem, to get a model of size $|S|$, we start with an infinite model $M \vDash T$, grow it with Upward Löwenheim-Skolem until it has size at least $|S|$, and then shrink it until it has size exactly $|S|$.

In lecture, we will just show the following simpler versions:

Theorem 3.7 (Uncountable Upward Löwenheim-Skolem)**.** *If T has an infinite model, then T has an uncountable model.*

Theorem 3.8 (Countable Downward Löwenheim-Skolem)**.** *If T has an infinite model $\mathcal{M} \vDash T$, then T has a countable model, which is a subset of M.*

3.2 Proving Upward Löwenheim-Skolem

To prove Uncountable Upward Löwenheim-Skolem, we will we want to find a theory, all of whose models are uncountable models of T. We will then show this theory is consistent, using compactness.

To do this, we will expand our language by adding a whole bunch of constants.

Let $\mathcal{L}_{\mathbb{R}}$ be the language obtained by adding constant symbols $\{c_r : r \in \mathbb{R}\}$ to $\mathcal{L}$. Let $\tilde{T}$ be the theory in $\mathcal{L}_{\mathbb{R}}$ consisting of T with the following new sentences added:

$$\tilde{T} = T \cup \{c_r \neq c_s : r, s \in \mathbb{R}, r \neq s\}.$$

Any model $\mathcal{M} \vDash \tilde{T}$ of this theory must be a model of T, and the function $\mathbb{R} \mapsto (c_r)^{\mathcal{M}}$ will be injective, by the fact that $\mathcal{M} \vDash c_r \neq c_s$ for any $r \neq s$. Thus $|\mathbb{R}| \leq |M|$, and in particular, M is uncountable.

To get this model of $\tilde{T}$, we need to show $\tilde{T}$ is consistent, and to do this, it's enough to show that any finite subtheory $T_0 \subseteq \tilde{T}$ is consistent. Any finite subtheory involves only a finite set of the constants - say $\{c_r : r \in F\}$ for some finite set $F \subseteq \mathbb{R}$. Thus we can pick any infinite model $\mathcal{M} \vDash T$, and label finitely many distinct elements with the constants $\{c_r : r \in F\}$. Because these are distinct, any sentence $c_r \neq c_s$ for $r, s \in F$ distinct will be satisfied, and all of T is satisfied, so T_0 is satisfied, and thus T_0 is consistent.

Homework

I recommend doing Problem 3.9, as well as revisiting the first few problems from yesterday's homework if you did not do them. Make sure you understand how we can use the compactness theorem to find models with interesting properties.

I also have some hints at the proof of Lemma 3.3, some problems about isomorphisms in interesting structures, and a review of cardinality, if you want to prove some of the stronger statements of Löwenheim-Skolem.

Problem 3.9. In general, given two sets A, B, we say that $|A| \leq |B|$ when there exists an injection $i : A \hookrightarrow B$, or equivalently, there exists a bijection between A and a subset of B.

Prove the stronger version of Upward Löwenheim-Skolem from the notes, using the same proof idea as Uncountable Upward Löwenheim-Skolem.

3.3 Proving Isomorphisms Preserve Formulas

Here's are some hints to get started proving Lemma 3.3. Throughout, let g be an isomorphism between $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$.

Problem 3.10. Suppose that $f \in \mathcal{F}$, and $t_1, \ldots, t_{n_f}$ are terms such that for each $1 \leq i \leq n_f$, using variables $x_1, \ldots, x_k$, and $m_1, \ldots, m_k \in M$, then

$$g\left(t_i^{\mathcal{M}}(m_1, \ldots, m_k)\right) = t_i^{\mathcal{N}}(g(m_1), \ldots, g(m_k)).$$

Show that

$$g\left(f(t_1, \ldots, t_{n_f})^{\mathcal{M}}(m_1, \ldots, m_k)\right) = (f(t_1, \ldots, t_{n_f}))^{\mathcal{N}}(g(m_1), \ldots, g(m_k))$$

Problem 3.11. Show by induction that if t is a term using variables $x_1, \ldots, x_k$, and $m_1, \ldots, m_k \in M$, then

$$g\left(t^{\mathcal{M}}(m_1, \ldots, m_k)\right) = t^{\mathcal{N}}(g(m_1), \ldots, g(m_k)).$$

Problem 3.12. Show that if $r \in \mathcal{R}$, and $t_1, \ldots, t_{n_r}$ are terms, using variables $x_1, \ldots, x_k$, and $m_1, \ldots, m_k \in M$, then

$$\mathcal{M} \vDash r(t_1, \ldots, t_{n_r})(m_1, \ldots, m_k) \iff \mathcal{N} \vDash r(t_1, \ldots, t_{n_r})(g(m_1), \ldots, g(m_k)).$$

Problem 3.13. Show that if ϕ, ψ are formulas using free variables $x_1, \ldots, x_k$ and $m_1, \ldots, m_k \in M$, such that

$$\mathcal{M} \vDash \phi(m_1, \ldots, m_k) \iff \mathcal{N} \vDash \phi(g(m_1), \ldots, g(m_k))$$

and

$$\mathcal{M} \vDash \psi(m_1, \ldots, m_k) \iff \mathcal{N} \vDash \psi(g(m_1), \ldots, g(m_k)),$$

then

$$\mathcal{M} \vDash (\phi \wedge \psi)(m_1, \ldots, m_k) \iff \mathcal{N} \vDash (\phi \wedge \psi)(g(m_1), \ldots, g(m_k)).$$

Problem 3.14. Show that if ϕ is a formula using free variables $x_0, \ldots, x_k$ and $m_1, \ldots, m_k \in M$, such that for all $m_0 \in M$,

$$\mathcal{M} \vDash \phi(m_0, m_1, \ldots, m_k) \iff \mathcal{N} \vDash \phi(g(m_0), g(m_1), \ldots, g(m_k)),$$

then

$$\mathcal{M} \vDash \exists x_0, \phi(x_0, m_1, \ldots, m_k) \iff \mathcal{N} \vDash \exists x_0, \phi(x_0, g(m_1), \ldots, g(m_k)).$$

3.4 Orders

Problem 3.15. Let $\mathcal{M}, \mathcal{N}$ be two linear orders (models of the theory of linear orders in the language $\{<\}$).

Show that any increasing bijection $g : M \to N$ is an isomorphism in the language $\{<\}$.

Problem 3.16. Show that for any $r \in \mathbb{R}$ and $q_1, \ldots, q_n \in \mathbb{Q}$, there is an isomorphism $g : \mathbb{R} \to \mathbb{R}$ such that for each $1 \leq i \leq n$, $g(q_1) = q_1$, and $g(r) \in \mathbb{Q}$.

3.5 Groups

Problem 3.17. Show that any group isomorphism is an isomorphism in the language $\{1, \times, ^{-1}\}$.

3.6 Rings

Problem 3.18. Show that any ring isomorphism is an isomorphism in the language $\{0, 1, +, \times\}$.

3.7 Linear Algebra

Problem 3.19. Show that any bijective linear transformation between K-vector spaces is an isomorphism.

Problem 3.20. Show that two vector spaces are isomorphic in that language if and only if they have the same dimension. (For infinite-dimensional vector spaces, V and W have the same dimension when there is a bijection between a basis of V and a basis of W.)

This is particularly cool because of the following theorem:

Theorem 3.21. *The theory $T_{K-\mathsf{vs}}$ of K-vector spaces is complete.*

Problem 3.22. Using Theorem 3.21, show that there are two $\mathbb{Q}$-vector spaces of the same cardinality that are elementarily equivalent but not isomorphic.

Problem 3.23. Challenge: Assuming the full version of Löwenheim-Skolem, prove Theorem 3.21.

3.8 Cardinality Review

The best way to compare the sizes of infinite sets (it also works for finite sets) is to look at functions between them. If A, B are sets, and $f : A \to B$ is an *injection* (recall that this means that any distinct $a_1 \neq a_2$ map to distinct $f(a_1) \neq f(a_2)$), we say that $|A| \leq |B|$. The injection f is basically identifying the elements of A with *some* elements of B, which are the range of f.

If there is a *bijection* $f : A \to B$, then we say $|A| = |B|$. Recall that this means an injection which is also a *surjection*, so every $b \in B$ is $f(a) = b$ for some $a \in A$, or equivalently a function with an inverse.

If $f : A \to B$ is an injection, we can think of f as a bijection between A and its range, which is a subset of B. Thus saying $|A| \leq |B|$ is the same thing as saying that $|A|$ is the size of a subset of B.

This $\leq$ relation on sizes of sets is reflexive, and transitive, but is it antisymmetric?

Theorem 3.24 (Cantor-Schröder-Bernstein)**.** *If $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.*

This means that $\leq$ is a partial ordering on the sizes, or *cardinalities*, of sets, which we usually call the *cardinal numbers*.

Assuming the AXIOM OF CHOICE, which we always do in this class, this is a linear order, so if A, B are sets, either $|A| \leq |B|$, or $|B| \leq |A|$, so we can think of the cardinal numbers as one big number line of (mostly infinite) sizes.

Of particular interest to us are *countably infinite* sets. A set A is *countable* if $|A| \leq |\mathbb{N}|$. This is equivalent to there being a surjection $f : \mathbb{N} \to A$, which we can think of as counting out the elements of A: $f(0), f(1), f(2), \ldots$

Problem 3.25. Show that if X, Y are infinite, then $|X \times Y| = \max(|X|, |Y|)$. (Recall the proof that $|\mathbb{Q}| = |\mathbb{N}|$, which is really showing $|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$.)

Problem 3.26. Show by induction that if X is infinite, then $|X^n| = |X|$.

Problem 3.27. Show that if $X_1, X_2, X_3, \ldots$ have the same infinite cardinality, then $\left|\bigcup_{n\in\mathbb{N}} X_n\right| = |X_1|$. (Use a bijection with $\mathbb{N} \times X_1$.)

4 Countable Categoricity and Substructures

Why do we care about Löwenheim-Skolem? One reason is that it can help us show certain theories are complete.

To show that a consistent theory is complete, we need to show that all of its models are elementarily equivalent. One way to do that *would* be to show that all of its models are isomorphic, because isomorphic structures are elementarily equivalent. However, Löwenheim-Skolem tells us that'll never happen, as there will be models of different cardinalities, which cannot be isomorphic.

In this case, we can see right away that countable $\mathbb{Q}$ is not isomorphic to uncountable $\mathbb{R}$. For these in the language $\{<\}$, we can do the next best thing. They both satisfy the theory of *dense linear orders without endpoints*, which we will abbreviate as DLO. We could write this theory in the language $\{\leq\}$, but it's slightly easier to axiomatize in $\{<\}$. It starts with the theory of linear orders:

- $\forall x, \neg x < x$ (irreflexivity)
- $\forall x, y, \neg(x < y \wedge y < x)$ (antisymmetry)
- $\forall x, y, z, x < y \wedge y < z \rightarrow x < z$ (transitivity)
- $\forall x, y, x < y \vee y < x \vee x = y$ (linearity)

and then we add axioms ensuring that it is *dense* (any two elements of the order have another between them) and lacks endpoints:

- $\forall x, y, x < y \rightarrow \exists z, x < z \wedge z < y$ (density)
- $\forall x, \exists y, y < x$ (no bottom endpoint)
- $\forall x, \exists y, x < y$ (no top endpoint).

As DLO consists of precisely 7 axioms the way we've expressed it, it's finitely axiomatizable. It also has two very famous models, $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$, which we know to be elementarily equivalent. This isn't an accident - DLO is complete. This means these 7 axioms are all we need to say to understand everything the language $\{<\}$ can say about $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$.

All of the *countable* models of DLO are isomorphic to $(\mathbb{Q}, <)$, making this an example of a *countably categorical* theory:

Definition 4.1. If all countable infinite models of a theory T are isomorphic, we call T *countably categorical.*
In general, if all models $\mathcal{M} \vDash T$ with $|M| = \kappa$ are isomorphic, then we call T *κ-categorical.*

In this case, countably categorical is the same thing as $|\mathbb{N}|$-categorical.
When we combine this with Löwenheim-Skolem, we can prove completeness:

Theorem 4.2 (Łoś-Vaught Test)**.** *If T is a consistent $\mathcal{L}$-theory with no finite models, and T is κ-categorical for some $\kappa \geq |\mathcal{L}|$, then T is complete.*

Proof. We will prove this in the case where $\kappa = |\mathbb{N}|$ - that is, for countably categorical T.
If T is *not* complete, then we can suppose for contradiction that there are models $\mathcal{M} \vDash \phi$ while $\mathcal{N} \vDash \neg\phi$.
Because T has no finite models, $\mathcal{M}$ is infinite, so it is an infinite model of $T \cup \{\phi\}$. If we apply Countable Downward Löwenheim-Skolem to $T \cup \{\phi\}$, we find that there is some countable model $\mathcal{M}^* \vDash T \cup \{\phi\}$, meaning $\mathcal{M}^* \vDash T$ and $\mathcal{M}^* \vDash \phi$. Similarly, there is a countable model $\mathcal{N}^* \vDash T$ such that $\mathcal{N}^* \vDash \neg\phi$, by starting with $\mathcal{N}$.

Because $\mathcal{M}^*, \mathcal{N}^* \models T$, both are countable, and T is countably categorical, $\mathcal{M}^*, \mathcal{N}^*$ are isomorphic, so $\mathcal{M}^* \equiv \mathcal{N}^*$.
However, this contradicts our observation that $\mathcal{M}^* \models \phi$ while $\mathcal{N}^* \models \neg\phi$. □

4.1 Substructure Basics

We're able to grow a model to arbitrary sizes with compactness, but for Downward Löwenheim-Skolem, we want to shrink a model down to a subset of a particular size. To do this, we need to understand, given a structure $\mathcal{M}$ in a language $\mathcal{L}$, how to look at certain subsets of M as structures themselves.

Definition 4.3. Let $\mathcal{M}$ be a $\mathcal{L}$-structure, and let $S \subseteq M$. We say that S is an $\mathcal{L}$-*substructure* of $\mathcal{N}$ when

- For every constant symbol c, $c^{\mathcal{M}} \in S$
- For every function symbol f, and $s_1, \dots, s_{n_f} \in S$, $f^{\mathcal{M}}(s_1, \dots, s_{n_f}) \in S$ (S is *closed* under f).

If this happens, then there is a unique way to interpret S as a $\mathcal{L}$-structure $\mathcal{S}$ in such a way that every constant, function, and relation symbol will agree with its interpretation in $\mathcal{M}$.

- For every constant symbol c, $c^{\mathcal{S}} = c^{\mathcal{M}}$
- For every function symbol f, and each $s_1, \dots, s_{n_f} \in S$, $f^{\mathcal{S}}(s_1, \dots, s_{n_f}) = f^{\mathcal{M}}(s_1, \dots, s_{n_f})$
- For every relation symbol r, and each $s_1, \dots, s_{n_r} \in S$, $r^{\mathcal{S}}(s_1, \dots, s_{n_r})$ is true if and only if $r^{\mathcal{M}}(s_1, \dots, s_{n_r})$ is.

If a structure $\mathcal{M}$ happens to be a substructure of a larger structure $\mathcal{N}$, with all symbols interpreted the same way in both structures, we call $\mathcal{N}$ an *extension* of $\mathcal{M}$.

In a purely relational language, such as the languages of orders or graphs, there are no constants or function symbols to worry about, so any subset is a substructure. In the case of graphs, these substructures are *induced subgraphs*, where vertices are adjacent in the subgraph if and only if they are in the original graph. In algebra, you may already be familiar with some kinds of substructures, like subgroups, subrings, and subfields. Our definition of substructure can align with those, so that, for instance, a subset of a group is an $\mathcal{L}$-substructure if and only if it's a subgroup. However, we need to be careful which language we use!

For some examples in the language $\{0, 1, +, \times, \leq\}$, $\mathbb{N}$ is a substructure of $\mathbb{Z}$ is a substructure of $\mathbb{Q}$ is a substructure of $\mathbb{R}$. Note that while we often call this the language of "ordered rings", $\mathbb{N}$ is not a subring of $\mathbb{Z}$, because it is not closed under subtraction! We would need to add in a function symbol for subtraction to make sure that subrings are closed under subtraction.

For any structure, it is easy to construct a very small substructure, using terms.

Lemma 4.4. *If $\mathcal{M}$ is a $\mathcal{L}$-structure, the subset*

$$S = \{t^{\mathcal{M}} : t \text{ is a term with no variables}\}$$

is a substructure.

Proof. For each constant c, c is a term, so $c^{\mathcal{M}} \in S$ by definition.
For each function symbol f and $s_1, \dots, s_{n_f} \in S$, we can find terms $t_1, \dots, t_{n_f}$ such that $s_i = t_i^{\mathcal{M}}$,

so then we build a term $t = f(t_1, \ldots, t_{n_f})$, and then

$$t^{\mathcal{M}} = f^{\mathcal{M}}(t_1^{\mathcal{M}}, \ldots, t_{n_f}^{\mathcal{M}}) = f^{\mathcal{M}}(s_1, \ldots, s_{n_f}).$$

This is in S, so S is closed under f. $\square$

Now we want to show this substructure is countable. Do do that, we just need to show the set of terms is countable. This is true when $\mathcal{L}$ is countable - for now, we will just assume $\mathcal{L}$ is finite.

In this case, a term is just a string of symbols from a finite set. We can encode every element of that finite set with a binary number of sufficient length, as in the ASCII encoding that computers use to store letters. Then we chain those together to get a really big binary string, which can be interpreted as a number. This means that we can assign each term a unique number, so there's an injection from the set of terms into $\mathbb{N}$, meaning the set of terms is countable.

If the language $\mathcal{L}$ is countably infinite, you can still encode every string as a natural number in a reasonable way, but the procedure is a little more complicated.

This argument shows us that if $\mathcal{L}$ is countable, then the set of terms is countable, and the set of formulas is countable, but also the substructure S we constructed is countable.

4.2 Elementary Substructures

We've seen that given S with $\max(|\mathbb{N}|, |\mathcal{L}|) \leq |S| \leq |M|$, then $\mathcal{M}$ has a substructure of size $|S|$.

However, for Downward Löwenheim-Skolem, we want to take a model of a theory T and shrink it down to *another model of* T, and a substructure of a model can satisfy completely different sentences!

To fix this, let's look at substructures that are as similar as possible to the larger structure, at least as far as formulas are concerned.

Definition 4.5. We say that $\mathcal{M}$ is an *elementary* $\mathcal{L}$-substructure of $\mathcal{N}$ (and that $\mathcal{N}$ is an *elementary* extension of $\mathcal{M}$) when $\mathcal{M}$ is a substructure of $\mathcal{N}$ such that for all formulas $\phi(x_1, \ldots, x_k)$, and all $m_1, \ldots, m_k \in M$, $\mathcal{M} \vDash \phi(m_1, \ldots, m_k)$ if and only if $\mathcal{N} \vDash \phi(m_1, \ldots, m_k)$. This gets the notation $\mathcal{M} \preceq \mathcal{N}$.

If $\mathcal{M} \preceq \mathcal{N}$, then among other things, $\mathcal{M} \equiv \mathcal{N}$, because sentences are formulas too. This lets us do a fun proof technique where we prove something is true in a structure by first proving it in some elementary extension, which we've specially constructed to make the proof easier. This is the model-theoretic approach to nonstandard analysis.

For an example of an elementary substructure, let's strip away all the arithmetic, and just look at the ordering, on $\mathbb{Q}$ and $\mathbb{R}$. In just the language $\{\leq\}$ of orders, $\mathbb{Q} \preceq \mathbb{R}$. This may not be obvious yet, but there is a test for determining if a substructure is elementary, which we can also use to show $\mathbb{Q} \preceq \mathbb{R}$:

Theorem 4.6 (Tarski-Vaught Test)**.** *Let $\mathcal{M}$ be a substructure of $\mathcal{N}$. We have $\mathcal{M} \preceq \mathcal{N}$ if and only if for every formula $\phi(x, y_1, \ldots, y_k)$ and $m_1, \ldots, m_k \in M$, if there exists $n \in N$ such that $\mathcal{N} \vDash \phi(n, m_1, \ldots, m_k)$, then there also exists $m \in M$ such that $\mathcal{N} \vDash \phi(m, m_1, \ldots, m_k)$.*

Basically, the criterion here says that any time $\mathcal{N} \vDash \exists x, \phi(x, m_1, \ldots, m_k)$, we can choose the witness to be in M.

The proof of this is in the homework! It starts by showing that if a sentence has no quantifiers, then it's always true in $\mathcal{M}$ if and only if it's true in $\mathcal{N}$. Then if we rewrite all the $\forall$s using $\exists$ and $\neg$, we can use the quantifier-free formulas as a base case, and induct on the number of $\exists$s. The inductive step has two directions - one direction is easy (if something exists in M, then it exists in N), and the other direction is given by the Tarski-Vaught Test criterion.

Homework

4.3 $\mathbb{Q} \preceq \mathbb{R}$

Let's use the Tarski-Vaught test to show that $(\mathbb{Q}, <) \preceq (\mathbb{R}, <)$.

To do this, we suppose $q_1, \ldots, q_k \in \mathbb{Q}$, and let $\phi(x, y_1, \ldots, y_k)$ be a formula in the language $\{<\}$, such that there is some $r \in \mathbb{R}$ with $\mathbb{R} \vDash \phi(r, q_1, \ldots, q_k)$. We want to show that there is also some $q \in \mathbb{Q}$ such that $\mathbb{R} \vDash \phi(q, q_1, \ldots, q_k)$.

Problem 4.7. Show that there is an isomorphism (just an increasing bijection) $g : \mathbb{R} \to \mathbb{R}$ such that for each $1 \leq i \leq n$, $g(q_1) = q_1$, and $g(r) \in \mathbb{Q}$.

Hint: Either $r \in \{q_1, \ldots, q_k\}$, r is less than all q_i, r is greater than all q_i, or r is directly between two elements of $\{q_1, \ldots, q_k\}$, call them $q_i < r < q_j$. In the latter case, find an isomorphism such that $g(x) = x$ unless $q_i < x < q_j$, and stretch the interval (q_i, q_j) around a bit to send r to some $q \in \mathbb{Q} \cap (q_i, q_j)$.

Problem 4.8. Show that

$$\mathbb{R} \vDash \phi(r, q_1, \ldots, q_k) \iff \mathbb{R} \vDash \phi(g(r), q_1, \ldots, q_k).$$

4.4 Checking the Tarski-Vaught Test

Let $\mathcal{M}$ be a substructure of $\mathcal{N}$.

Problem 4.9. Let $\phi(x_1, \ldots, x_k)$ be an atomic formula, and let $m_1, \ldots, m_k \in M$. Show that $\mathcal{M} \models \phi(m_1, \ldots, m_k)$ if and only if $\mathcal{N} \models \phi(m_1, \ldots, m_k)$.

Problem 4.10. Let $\phi(x_1, \ldots, x_k)$ be a quantifier-free formula, and let $m_1, \ldots, m_k \in M$. Show that $\mathcal{M} \models \phi(m_1, \ldots, m_k)$ if and only if $\mathcal{N} \models \phi(m_1, \ldots, m_k)$.

Problem 4.11. Finish the proof of the Tarski-Vaught Test:

Suppose that for every formula $\phi(x, y_1, \ldots, y_k)$ and all $m_1, \ldots, m_k \in M$, if there exists $n \in N$ such that $\mathcal{N} \models \phi(n, m_1, \ldots, m_k)$, then there also exists $m \in M$ such that $\mathcal{N} \models \phi(m, m_1, \ldots, m_k)$. Show that $\mathcal{M} \preceq \mathcal{N}$.

(Remember that you will want to swap out $\forall \phi$ with $\nexists \not\phi$, and induct on the number of quantifiers.)

4.5 Generating Substructures

Given a subset $S \subseteq M$ of an $\mathcal{L}$-structure $\mathcal{M}$, what's the smallest substructure of $\mathcal{M}$ containing S?

Problem 4.12. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Let $\langle S \rangle = \{t^{\mathcal{M}}(s_1, \ldots, s_n) : \text{t is a term}, s_1, \ldots, s_n \in S\}$. Prove that $\langle S \rangle$ is a substructure of $\mathcal{M}$, containing S.

Problem 4.13. If S is infinite, show that $|\langle S \rangle| = \max(|\mathcal{L}|, |S|)$. (Every element of $\langle S \rangle$ can be written with a term $t(s_1, \ldots, s_n)$, which is a finite string of characters, coming from $\mathcal{L} \cup \{0, 1, +, \times, \leq\} \cup S$. How many such finite strings are there?)

Problem 4.14. Show that if $\max(|\mathbb{N}|, |\mathcal{L}|) \leq |S| \leq |M|$, then $\mathcal{M}$ has a substructure of size $|S|$.

4.6 Groups

Problem 4.15. Show that if $\mathcal{M}$ is a multiplicative group, viewed as a structure in the language $\{1, \times, ^{-1}\}$ of multiplicative groups, the substructures of $\mathcal{M}$ are precisely the subgroups.

4.7 Linear Algebra

Problem 4.16. Let's revisit a challenge problem from yesterday: Showing the theory $T_{\mathbb{Q}-\mathsf{vs}}$ of $\mathbb{Q}$-vector spaces is complete.

Suppose $\mathcal{M}, \mathcal{N} \vDash T_{\mathbb{Q}-\mathsf{vs}}$ violate completeness, in that there is a sentence ϕ such that $\mathcal{M} \vDash \phi$ and $\mathcal{N} \vDash \neg\phi$.

- Show that there are $\mathcal{M}' \equiv \mathcal{M}, \mathcal{N}' \equiv \mathcal{N}$ such that $|\mathcal{M}'| = |\mathcal{N}'| = |\mathbb{R}|$ using Löwenheim-Skolem.
- Show that $\mathcal{M}'$ and $\mathcal{N}'$ have dimension $|\mathbb{R}|$, and must be isomorphic.
- Conclude a contradiction.

4.8 Strengthening Upward Löwenheim-Skolem

Here's the strong version of Upward Löwenheim-Skolem:

Theorem 4.17 (Upward Löwenheim-Skolem)**.** *If $\mathcal{M}$ is an infinite $\mathcal{L}$-structure and S is a set with $|M| \leq |S|$, then $\mathcal{M}$ has an elementary extension of size $|S|$.*

In some optional exercises, we will prove this, with a tool for building elementary extensions.

Given a language $\mathcal{L}$ and an $\mathcal{L}$-structure $\mathcal{M}$, define the language $\mathcal{L}_\mathcal{M}$ by adding a new constant symbol c_m for every $m \in M$, getting $\mathcal{L}_\mathcal{M} := \mathcal{L} \cup \{c_m : m \in M\}$. Then we can interpret $\mathcal{M}$ as an $\mathcal{L}_\mathcal{M}$-structure by interpreting each c_m as m, let's call this structure $\mathcal{M}^*$.

Now let the *elementary diagram* of $\mathcal{M}$ be $\mathrm{Th}(\mathcal{M}^*)$. What kind of sentences are in this theory? In general, a $\mathcal{L}_\mathcal{M}$-formula is just an $\mathcal{L}$-formula where we've replaced some of the variables with elements of M. That means that the sentences will all take the form $\phi(c_{m_1}, \ldots, c_{m_n})$ where $\phi(x_1, \ldots, x_n)$ is an $\mathcal{L}$-formula. So really, the elementary diagram of $\mathcal{M}$ is the set of all formulas $\phi(c_{m_1}, \ldots, c_{m_n})$ where ϕ is an $\mathcal{L}$-formula and $\mathcal{M} \models \phi(m_1, \ldots, m_n)$.

Problem 4.18. Let $\mathcal{N}$ be a $\mathcal{L}_\mathcal{M}$-structure, where $M \subseteq N$ and $c_m^\mathcal{N} = m$ for each $m \in M$.

Show that $\mathcal{N} \vDash \mathrm{Th}(\mathcal{M}^*)$ if and only if $\mathcal{M} \preceq \mathcal{N}$ as $\mathcal{L}$-structures.

Problem 4.19. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Show that if $\mathcal{N}$ is a model of the elementary diagram of $\mathcal{M}$, then there is an elementary extension of $\mathcal{M}$ with the same cardinality as $\mathcal{N}$.

Problem 4.20. Finish the proof of the strong version of Upward Löwenheim-Skolem.

5 Finishing Downward Löwenheim-Skolem: Skolem Functions

In our proof of Downward Löwenheim-Skolem, we know that we can construct a substructure of our chosen size, but it's not necessarily. We want to construct a small substructure that *is* elementary, and to check that it's elementary, we'll use the Tarski-Vaught Test. What follows is an attempt to custom-build a substructure to satisfy the Tarski-Vaught Test.

Fundamentally what we want to do is to extend the language $\mathcal{L}$ to a language $\mathcal{L}^*$ by adding a small number of function symbols, and extending $\mathcal{M}$ to a structure $\mathcal{M}^*$ by adding interpretations of those function symbols. Then we will find a small substructure of $\mathcal{M}^*$, which we call $\mathcal{N}^*$. If we peel away the extra function symbols, we can also view $\mathcal{N}^*$ as an $\mathcal{L}$-substructure, $\mathcal{N}$. If we are careful with how we add the function symbols and their interpretations, we can ensure that $\mathcal{N}$ satisfies the requirements of the Tarski-Vaught Test, so $\mathcal{N} \preceq \mathcal{M}$ (at least in the original language $\mathcal{L}$), and we are done.

The basic idea is that of *Skolem functions.* Let $\phi(x, y_1, \ldots, y_k)$ be an $\mathcal{L}$-formula. Then f_ϕ is a *Skolem function for ϕ in $\mathcal{M}$* if for all $m, m_1, \ldots, m_k \in M$ such that $\mathcal{M} \models \phi(m, m_1, \ldots, m_k)$, we have $\mathcal{M} \models \phi(f_\phi(m_1, \ldots, m_k), m_1, \ldots, m_k)$. We can always define such a Skolem function, defining $f_\phi(m_1, \ldots, m_k)$ to be an arbitrary value that makes $\mathcal{M} \models \phi(f_\phi(m_1, \ldots, m_k), m_1, \ldots, m_k)$ work if there is one, and defining it to be any element of M otherwise.

Define $\mathcal{L}^*$ to be the union $\mathcal{L} \cup \{f_\phi : \phi \text{ is an } \mathcal{L}\text{-formula}\}$. Define $\mathcal{M}^*$ by letting $f_\phi^{\mathcal{M}}$ be a Skolem function for ϕ in $\mathcal{M}$.

We've added a symbol for every formula, and there are countably many of these, so our language $\mathcal{L}^*$ is still countable. Thus $\mathcal{M}^*$ has a countable $\mathcal{L}^*$-substructure $\mathcal{N}^*$, which is also a $\mathcal{L}$-substructure. We now use Tarski-Vaught to check that $\mathcal{N}^* \preceq \mathcal{M}^*$.

For every $\phi(x, y_1, \ldots, y_k)$ and $n_1, \ldots, n_k \in N$, if there is $m \in M$ such that $\mathcal{M} \models \phi(m, n_1, \ldots, n_k)$, then $f_\varphi(n_1, \ldots, n_k) \in N$, because $\mathcal{N}^*$ is an $\mathcal{L}^*$-substructure, and by construction, $\mathcal{M} \models \phi(f_\varphi(n_1, \ldots, n_k), n_1, \ldots, n_k)$, satisfying Tarski-Vaught.

6 Back and Forth

Let's suppose $\mathcal{M}, \mathcal{N} \models \mathsf{DLO}$ are two countably infinite dense linear orders. To take advantage of countability, enumerate their elements as $M = \{m_0, m_1, m_2, \ldots\}$ and $N = \{n_0, n_1, n_2, \ldots\}$. We'll build an isomorphism between them step-by-step, and back-and-forth, by mapping m_0 into N, mapping n_0 into M, mapping m_1 into N, mapping n_1 into M, and so on. Remember as we go that an isomorphism in the language $\{<\}$ is just an increasing bijection. After we finish this proof, we can think about how we could extend it to other theories in other languages, where isomorphisms are more complicated. If you went to my colloquium, you have another example to generalize from: the theory of random graphs.

At step k, we will define finite substructures $A_k \subset M, B_k \subset N$, and an isomorphism between them: $f_k : A_k \to B_k$. To get to step $k+1$, we will try to add an element to each substructure, while extending the isomorphism.

Lemma 6.1. *If $f : A \to B$ is an isomorphism between finite substructures of $\mathcal{M}, \mathcal{N}$, and $m \in M$, we can extend A, B, f to A', B', f', where $f' : A' \to B'$ is still an isomorphism between finite substructures, but $m \in A'$.*

Proof. If $m \in A$ already, we're done. Otherwise, we just need to figure out what $f'(m) \in N$ should be, and add that to B'. We need to do this in a way that preserves the isomorphism, so it just needs to preserve the order.

As A, B are finite substructures of a model of DLO, they are finite linear order. We can write each of them as $A = \{a_1, \ldots, a_k\}, B = \{b_1, \ldots, b_k\}$, where $a_1 < a_2 < \cdots < a_k$ and $b_1 < b_2 < \cdots < b_k$, and we know that our isomorphism must send $f(a_i) = b_i$ to preserve the order.

Now m must fit somewhere into the order of A. Assuming $m \notin A$, then either $m < a_1$, for some i, $a_i < m < a_{i+1}$, or $a_k < m$. In each case, we find n that fits into the same spot in the order of B. In the case where $m < a_1$, we find $n < b_1$, because $\mathcal{N}$ has no least element. In the case where $m > a_k$, we find $n > b_k$, because $\mathcal{N}$ has no greatest element. In the case where $a_i < m < a_{i+1}$, we find $b_i < n < b_{i+1}$, because $\mathcal{N}$ is dense.

Now we see that if we let $A' = A \cup \{m\}, B' = B \cup \{n\}$, and extend the definition of f to f' by defining $f'(m) = n$, then f' is a bijection, and preserves the order, so it is an isomorphism. $\square$

Now to finish our proof. Start with any isomorphism $f_0 : A_0 \to B_0$ between finite substructures $A_0 \subset M$, $B_0 \subset N$. (We can start with $A_0 = B_0 = \emptyset$.)

We behave differently on even and odd steps. If $f_{2k} : A_{2k} \to B_{2k}$ has been defined, we want to define A_{2k+1} so that $m_k \in A_{2k+1}$. We can do this by using Lemma 6.1.

If $f_{2k+1} : A_{2k+1} \to B_{2k+1}$ has been defined, we use Lemma 6.1 to define B_{2k+2} so that $n_k \in B_{2k+2}$.

We now fast-forward until this countably infinite ladder of isomorphisms between substructures is defined. We checked that each $m_k \in \bigcup_{k=1}^{\infty} A_k$, so this union is M, and similarly, $\bigcup_{k=1}^{\infty} B_k = N$.

We now define $f : M \to N$ by saying that if $m \in A_k$, then $f(m) = f_k(m)$. This is a valid definition, and doesn't actually depend on the choice of k. If $m \in A_j$ and $m \in A_k$, then we constructed our functions such that $f_j(m) = f_k(m)$. Lastly, we check that f is increasing. For any two elements $x, y \in M$, assume without loss of generality that $x < y$. There is some j such that $x \in A_j$, and k such that $y \in A_k$, and if we let $\ell = \max(j, k)$, then $x, y \in A_\ell$. We know that f_ℓ is an isomorphism, so $f(x) = f_\ell(x) < f_\ell(y) = f(y)$, and we're done.

This shows that DLO is countably categorical. By the Łoś-Vaught test, this means that DLO is complete.

7 Bonus Material

Here is some other content that I prepared, but which didn't make it into the 1-week class.

7.1 Quantifier Elimination

The proof that DLO is countably categorical gives us something much more powerful. Suppose that $f_0 : A_0 \to B_0$ is any isomorphism of finite substructures $A \subseteq M$, $B \subseteq N$. We could start our proof with this isomorphism and proceed with all the steps, and construct from them an isomorphism $f : \mathcal{M} \to \mathcal{N}$. By the way this proof works, for any $a \in A_0$, we'd end up with $f(a) = f_0(a)$. This means we can extend any isomorphism of finite substructures of countable models of DLO to an entire isomorphism!

This lets us prove something that makes it drastically easier to study DLO: *quantifier elimination*.

Definition 7.1. We say a theory T *eliminates quantifiers* (or *has QE*) if for any formula $\phi(x_1, \ldots, x_n)$, there is a formula $\psi(x_1, \ldots, x_n)$ constructed without $\forall$ or $\exists$ such that

$$T \vDash \forall x_1, \ldots, \forall x_n, \phi \leftrightarrow \psi.$$

Theorem 7.2. DLO *eliminates quantifiers.*

Proof. Fix a formula $\phi(x_1, \ldots, x_n)$.

Call a formula $\sigma(x_1, \ldots, x_n)$ a *sign condition* if it is built from $\wedge$-ing together a bunch of formulas of the forms $x_i = x_j$ or $x_i < x_j$, with each pair (i, j) occuring exactly once.

Not every $\sigma(x_1, \ldots, x_n)$ actually has some $a_1, \ldots, a_n$ satisfying it, but every $a_1, \ldots, a_n$ in a linear order satisfies some unique sign condition. If $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ satisfy the same sign condition, this means that they have the exact same order type, so mapping $a_1 \mapsto b_1$ and so on to $a_n \mapsto b_n$ defines an isomorphism between finite substructures. Thus there is an isomorphism $f : \mathbb{Q} \to \mathbb{Q}$ with $f(a_i) = b_i$ for $1 \leq i \leq n$.

Let $\psi(x_1, \ldots, x_n)$ be the disjunction (a big $\bigvee$, an `OR`) of all the sign conditions satisfied by some $a_1, \ldots, a_n \in \mathbb{Q}$ with $\mathbb{Q} \vDash \phi(a_1, \ldots, a_n)$. Note that this has no quantifiers!

By definition, if $\mathbb{Q} \vDash \phi(b_1, \ldots, b_n)$, then $\mathbb{Q} \vDash \psi(b_1, \ldots, b_n)$.

Conversely, if $\mathbb{Q} \vDash \psi(b_1, \ldots, b_n)$, there is some sign condition σ in the big $\bigvee$ such that $\mathbb{Q} \vDash \sigma(b_1, \ldots, b_n)$. By definition, for σ to be put on the list, there are some $a_1, \ldots, a_n \in \mathbb{Q}$ with $\mathbb{Q} \vDash \phi(a_1, \ldots, a_n)$ such that $\mathbb{Q} \vDash \sigma(a_1, \ldots, a_n)$, so there is an isomorphism $f : \mathbb{Q} \to \mathbb{Q}$ with $f(a_i) = b_i$ for $1 \leq i \leq n$. This tells us that $\mathbb{Q} \vDash \phi(f(a_1), \ldots, f(a_n))$, so $\mathbb{Q} \vDash \phi(b_1, \ldots, b_n)$. $\square$

Quantifiers are a pain to deal with, so in practical model theory, the first thing you want to do when encountering a new structure is to try and prove a quantifier-elimination result. You may have to add some more symbols to the language then you started with, but in most nice structures, there is a nice language with quantifier-elimination, which is usually what we think of as the "correct" language to default to.

Here are some classic structures that have (or rather, their theories have) quantifier elimination, although the proofs are more involved:

- $(\mathbb{Q}, 0, +, <)$, an *ordered divisible abelian group*
- $(\mathbb{R}, 0, 1, +, \times, <)$
- $(\mathbb{C}, 0, 1, +, <)$
- Any vector space in the language from the homework.

7.2 PAradoxes

(Discuss Skolem's Paradox if not done before - probably should be done before.)

Yesterday, we proved that a particularly nice theory, DLO, is complete. From our back-and-forth argument, we could even prove quantifier elimination, which allows us to understand its definable sets very well. DLO is a very tame theory, and today we will talk about one of the wildest: Peano Arithmetic, or PA.

PA was an attempt to axiomatize $(\mathbb{N}, 0, 1, +, \times, \leq)$. It starts with some normal algebraic statements: (Do I want to get into these?)

However, what really sets the natural numbers apart from any similar algebraic structure is *induction*. Any subset $S \subseteq \mathbb{N}$ satisfies the following: If $0 \in S$ and $\forall x, x \in S \to x + 1 \in S$, then $S = \mathbb{N}$. We call such a set *inductive*.

This would be a neat way of describing what the natural numbers really are, from which we could prove everything else by induction, *but we'd have to quantify over sets.* The system of formulas we've built up this week (called *first-order logic*) has no symbols for that. (There's a system that does, called *second-order logic*, but almost nothing we've discussed this week works there.)

What we *can* do is notice that *definable* sets have to satisfy this property. If $\phi(x, y_1, \ldots, y_n)$ is a formula, then we can write down a sentence stating that any set defined by $\phi(x, b_1, \ldots, b_n)$ is inductive:

$$\forall y_1, \ldots, y_n, \phi(0, y_1, \ldots, y_n) \wedge (\forall x, \phi(x, y_1, \ldots, y_n) \to \phi(x + 1, y_1, \ldots, y_n)) \to \forall x, \phi(x, y_1, \ldots, y_n).$$

Call that sentence $\texttt{Ind}_\phi$. The theory PA consists of the above algebraic axioms (Check that this makes sense) together with $\texttt{Ind}_\phi$ for every formula $\phi(x)$.

Homework

7.3 Quantifier Elimination

Let me expand the definition of definable sets from the first homework:

Definition 7.3. Let $\mathcal{M}$ be an $\mathcal{L}$-structure.
A subset $A \subseteq \mathcal{M}^k$ is called *definable* when there is some $\mathcal{L}$-formula $\phi(x_1, \ldots, x_j, y_1, \ldots, y_k)$ and $b_1, \ldots, b_k \in M$ such that for all $a_1, \ldots, a_j \in M$, $\mathcal{M} \vDash \phi(a_1, \ldots, a_j, b_1, \ldots, b_k)$ if and only if $(a_1, \ldots, a_k) \in A$.

Problem 7.4. Using quantifier elimination, show that any definable subset $A \subseteq \mathbb{Q}^1$ in the language $\{<\}$ is a finite union of points and intervals.

Such a linearly-ordered structure is called *o-minimal*, and these are all the rage in number theory these days.

Problem 7.5. Show that any definable subset (again, in one variable) of $(\mathbb{C}, 0, 1, +, <)$ or a vector space is either finite or the complement of a finite set.

Any theory whose models all have this property is called *strongly minimal*, and understanding these revolutionized model theory in the 60's and 70's.

Problem 7.6. Assuming the quantifier elimination remark from the end of class, show that $(\mathbb{Q}, 0, +, <)$ and $(\mathbb{R}, 0, 1, +, \times, <)$ are o-minimal.

Problem 7.7. Show that $(\mathbb{N}, 0, 1, +, \times, <)$ and $(\mathbb{R}, \sin(x), <)$ are not o-minimal and does not have quantifier elimination.

7.4 Orders

Problem 7.8. Show that DLO is *not* $|\mathbb{R}|$-categorical.

(Hint: If A, B are linear orders, then the *lexicographical ordering* on $A \times B$, given by $(a_1, b_1) < (a_2, b_2)$ when $a_1 < a_2$ or $a_1 = a_2$ and $b_1 < b_2$ is also a linear order.)

8 Outro

What have we learned?

We've learned how to describe many different kinds of mathematical widget as the models of a particular theory.

We've learned that the compactness theorem can be used to show certain structures exist by only checking finitely many properties at a time. To do so, we cited the completeness theorem - if you want to understand the completeness theorem in greater detail, take Maya's class next week.

We've learned that infinite models can be grown or shrunk to whatever sizes we like. I wrote a song parody about this in Mathcamp 2021, and you can see the lyrics here: https://awainverse.github.io/misc/whiterabbit/

We've learned one way to show that a theory is complete, using the Łoś-Vaught test (which also can apply to the theory of the random graph, and as we've seen in homework, to vector spaces). There are other ways, including using quantifier elimination directly.

We have not discussed much about how to show a theory is incomplete. We've discussed that the theory of $(\mathbb{N}, 0, 1, +, \times, \leq)$ is complicated, and there is no good way of determining the truth of all sentences there. In fact, any "good" axiomatization of this theory is incomplete (basically Gödel's Incompleteness Theorems), and the set of binary strings representing sentences ϕ such that $\mathbb{N} \vDash \phi$ is not definable in $(\mathbb{N}, 0, 1, +, \times, \leq)$ (Tarski's Undefinability of Truth). You can probably ask any logic-interested staff about this, but Steve has taught classes on these topics before, so you should ask Steve about that.