141A MATHEMATICAL LOGIC I

ASSAF SHANI

This course is an introduction to mathematical logic, more specifically, to the branch of logic called *Model Theory*. At times we will need some basic results and notions from *Set Theory*, as well as some basic concepts from *Proof Theory* (other branches of mathematical logic), which we will cover accordingly.

1. Introduction

Before getting formal, let us recall some examples of mathematical structures.

1.1. Some mathematical structures.

Here are some examples of mathematical structures, some of which you may have seen, and how we think of them in the general context of **alphabets**, **axioms**, and **models**. We will make this more formal and precise soon.

Example 1.1. A **Linear order** is a pair $(X, <)$ so that

- X is a set (just some collection of objects);
- $<$ is a binary relation on X (meaning for $x, y \in X$ we may ask if $x < y$ is true or not),

and such that the following statements hold:

(LO 1) (Strictness) for any members x, y of the set X, if $x < y$ holds, then $y < x$ fails (we may denote this by $y \not< x$;

(LO 2) (Transitivity) for any members x, y, z of the set X, if $x < y$ and $y < z$, then $x < z$;

(LO 3) (Total (linear) ordering) for any members x, y of X, either $x < y$ or $y < x$ (or $x = y$).

Examples of linear orders include the natural numbers $(\mathbb{N}, <)$, the integers $(\mathbb{Z}, <)$, the rational numbers $(\mathbb{Q}, <)$, and the real numbers $(\mathbb{R}, <)$, all with the usual order $<$ which you are familiar with.

In this case, we will say that the **signature** (or **alphabet**) consists of one binary relation symbol "$<$"; conditions (LO 1),(LO 2),(LO 3), above are **axioms** in this language (for this alphabet); and a linear order $(X, <)$ is a **model** for this alphabet, which **satisfies these axioms**.

Let us emphasize that by a model, a mathematical structure, we really mean anything, in the most abstract sense, and not just familiar objects such as the natural numbers or the real numbers. Here are some examples.

Define a relation $<^*$ on the set $\mathbb{N}$ as follows:

$$n <^* m \text{ if and only if } n > m\text{, for any natural numbers } n, m.$$

Date: Spring 2022.

Then $(\mathbb{N}, <^*)$ is again a linear order, quite different than $(\mathbb{N}, <)$. (Note that the set of objects, in this case the natural numbers $\mathbb{N}$, has very little to do with the model (linear order). It is how we **interpret** the relation $<$ in it that matters.)

Another example would be to take $X = \{\square, \triangle\}$ and declare that $\square < \triangle$ and that $\triangle \not< \square$. Then $(X, <)$ is a linear order.

A "non-example" would be to take $X = \{\square, \triangle\}$ and declare that $\square \not< \triangle$ and that $\triangle \not< \square$. Then $(X, <)$ fails clause (3), so it does not satisfy the axioms of a linear order. It is just some structure for the alphabet "$<$".

Example 1.2. A **Graph** is a pair (V, E) so that

- V is a set (some collection of objects);
- E is a binary relation on V,

and such that the following statement holds:

(Graph 1) (Symmetry) for any x, y in V, if $x\ E\ y$ then $y\ E\ x$.

The members of V are often called the *vertices* of the graph and we say that there is an *edge* between x and y if $x\ E\ y$.

In this case, we will say that the **signature** consists of one binary relation symbol E; the condition (Graph 1) above is an **axiom** in this language; and a graph (V, E) is a **model** in this language, which **satisfies this axiom**.

Note that in both cases the models are just some collections of objects with a binary relation. That is, the signature for both graphs and linear orders is just a single binary relation. Whether we call it E or $<$ does not really matter. It is the different axioms that make the difference.

Example 1.3. A **Group** is a triplet $(G, \cdot, e)$ so that

- G is a set;
- $\cdot$ is a binary operation on G (meaning for any x, y in G the operation products another member $x \cdot y$ in G);
- e is a member of G,

so that

(Group 1) (Identity) for any x in G, $x \cdot e = x$ and $e \cdot x = x$;
(Group 2) (Associativity) for any x, y, z in G, $(x \cdot y) \cdot z = x \cdot (y \cdot z)$;
(Group 3) (Inverses) for any x in G there is some y in G (usually denoted x^{-1}) so that $x \cdot y = y \cdot x = e$.

Here the signature consists of two symbols, a **binary function symbol** $\cdot$, and a **constant symbol** e. A model $(G, \cdot, e)$ needs to interpret the function symbol $\cdot$ as a function, taking two members x, y in G and producing a third, z in G, and to interpret the constant symbol is a single member of G.

Examples of groups include

- $(\mathbb{Q}, +, 0)$;
- $(\mathbb{R}, +, 0)$;
- $(\{0, 1\}, \text{"}+ \mod 2\text{"}, 0)$ (addition mod 2, meaning $1 + 1 = 0$);
- $(\{1, -1\}, \cdot, 1)$ (usual product);
- $([0, 1), \text{"}+ \mod 1\text{"}, 0)$.

1.2. **Realization and formal proofs.** Fix now some alphabet, meaning some symbols for relations or functions. (For now, just think about one binary relation, as for orders and graphs.)

A **theory** is just a set of axioms (sentences) in the language, using this alphabet. For example, the theory for graphs is the single axiom (Graph 1) above, and the theory for linear orders is the conditions (LO 1), (LO 2), (LO 3) above.

In all the examples above, the axioms are simply some key properties of common structures which we encounter often in mathematics. One may look at things the other way around. Suppose you give me some axioms, can I find a mathematical model satisfying these axioms? Can these axioms be **realized**? For example, you may want to find a graph with some additional properties (say, "a graph with no triangles").

Of course, if your demands are outrageous, the answer will be no. You may include in your axioms the statement $\Phi =$"there exist some x so that $x \mathrel{E} x$ and $x \not\mathrel{E} x$". This statement contradicts itself, and you cannot find any structure satisfying Φ.

Once we formalize things more, there is actually something to prove here. This is the so called **Soundness theorem**, saying that nothing outrageously false can be true in an actual mathematical model. Something is "outrageously false" if we can use it to formally prove a contradiction: a statement and its negation. One thing we need to do is to formalize what a proof is. A key point here is that a proof will be a purely *syntactic* entity, just a sequence of steps following simple rules of deduction, *something a computer can do.*

Back to the question: suppose you came up with some theory T (a collection of axioms), which seems legit. Could it be that there *are no* models for it? Could it be that there is some inherent falsness in the axioms, beyond what we, or a computer, can see from the axioms themselves? The answer is no!

Theorem (The completeness theorem)**.** If the theory T has no mathematical models, then necessarily there is a formal proof of contradiction using the axioms in T.

Let us emphasize again: finding a proof of contradiction is something that can be verified syntactically. It is something a computer can find. Very much to the contrary, arguing that "a model does not exist" is not something a computer can even think about.

Another way of phrasing the completeness theorem is as follows: Suppose T is a theory and Φ is a statement. Assume that in any model satisfying T, the statement Φ holds true as well. Then there is in fact a formal proof of Φ from T. For example, if there is a statement that is true about *all groups*, then necessarily we can find a proof for it (using just the group axioms)! This is quite uplifting.

Closely related to the completeness theorem is the compactness theorem.

Theorem (The Compactness Theorem)**.** Suppose T is a theory (a collection of axioms, possibly infinitely many axioms). Assume that there is no mathematical model satisfying all the axioms in T (that is, T *cannot be realized*). Then in fact there are *finitely many* axioms $\Phi_0, ..., \Phi_N$ in T so that just the axioms $\Phi_0, ..., \Phi_N$ already cannot be realized.

You may think about it this way. Suppose you are trying to construct some (infinite) mathematical object. There are many (infinitely many) specific properties which you want this object to satisfy. Say you have a list of these requirements $\Phi_0, \Phi_1, \Phi_2, ...$ It may be hard to construct this object by hand, but you want to say that such an object exists. *By* *the compactness theorem, it is enough to show that any finitely many of your requirements,* $\Phi_0, ..., \Phi_N$*, can be realized.*

To emphasize that this is highly non-trivial, note that the realization of each finite chunk $\Phi_0, ..., \Phi_N$ can be by a different structure, depending on N, while at the end you get one object realizing all the requirements together. In fact, it may be the case that each finite chunk can be realized by a finite object, while at the end you get (from the theorem) an infinite object.

1.3. **Models.** Suppose we have a theory T, which is not inherently contradictory, so it does have models.

How many models are there?

How do we count?

First, if we just re-label a structure, it is not really different. Let $\mathbb{N}'$ be the set of objects $0', 1', 2', 3', ...$ (where $\mathbb{N}$ is the set of objects $0, 1, 2, 3, ...$). Define $n' <' k'$ if and only if $n < k$. Then $(\mathbb{N}', <')$ is a linear order. Technically, it is different than $(\mathbb{N}, <)$, in the sense than the objects are different symbols. Clearly, however, this is the same linear order (mathematical structure), just relabelled!

More specifically, there is a re-labelling function $f : \mathbb{N} \to \mathbb{N}'$, $f(n) = n'$, which is one-to-one and onto, and it respects the structures: given n, k in $\mathbb{N}$, $n < k$ if and only if $f(n) <' f(k)$. We will say that such a function is an **isomorphism** of linear orders. (We will define more generally, in a similar way, what is an isomorphism between arbitrary structures.)

We will say that two linear orders (two mathematical structures) $(X, <^X)$, $(Y, <^Y)$, are **isomorphic** if there is a relabelling map (an isomorphism) between them. In this case, we consider them as essentially the same structure.

So we want to know how many *different* structures there are "up to re-labelling".

Example 1.4. The linear orders $(\mathbb{Q}, <)$ and $(\mathbb{N}, <)$ are *not* isomorphic.

Proof. Assume, towards a contradiction, that $f : \mathbb{N} \to \mathbb{Q}$ is a one-to-one and onto function which satisfies

$$(\star) \;\; n < m \iff f(n) < f(m) \text{ for any } n, m \in \mathbb{N}.$$

Take 0 in $\mathbb{N}$, the smallest member, and let $q = f(0)$, a rational number. Define $p = q - 1$, so p is a rational number and $p < q$. By our assumption of f, there is some m so that $f(m) = p$. Applying $(\star)$ we conclude that $m < 0$, a contradiction! □

Example 1.5. Let $\mathbb{Z} \setminus \{0\}$ be the set of non-zero integers, and the usual ordering on them. Then $(\mathbb{Z} \setminus \{0\}, <)$ and $(\mathbb{Z}, <)$ *are* isomorphic.

Proof. Define $f : \mathbb{Z} \to \mathbb{Z} \setminus \{0\}$ as follows. For a negative integer n, define $f(n) = n$. For a non-negative integer n, define $f(n) = n + 1$. You can check that f is one-to-one, onto, and it satisfies that $n < k \iff f(n) < f(k)$, for any integers n, k. (Make sure you know how to argue for this.) □

We will be particularly interested in infinite structure. Another issue that comes up is that of *size* (cardinality). Recall that two (infinite) sets X and Y **have the same size** if there is a one-to-one and onto function $f : X \to Y$. So, if two structures $(X, <^X)$ and $(Y, <^Y)$ are isomorphic, necessarily the sets X and Y have the same size.

We will focus on the smallest infinite size, the countable infinite. (We will go over, and develop, these things in more detail. For now I assume you have heard a little about it in some way. Otherwise you may ignore this subtlety or think of it informally.) Recall that a set X is **countable** if there is an onto function $f\colon \mathbb{N} \to X$. (That is, we can enumerate all the members of X using the counting numbers.) For example, each of the sets $\mathbb{N}$,$\mathbb{Z}$, and $\mathbb{Q}$ is countable, while the set of real numbers $\mathbb{R}$ is famously not countable.

How many *countable* models are there, up to isomorphism?

Just one model. Here is an interesting and very important example.

Definition 1.6 (Dense (unbounded) linear orders)**.** Let the theory DLO (dense linear orders) consist of the three axioms of a linear order (LO 1), (LO 2), (LO 3) above, together with:

(Density) for any x and y, if $x < y$ then there is some z so that $x < z < y$;
(No min) for any x there is some y with $y < x$;
(No max) for every x there is some y with $x < y$.

These are some of the key properties of the rational numbers $\mathbb{Q}$, as an order (compared to $\mathbb{N}$ or $\mathbb{Z}$). In fact, these simple properties completely capture *everything* about the rational numbers.

Theorem 1.7 (Cantor's isomorphism theorem)**.** Suppose $(X, <^X)$ and $(Y, <^Y)$ are two linear orders satisfying the DLO axioms, and X and Y are countable infinite sets. Then they are isomorphic. That is, there is a relabelling function f as above.

So for the theory DLO of dense linear orders, there is only **one** (countable) model, up to isomorphism. In this case we will say that the theory is (countably) **categorical**.

Example 1.8. Each of the following two are isomorphic.

(1) $(\mathbb{Q}, <)$;
(2) $(\mathbb{Q}^+, <)$, where $\mathbb{Q}^+$ is the set of positive rationals;
(3) $(\mathbb{Q} \setminus \{0\}, <)$, where $\mathbb{Q} \setminus \{0\}$ is the set of non-zero rational numbers.
(4) $(\mathbb{Q} \setminus \mathbb{Z}, <)$, where $\mathbb{Q} \setminus \mathbb{Z}$ is the set of rationals which are not integers.

Exercise 1.9. Show that the structures in (2), (3), and (4) above satisfy the DLO axioms.

As we will see, the following are curious consequences of Cantor's isomorphism theorem.

Corollary 1.10. For any statement Φ (stated in the language using the relation $<$), either

- there is a formal proof of Φ, using the axioms DLO, or
- there is a formal proof that Φ is false, using the axioms DLO.

In this case we will say that the axioms DLO are **complete**. They "decide" every statement. We will investigate complete theories a lot later on.

Corollary 1.11. For any statement Φ (stated in the language using the relation $<$), the following are equivalent:

- The statement Φ is true in the structure $(\mathbb{Q}, <)$;
- The statement Φ is true in the structure $(\mathbb{R}, <)$.

So even though the two structures are not isomorphic (based on size consideration), one cannot really see the difference between them, *using the language of orders*.

Another consequence is that the "*theory of* $(\mathbb{R}, <)$" is decidable, meaning you can run a computer program that will spit out statements (in the language for $<$), so that the statements it spits out are precisely those statements true in $(\mathbb{R}, <)$. Other very interesting mathematical examples of structures with "decidable theories" are $(\mathbb{R}, +, \cdot, 0, 1)$ and $(\mathbb{C}, +, \cdot, 0, 1)$ (complex numbers).

Example 1.12. We will soon carefully formalize what "a statement in the language" means. For now, here are statements that can be made in the language using one binary relation "$<$":

"for any x and y there exists z so that $z < x$ and $z < y$";

"for any x there is some y so that $x < y$ and there is no z satisfying both $x < z$ and $z < y$".

Using the axioms in DLO, you can see how to prove that the first statement is true, and that the second statement is false.

Exercise 1.13.
1. Explain (informally) what each statement above means in terms of the order.
2. Determine whether each statement is true for the following linear orders:
   - $(\mathbb{N}, <)$;
   - $(\mathbb{Z}, <)$;
   - $(\mathbb{N}, <^*)$.

1.4. **More on isomorphisms.**

Proposition 1.14. Suppose $f\colon \mathbb{N} \to \mathbb{N}$ is an isomorphism of the structure $(\mathbb{N}, <)$ and (the same structure) $(\mathbb{N}, <)$. Then necessarily f is the identity map: $f(n) = n$ for all n in $\mathbb{N}$.

An isomorphism from a structure into itself is called an **automorphism**.

Proof. Let f be an isomorphism as in the statement. We prove by induction on $n = 0, 1, 2, ...$ that $f(n) = n$ for each n.

Let us start with $n = 0$. Assume towards a contradiction that $f(0) \neq 0$. Then $f(0) = k$ for some $k > 0$. Since f is onto, there is some m so that $f(m) = 0$. Since $f(m) = 0 \neq k = f(0)$, necessarily $m \neq 0$ (f is injective), and so $m > 0$. Now $0 < m$ yet $f(0) > f(m)$, a contradiction.

Assume that we know $f(0) = 0, ..., f(n) = n$, and we prove that $f(n+1) = n+1$. Again, assume towards a contradiction that $f(n+1) \neq n+1$. Since f is onto-to-one, it must be that $f(n+1) = k > n+1$. Since f is onto there must be some m so that $f(m) = n+1$. Again since f is injective we know that $m \neq n+1$ and also that $m \neq 0, ..., n$, so it must be that $m > n+1$. Again we arrive at a contradiction as $n+1 < m$ yet $f(n+1) > f(m)$.

We conclude that $f(n) = n$ for all n in $\mathbb{N}$, as required. □

A structure such as $(\mathbb{N}, <)$, which does not have any *non-trivial automorphism* (non-identity automorphisms) is called **rigid**.

Example 1.15. Consider the linear order $(\mathbb{Z}, <)$. Then for any a and b in $\mathbb{Z}$, there is an automorphism of $(\mathbb{Z}, <)$, a map $f\colon \mathbb{Z} \to \mathbb{Z}$ which is an isomorphism of orders, so that

$f(a) = b$. Simply define $f(x) = x + (b - a)$. Then f is a one-to-one and onto map from $\mathbb{Z}$ to $\mathbb{Z}$. Furthermore, for any x, y in $\mathbb{Z}$, $x < y \iff x + (b - a) < y + (b - a)$. Finally, $f(a) = a + (b - a) = b$.

Exercise 1.16.
(1) Prove that there *does not exist* an automorphism f of the linear order $(\mathbb{Z}, <)$ such that $f(0) = 1$ and $f(2) = 5$.
(2) ($\star$) In fact, the automorphisms of $(\mathbb{Z}, <)$ which we described above are *all* the possible automorphisms of $(\mathbb{Z}, <)$. That is, if f is an automorphism of $(\mathbb{Z}, <)$, prove that there is some integer c in $\mathbb{Z}$ so that for all x in $\mathbb{Z}$, $f(x) = x + c$.

Exercise 1.17 (If you are familiar with groups)**.** For groups $(G, \cdot^G, e^G)$ and $(H, \cdot^H, e^H)$ an isomorphism between them is a one-to-one and onto function $f \colon G \to H$ so that $f(e^G) = e^H$ and $f(x \cdot y) = f(x) \cdot f(y)$ for any x, y in G. [The first $\cdot$ is $\cdot^G$, the second is $\cdot^H$.] Similarly we define an automorphism of a group as an isomorphism from it to itself. The trivial automorphism is the identity map.

(1) Show that the group $(\mathbb{Z}, +, 0)$ is rigid, meaning it has no non-trivial automorphisms.
(2) Suppose that $(G, \cdot, e)$ is a *non-abelian* (non-commutative) group. Show that $(G, \cdot, e)$ is not rigid (there is a non-trivial automorphism).

1.5. **Proof of the isomorphism theorem for dense linear orders.** Suppose $(X, <^X)$ and $(Y, <^Y)$ are linear orders. Given sequences $\bar{a} = a_0, a_1, ..., a_{n-1}$ from X and $\bar{b} = b_0, b_1, ..., b_{n-1}$ from Y, say that $\bar{a}$ and $\bar{b}$ **have the same type** if

$$a_i < a_j \iff b_i <' b_j$$

for any $i, j \in \{0, ..., n-1\}$. (Equivalently: $\bar{a}$ and $\bar{b}$ have the same type if the map sending a_i to b_i is order preserving.) For a sequence $\bar{a} = a_0, a_1, ..., a_{n-1}$ from A, and some a in A, define $\bar{a}^\frown a$ as the sequence $a_0, a_1, ..., a_{n-1}, a_n$ with $a_n = a$.

Lemma 1.18 (Proved in Pset 1)**.** Suppose $(X, <^X)$ is some linear order, and $(Y, <^Y)$ is a dense linear order. Let $\bar{a}$ and $\bar{b}$ be sequences from X and Y accordingly, and assume that they have the same type. Then for any $a \in X$ there exists some $b \in Y$ such that the sequences $\bar{a}^\frown a$ and $\bar{b}^\frown b$ also have the same type.

Proof of Theorem 1.7. Let $(X, <^X)$ and $(Y, <^Y)$ be two dense linear orders, where X and Y are countable sets. We will prove that there is an isomorphism between the two structures. The key ideas in the proof are very important in model theory. This type of proof is often called a “back and forth construction”.

By assumption we have a list of *all* the members of X: $x_0, x_1, x_2, ...$ and a list of *all* the members of Y: $y_0, y_1, y_2, ...$ We may also assume that each member of X appears precisely once in the list $x_0, x_1, ...$ and similar for Y.

It is worth noting that the “order of enumeration” has nothing to do with the actual order in the structure. That is, we do not know if $x_0 <^X x_1$ or $x_1 <^X x_0$. For example, an attempt to define f from X to Y by sending x_n to y_n, will most likely fail to be an isomorphism (fail to respect the order). The proof will simultanously use these two very different structures, the orders $<^X$ and $<^Y$ which we are interested in, and the external “enumeration orders”. The function f will be defined in stages, where only finitely many values are dealt with at each stage, according to the enumerations. At each stage we make sure that the $x_1 <^X x_2 \iff f(x_1) <^Y f(x_2)$, for the finitely many values we are dealing with.

Let us start, a little informally, with a sketch of how to start this construction, and where it is going. First define $f(x_0) = y_0$. [Define also $a_0 = x_0$ and $b_0 = y_0$]

Next, where can we send x_1? First check its relation to x_0.
If $x_0 < x_1$, let m be the smallest number so that $m > 0$ and $y_0 < y_m$. (m may be 1, but may not.) Now define $f(x_1) = y_m$. [Define $a_1 = x_1$ and $b_1 = y_m$.]
If $a_1 < a_0$, let m be the smallest number so that $m > 0$ and $b_0 > b_m$, and define $f(a_1) = b_m$. [Define $a_1 = x_1$ and $b_1 = y_m$.]
So far, the function f does respect the orders, defined only on $\{x_0, x_1\}$. [Equivalently, the sequences a_0, a_1 and b_0, b_1 have the same type.]

We can play the same game with x_2, asking on its relationship with x_0, x_1. However, before doing that, recall that we want f to be not only order-preserving and injective, but also onto. What if we skipped y_1?
Suppose we did, that is, $m > 1$. Now there are three options for the relationship between y_1 and y_0, y_m:
If y_1 is smaller than both y_0 and y_m: Since $(X, <^X)$ is a DLO, there is some x_k so that x_k is smaller than both x_0, x_1. (Take the first such k we can find.) Define $f(x_k) = y_1$. [Define $a_2 = x_k$ and $b_2 = y_1$.]
If y_1 is above both y_0, y_m: We may find x_k above x_0, x_1, and define $f(a_k) = b_1$. [Define $a_2 = x_k$ and $b_2 = y_1$.]
If y_1 is between y_0 and y_m, again there are two cases.
Either $y_0 < y_1 < y_m$, then we can find (using the DLO assumption) some x_k so that $x_0 < x_k < x_1$. (Note that if $y_0 < y_m$ then by the previous step necessarily $x_0 < x_1$.) Define $f(x_k) = x_1$. [Define $a_2 = x_k$ and $b_2 = y_1$]
Otherwise, $y_m < y_1 < y_0$, and similarly we may find k so that $x_1 < x_k < x_0$, and define $f(x_k) = x_1$.
In any of these cases, f is still injective and respects the orders. [The sequences a_0, a_1, a_2 and b_0, b_1, b_2 have the same type.] Furthermore now f is defined on x_0, x_1, and y_0, y_1 are both in the image of f.

Formally, our construction is done recursively as follows. Assume that at stage n of our construction we defined a pair of sequences $\bar{a} = a_0, ..., a_{2n}$ from X and $\bar{b} = b_0, ..., b_{2n}$ from Y so that $\bar{a}$ and $\bar{b}$ *have the same type*. We define a_{2n+1}, a_{2n+2} and b_{2n+1}, b_{2n+2} as follows.

Let $t \in \mathbb{N}$ be the smallest natural number so that x_t is not one of $\{a_0, ..., a_{2n}\}$. Define $a_{2n+1} = x_t$. Find some b_{2n+1} in Y so that $a_0, ..., a_{2n}, a_{2n+1}$ and $b_0, ..., b_{2n}, b_{2n+1}$ have the same type. [Possible by Lemma 1.18, as $(Y, <^Y)$ is dense.] Now let $u \in \mathbb{N}$ be the smallest natural number so that y_u is not one of $\{b_0, ..., b_{2n}, b_{2n+1}\}$. Define $b_{2n+2} = y_u$. Find some a_{2n+2} in X so that $b_0, ..., b_{2n}, b_{2n+1}, b_{2n+2}$ and $a_0, ..., a_{2n}, a_{2n+1}, a_{2n+2}$ have the same type. [Possible by Lemma 1.18, as $(X, <^X)$ is dense.]

Continuing this "back and forth" process indefinitely, we end up with infinite sequences $a_0, a_1, a_2, ..., b_0, b_1, b_2, ...$ so that

- for any k, x_k appears in $\{a_0, ..., a_{2k}\}$ and y_k appears in $\{b_0, ..., b_{2k+2}\}$;
- for any k, the sequences $a_0, ..., a_{2k}$ and $b_0, ..., b_{2k}$ have the same type.

It now follows that the function $f\colon X \to Y$ defined by $f(a_i) = b_i$ is defined on all members of X, is onto, and is order-preserving:

$$x <^X x' \iff f(x) <^Y f(x') \text{ for any } x, x' \in X.$$

We are done, by the following exercise.

Exercise 1.19. If f is an order-preserving map between two linear orders, then necessarily f is one-to-one. (Recall that the order is strict.)

and at the end we arrive at a function defined on all of X, with all of Y in the image, which respects the order. □

1.6. **Propositional logic.** Please look over Chapter 1 parts 1.1 and 1.2 of [Enderton] for the basic definitions regarding propositional logic and connectives. I suspect most of you will find it familiar (in essence, if not in notation), and easy to understand. You can also find these in [Woodin-Slaman, 1.1 and 1.2]. (As usual, please ask if you have any questions.)

In particular recall the commonly used connectives $\neg$, $\wedge$, $\vee$, $\to$, where given propositions A and B (statements which could be true or false)

- $\neg A$ means "not A";
- $A \wedge B$ means "A and B";
- $A \vee B$ means "A or B";
- $A \to B$ means "A implies B".

We will also use the symbol "$\perp$" represent the statement "*False*".

2. First order logic

2.1. **The language.** The basic building blocks for our language are the following.

Non-logical symbols. The **alphabet** $\mathcal{A}$ is a collection of **relation symbols** and **function symbols** ($\mathcal{A}$ may be infinite). Each relation symbol R in $\mathcal{A}$ comes with a fixed arity n, in which case we say that R is an "n-ary" relation. Similarly, each function symbol F in $\mathcal{A}$ has a fixed arity.

If R is a 0-ary relation symbol, you may think of it as a proposition, either true or false.

If F is a 0-ary function symbol, we think of it as a **constant symbol**. That is, the function represented by F only spits out one value, so this function simply corresponds to a "constant symbol" for this value.

Terminology: The alphabet is often referred to as the **Vocabulary**, or the **Signature**.

Logical symbols.

- We will use the symbols (and) to parse formulas
- **Logical connectives**: $\neg$ $\to$ $\wedge$ $\vee$;
- **Quantifiers**: $\exists$ $\forall$;
- We will have an **equality symbol** $\approx$ (which is to be always interpreted as a binary relation representing true equality);
- We use **variable** x, y, z,... x_0, x_1,... etc. (Technically, we should fixed some infinite set of variable in advance and only use those, but as customary, in different situations one likes to use different symbols for the variables.)

The variables $x, y, z, \ldots$ are meant to represent some members of our structure. Similarly, the constant symbols in $\mathcal{A}$ represent some members of our structure.

Definition 2.1 (Terms)**.** The **Terms** in the language are defined (recursively) as follows:

- Each variable is a term;
- Each constant symbol is a term;
- Given terms $t_1, \ldots, t_n$ and an n-ary function symbol F then $t = F(t_1, \ldots, t_n)$ is a term.

Example 2.2. Consider two binary function symbols $+$ and $\cdot$. Then the term "$(x+y)^2$" in our formal language is denoted by $\cdot(+(x,y),+(x,y))$.

The term for "$(x+y)+z$" will be $+(+(x,y),z)$. Note that this is *not* the same as the terms $+(z,+(x,y))$ or $+(x,+(y,z))$. These are just (different) formal strings of symbols. The "identification" between them only happens under certain assumptions (axioms of commutatity / associativity).

We will not actually write things like that often...

Remark 2.3. The important thing is that given a term t of the form $t = F(t_1, \ldots, t_n)$, we can find F, $t_1, \ldots, t_n$, just by looking at t.

The terms, like the variables, are supposed to represent members of our structures. We get formulas by plugging terms into relatoins. That is, asking whether the terms satisfy the relation.

Definition 2.4 (Atomic formulas)**.** An **atomic formula** is an expression of the form $P(t_1, \ldots, t_n)$ where P is an n-ary relation symbol in $\mathcal{A}$ (or the relation $\approx$) and $t_1, \ldots, t_n$ are terms.

Example 2.5. Consider two binary function symbols $+$ and $\cdot$ and a constant symbol "1". Then $\approx(\cdot(x,x),+(1,1))$ is an atomic formula, with the intended meaning "$x^2 = 2$".

Consider one binary relation symbol E. Then for any two variables x, y, $E(x,y)$ is an atomic formula. (Before we wrote it as $x \; E \; y$.)

Remark 2.6. We will often (as normal humans do) use short-hand notations. For example, if we use the symbols $+, \cdot, 0, 1$, we may write "2" to be understood as $+(1,1)$. If we also are working in some associative number system (that is, we assume that $(x+y)+z = x+(y+z)$ for any x, y, z), then we may freely write "3" to be understood as $+(+(1,1),1)$ (in this case it would be the same as $+(1,+(1,1))$. Similarly we may simply write $x+y+z$ instead of $+(x,+(y,z))$ or $+(+(x,y),z)$.

As long as we know what we mean, are what we are doing, there is no problem.

Note that atomic formulas can have *free variables.* For example $\Phi =$"$x > y$" is a formula, where x, y are variables and $>$ is a binary relation. If we also have constant symbols $1, 0$, then $\Psi =$"$1 > 0$" is a formula, with no free variables.

The point is: it does not make sense to ask whether the formula Φ is true or not, in a given structure (say the reals $(\mathbb{R}, <, 0, 1)$). It depends of course on the values for x and y. However, the interpretation of the constant symbols $0, 1$ will be part of our structure, so whether Ψ is true or not does have an answer in any given structure (it is true in the usual $(\mathbb{R}, <, 0, 1)$). We will say that Ψ is a **sentence**, while Ψ is a **formula**.

Definition 2.7 (Formulas)**.** The **Formulas** in the language are defined (recursively) as follows. At the same time, we also define what the **free variables** of a formula are

- Any atomic formula is a formula. It's free variables are defined to be all the variables appearing in it.
- If φ is a formula, then $\neg\varphi$ is a formula. The free variables of $\neg\varphi$ are the free variables of φ.
- If φ_1, φ_2 are both formulas, then $(\varphi_1 \wedge \varphi_2)$, $(\varphi_1 \vee \varphi_2)$, and $(\varphi_1 \to \varphi_2)$ are formulas. In each case the free variables of the formula consists of the free variables of φ_1 together with the free variables of φ_2.
- If φ is a formula, x is a variable, then $(\exists x)\varphi$ and $(\forall x)\varphi$ are formulas. The free variables are the free variables of φ *excluding* x.
- We also declare that $\perp$ is a formula, with no free variables.

Example 2.8.
- $(<(x, y) \wedge <(y, z))$ is a formula with 3 free variables, x, y, z. (Intended meaning: $x < y$ and $y < z$.)
- $(\exists x) < (0, x)$ is a formula with no free variables. (Intended meaning: there is some x with $0 < x$.)

Definition 2.9 (Sentences)**.** A **sentence** is a formula with no free variables.

Example 2.10. Consider the language of of set theory, with one binary symbol $\in$ (with inteded meaning "membership").

- $\in(x, y)$ is an atomic formula with free variables $\{x, y\}$ (intended meaning: x is a member of y, we usually write: $x \in y$);
- $(\exists x)(x \in y)$ is a formula with one free variable y (intended meaning: y is not an empty set);
- $\neg(\exists x)(x \in y)$ is a formula with one free variable y (intended meaning: y is the empty set);
- $(\exists y)\neg(\exists x)(x \in y)$ is a sentence (intended meaning: the empty set exists).

Exercise 2.11. Consider the informal formulas we wrote in Example 1.12. Write these formally in the language for the alphabet $\mathcal{A} = \{<\}$.

Example 2.12. Consider the binary operations $+, \cdot$, a binary relation $<$, and a constant "1", and assume that are talking about some number system, such as the real number.

Can we express, in our language, the statement "$\sqrt{2} < \frac{3}{2}$"?

2 and 3 are short-hand notations for specific terms in the language, so that's not the problem. We do not have a division operation, but that is easy to fix: we will consider instead "$2 \cdot \sqrt{2} < 3$", or "$\sqrt{2} + \sqrt{2} < 3$".

However, there is no term in our language that can capture $\sqrt{2}$. (This is not just a formal statement. Even if we take the true standard interpretation of things as real numbers, there is no term which will be interpreted as $\sqrt{2}$.) Nonetheless, we *can* still capture this expression in our language, using the power of quantifiers.

One example would be: "$(\exists x)((x^2 = 2) \wedge ((x + x) < 3))$ ". (Here x^2 is short-hand for $\cdot(x, x)$, etc...)

Another way would be: "$(\forall x)((x^2 = 2) \to (x + x < 3))$"

2.2. **Structures.** Fix a signature $\mathcal{A}$ (a collection of relations symbols and function symbols). (Sometimes we will call the signature "the language", identifying it with the language created from it as above.

Definition 2.13 (Structures)**.** A **structure** $\mathcal{A}$ for the signature $\mathcal{S}$ consists of the following information.

- A set A ("the universe (or domain) of the structure $\mathcal{A}$".
- for any n-ary relation symbol R in the signature S, an n-place relation $R^{\mathcal{A}}$ on A. That is $R^{\mathcal{A}}$ is a subset of A^n. We call $R^{\mathcal{A}}$ the interpretation of R in the structure $\mathcal{A}$.
- For any n-ary function symbol F in $\mathcal{S}$, an n-place function $F^{\mathcal{A}}\colon A^n \to A$. We call $F^{\mathcal{A}}$ the interpretation of F in the structure $\mathcal{A}$.
- The equality symbol $\approx$ is always interpreted as true equality, $\approx^{\mathcal{A}} = \{(a,b) \in A^2 : a = b\}$.

Given an n-place relation symbol R, for any n-many members of A, $a_1, ..., a_n$, we may ask whether they *satisfy the relation*, that is, whether $(a_1, ..., a_n)$ is a member of the subset $R^{\mathcal{A}}$. We will often say "$R^{\mathcal{A}}(a_1, ..., a_n)$ holds" to mean that, and "$R^{\mathcal{A}}(a_1, ..., a_n)$ fails" to mean $(a_1, ..., a_n) \notin R^{\mathcal{A}}$. When R is binary we often use $a_1 \ R^{\mathcal{A}} \ a_2$ instead.

For a 0-place relation symbol P, its interpretation $P^{\mathcal{A}}$ is either "true" or "false". (That is, it is just a predicate.)

For a 0-place function symbol c, its its interpretation $c^{\mathcal{A}}$ is a function with 0 variables as input, with output in A. That is, it just has one output in A. We identify $c^{\mathcal{A}}$ as this single member of A.

The examples of orders and graphs we have seen are all structure for the signature containing one binary relation symbol. A group is a structure for the language containing one binary function symbol $\cdot$ and one constant symbol e. A field can be seen as a structure in the language $\{+, \cdot, 0, 1\}$, two binary function symbols and two constant symbols.

Example 2.14. Unlike these examples, sometimes the question of which vocabulary to choose in order to formally present our structures can be more subtle, and there could be different ways.

Recall that a vector space (over the reals[1].) is a set V ("of vectors", for example $\mathbb{R}^3$) with addition and subtraction operations $+, -$ between vectors in V, as well as *scalar* multiplication: for $v \in V$ and $\alpha \in \mathbb{R}$ we have a vector $\alpha \cdot v \in V$.

Consider the vocabulary $\mathcal{A}$ consisting of the binary function symbols $+, -$ as well as a constant symbol $\bar{0}$, and (infinitely many) unary function symbols f_α for each real $\alpha \in \mathbb{R}$. The intended interpretation would be for $+, -$ to be the addition and subtraction, $\bar{0}$ to be the zero vector, and each f_α to be the function taking a vector v to $\alpha \cdot v$.

2.3. Interpreting the language in a structure.

Example 2.15. Consider the signature $\{s, \underline{0}\}$ where s is a 1-place function symbol and $\underline{0}$ is a constant symbol (a 0-place function symbol). Define a structure $\mathcal{A}$ as follows. The universe A is the set of integers $A = \mathbb{N} = \{0, 1, 2, 3, ...\}$. The interpretation of s is the function $s^{\mathcal{A}}\colon \mathbb{N} \to \mathbb{N}$ defined by $s^{\mathcal{A}}(n) = n+1$ (the successor function). The interpretation of $\underline{0}$ is $\underline{0}^{\mathcal{A}} = 0$.

$t = s(s(x))$ is a term in this language with one variable x. Its intended interpretation should be the function sending n to $n+2$. $t = s(s(0))$ is a term with no variables (term for a constant). Its intended interpretation is 2.

---

[1]You can do the same for vector spaces over some other field, for example $\mathbb{Q}$ or $\mathbb{C}$

Let t be a term and $\bar{x} = x_1, ..., x_n$ a list *which includes all variables appearing in t*. In this case we sometimes write $t(x_1, ..., x_n)$ for t. ("t is a term whose value depends on $x_1, ..., x_n$".) There is a minor abuse of notation here. For example, if t is simply the variable x, we may thing of t as $t(x)$, but also as $t(x, y)$. That is, we may always add "dummy variables".

Definition 2.16 (Interpretation of terms)**.** Fix a signature $\mathcal{S}$ and a structure $\mathcal{A}$. For a term t and variables $x_1, ..., x_n$ which include all variables appearing in t, we define **the realization** of t in $\mathcal{A}$ to be a function $t^{\mathcal{A}}(x_1, ..., x_n)$ from $A^n \to A$ as follows, recursively along the construction of terms.

[Case 1] t is a variable. Then t is x_i for some $i \in \{1, ..., n\}$. Then $t^{\mathcal{A}}(\bar{a}) = a_i$ for any $\bar{a} = (a_1, ..., a_n)$ in A^n.

[Case 2] t is $F(t_1, ..., t_k)$ for some k-ary function symbol F and some terms $t_1, ..., t_k$. Define

$$t^{\mathcal{A}}(\bar{a}) = F^{\mathcal{A}}(t_1^{\mathcal{A}}(\bar{a}), ..., t_k^{\mathcal{A}}(\bar{a})).$$

Remark 2.17. In Case 2, $t_j^{\mathcal{A}}$ is already defined inductively. Note also that we are using the fact that if $x_1, ..., x_n$ includes all the variables appearing in t, and t is of the form $F(t_1, ..., t_k)$, then also $x_1, ..., x_n$ includes all the variables appearing in t_j for each $j \in \{1, ..., k\}$. This does require a proof. We skip it here. You can find a more careful analysis of the syntax in [Woodin-Slaman, Chapter 2] or in [Enderton].

Similarly, in order to carry this recursive definition, we need to know that given t we can *uniquely* identify "where it came from", meaning finding the F, $t_1, ..., t_k$ for which $t = F(t_1, ..., t_k)$. The is not difficult to see here. (Similar facts are true for the construction of formulas.) Again we skip this syntax analysis here, and you can see more in [Woodin-Slaman, Chapter 2] or in [Enderton]. See in particular "unique readability" results.

Let us emphasize that the "unique readability" of this particular way of coding terms / formulas, is not important, but just that there is some reasonable way of doing so. For example, if you feel more comfortable adding some parenthesis, commas, semicolons, here and there, why not... [2]

Remark 2.18. Suppose c is a constant symbol in the signature (a 0-place function). Then c is a term. For any variables $x_1, ..., x_n$, we may view c as a term $c(x_1, ..., x_n)$. What is the function $c^{\mathcal{A}}(x_1, ..., x_n)\colon A^n \to A$? It is the constant function: $c^{\mathcal{A}}(\bar{a}) = c^{\mathcal{A}}$ for any $\bar{a}$ in A^n.

Example 2.19. Consider $\mathcal{A} = (\mathbb{R}, +, \cdot)$ (standard operations). Consider the term t "$x \cdot y + x$". It should be interpreted as a function from $\mathbb{R}^2 \to \mathbb{R}$. It is clear of course what the interpretation here is, but let us follow the definitions. Formally our term is $t = +(\cdot(x, y), x)$.

As the interpretation is done recursively along the construction, we must consider the whole construction of t.

Let t_1 be the variable x.

Let t_2 be the variable y.

[2] One can be more extreme, and replace each terms (and similarly formulas) with their entire sequence of construction. That is, instead of $\cdot(+(x, y), 1)$ we would use the "term" $\langle x; y; 1; +(x, y); \cdot(+(x, y), 1)\rangle$. You may also add some more information to the sequence coding the "instructions" of how to construct a term from previous terms in the sequence. This way, there is certainly no ambiguity on how the "definition along the construction" is done.

Let t_3 be $\cdot(t_1, t_2)$.
Let $t = t_4$ be $+(t_3, t_1)$
The interpretation of t_1 is $t_1^{\mathcal{A}}\colon \mathbb{R}^2 \to \mathbb{R}$, $t_1^{\mathcal{A}}(a,b) = a$. Similarly, $t_2^{\mathcal{A}}(a,b) = b$.
Now $t_3^{\mathcal{A}}(a,b) = \cdot^{\mathcal{A}}(t_1^{\mathcal{A}}(a,b), t_2^{\mathcal{A}}(a,b)) = \cdot^{\mathcal{A}}(a,b) = a \cdot b$.
$t^{\mathcal{A}}(a,b) = +^{\mathcal{A}}(t_3^{\mathcal{A}}(a,b), t_1^{\mathcal{A}}(a,b)) = +^{\mathcal{A}}(a \cdot b, a) = a \cdot b + a$.

Finally, we define the interpretation of formulas in a structure. For a formula $\varphi(x_1, ..., x_n)$ its interpretation in a structure $\mathcal{A}$ will be a function $\varphi^{\mathcal{A}}(x_1, ..., x_n)$ from A^n to $\{0,1\}$. For $\bar{a} = a_1, ..., a_n$ from A, either $\varphi^{\mathcal{A}}(\bar{a}) = 1$ (True) or $\varphi^{\mathcal{A}}(\bar{a}) = 0$ (False). In other words, the interpretation of $\varphi(x_1, ..., x_n)$ in $\mathcal{A}$ is an n-ary predicate. We sometimes say "$\varphi^{\mathcal{A}}(a_1, ..., a_n)$ holds" to mean that $\varphi^{\mathcal{A}}(a_1, ..., a_n) = 1$. Similarly, we sometimes identify $\varphi^{\mathcal{A}}$ as the subset of A^n, $\{\bar{a} \in A^n : \varphi(\bar{a}) = 1\}$.

Example 2.20. Consider the signature $\{s, \underline{0}\}$ where s is a 1-place function symbol and $\underline{0}$ is a constant symbol (a 0-place function symbol). Define the structure $\mathcal{A} = (\mathbb{N}, s^{\mathcal{A}}, 0)$ where $s^{\mathcal{A}}(n) = n + 1$ the successor, and $\underline{0}^{\mathcal{A}} = 0$.

The variables x, y are terms. $s(y)$ is a term. So $\varphi(x,y) = (x \approx s(y))$ is an atomic formula (with free variables x, y). For $n, m \in \mathbb{N}$, the interpretation $\varphi(n,m)$ is true if and only if $n = m + 1$.

$\psi(x) = (\exists y)\varphi(x,y)$ is a formula (with free variable x). For $n \in \mathbb{N}$, the interpretation $\psi(n)$ is true if and only if $n > 0$.

$\theta = (\forall x)\psi(x)$ is a formula with no free variables. Its interpretation is False in the structure $\mathcal{A}$.

Given a formula φ, and variables $\bar{x} = x_1, ..., x_n$, we write $\varphi(\bar{x})$ only when the list $x_1, ..., x_n$ includes all free variables of φ (possible more, "dummy variables").

Definition 2.21 (Interpretation of atomic formulas)**.** Fix a signature $\mathcal{S}$. Let $\varphi = R(t_1, ..., t_n)$ where R is an n-ary relation symbol in the language, $t_1, ..., t_n$ are terms. Let $x_1, ..., x_n$ be variables containing all variables appearing in φ. For $\bar{a} = a_1, ..., a_n$ in A, define

$$\varphi^{\mathcal{A}}(a_1, ..., a_n) = \begin{cases} 1\,(True) & (t_1^{\mathcal{A}}(\bar{a}), ..., t_n^{\mathcal{A}}(\bar{a})) \in R^{\mathcal{A}} \ (R^{\mathcal{A}}(t_1^{\mathcal{A}}(\bar{a}), ..., t_n^{\mathcal{A}}(\bar{a})) \text{ holds}); \\ 0\,(False) & (t_1^{\mathcal{A}}(\bar{a}), ..., t_n^{\mathcal{A}}(\bar{a})) \notin R^{\mathcal{A}} \ (R^{\mathcal{A}}(t_1^{\mathcal{A}}(\bar{a}), ..., t_n^{\mathcal{A}}(\bar{a})) \text{ fails}); \end{cases}$$

Technically we should write $\varphi^{\mathcal{A}}(x_1, ..., x_n)(a_1, ..., a_n)$, but we will omit the sequence $x_1, ..., x_n$ from the notation if it is clear from context.

Remark 2.22. The interpretations of relation symbols in the vocabulary $\mathcal{S}$ depend on the structure. Our equality symbol $\approx$ however is *always interpreted as true equality* (it is not up to the structure to interpret it).

We continue to define the interpretation of formulas in a structure $\mathcal{A}$, recursively along the construction of formulas. It will be notationally convenient to view $\varphi^{\mathcal{A}}(x_1, ..., x_n)$ as a subset of A^n. (The subset is all $(a_1, ..., a_n)$ for which $\varphi^{\mathcal{A}}(a_1, ..., a_n) = 1$.)

Logical connectives.

Assume $\varphi = \neg\psi$. The if $\bar{x} = x_1, ..., x_n$ includes all free variables of φ, it includes all free variables of ψ, and so (recursively) $\psi^{\mathcal{A}}(x_1, ..., x_n)$ is already defined. Define

$$\varphi^{\mathcal{A}}(x_1, ..., x_n) = A^n \setminus \psi^{\mathcal{A}}(x_1, ..., x_n). \text{ (The complement.)}$$

If $\varphi = (\psi_1 \wedge \psi_2)$. Define

$$\varphi^{\mathcal{A}}(x_1, ..., x_n) = \psi_1^{\mathcal{A}}(x_1, ..., x_n) \cap \psi_n^{\mathcal{A}}(x_1, ..., x_n). \text{ (The intersection.)}$$

If $\varphi = (\psi_1 \vee \psi_2)$. Define

$$\varphi^{\mathcal{A}}(x_1, ..., x_n) = \psi_1^{\mathcal{A}}(x_1, ..., x_n) \cup \psi_n^{\mathcal{A}}(x_1, ..., x_n). \text{ (The union.)}$$

If $\varphi = (\psi_1 \to \psi_2)$. Define

$$\varphi^{\mathcal{A}}(x_1, ..., x_n) = (A^n \setminus \psi_1^{\mathcal{A}}(x_1, ..., x_n)) \cup \psi_n^{\mathcal{A}}(x_1, ..., x_n).$$

If $\varphi = \perp$. $\varphi^{\mathcal{A}}(x_1, ..., x_n) = \emptyset$ (the empty set).

Quantifiers.

Assume $\varphi = (\exists z)\psi$*, and* *assume that* z *does not appear in* $\bar{x}$. If $\bar{x} = x_1, ..., x_n$ includes all free variables in φ, then $x_1, ..., x_n, z$ includes all free variables appearing in ψ. Define $\varphi^{\mathcal{A}}(a_1, ..., a_n) = 1$ if and only if there is some $c \in A$ so that $\psi^{\mathcal{A}}(a_1, ..., a_n, c) = 1$.

Note that $\varphi^{\mathcal{A}}(x_1, ..., x_n)$ as a subset of A^n is the projection of $\psi^{\mathcal{A}}(x_1, ..., x_n, z)$ (a subset of $A^n \times A$) to A^n. That is, define $\pi \colon A^n \times A \to A^n$ by $\pi(a_1, ..., a_n, c) = (a_1, ..., a_n)$. Then $(a_1, ..., a_n) \in \varphi^{\mathcal{A}}(x_1, ..., x_n)$ if and only if $(a_1, ..., a_n)$ is in $\pi[\psi^{\mathcal{A}}(x_1, ..., x_n, z)]$ (the image of the set $\psi^{\mathcal{A}}(x_1, ..., x_n, z)$ under the map π).

Remark 2.23. Here, before, and in the future, we use the common "identification" between, for example, $A^n \times A$ and A^{n+1}. Recall that all we mean by A^n is an (ordered) sequence of n members from A. How such sequences are formally coded, or what not, is not important. In particular, we identify (in the natural way) $A^n \times A^m$, which is formally a pair of sequences, one of length n and the other of length m, with A^{n+m}, a single sequence of length $n + m$.

2.3.1. *Assume* $\varphi = (\forall z)\psi$*, and assume that* z *does not appear in* $\bar{x}$. if $\bar{x} = x_1, ..., x_n$ includes all free variables in φ, then $x_1, ..., x_n, z$ include all free variables appearing in ψ. Define $\varphi^{\mathcal{A}}(a_1, ..., a_n) = 1$ if and only if there for all c in A, $\psi^{\mathcal{A}}(a_1, ..., a_n, c) = 1$.

Pictorially, viewing $\psi^{\mathcal{A}}(x_1, ..., x_n, z)$ as a subset of $A^n \times A$, $\varphi^{\mathcal{A}}(x_1, ..., x_n)$ are those elements in A^n whose "fiber" $\{c \in A : \psi^{\mathcal{A}}(a_1, ..., a_n, c) = 1\}$ is the entire A.

Another point of view is as a sort of "large conjunction". We may view $\psi^{\mathcal{A}}(x_1, ..., x_n, z)$ as a parametrized family of subsets of A^n: for each $c \in A$ the corresponding set is $\psi^{\mathcal{A}}(x_1, ..., x_n, c) = \{(a_1, ..., a_n) : \psi^{\mathcal{A}}(a_1, ..., a_n, c) = 1\}$, and

$$\varphi^{\mathcal{A}}(x_1, ..., x_n) = \bigcap_{c \in A} \psi^{\mathcal{A}}(x_1, ..., x_n, c).$$

Remark 2.24 (Sentences)**.** If φ is a sentence (formula with no free variables), we can interpret it with no variables at all, and simply get a truth value "1" (true) or "0" (false).

Example 2.25. Consider the structure $(\mathbb{R}, +, \cdot, 0, 1)$ with the usual interpretation. Consider the sentence

$$\varphi = (\forall y)(\exists x)(x \cdot x \approx y).$$

To interpret it we need its "entire construction". Begin with the term $t_1 = x \cdot x$ (formally $\cdot(x, x)$) whose interpretation, as a function of (x, y), is the function on $\mathbb{R}^2$ sending x, y to x^2. Similarly, the interpretation of $t_2 = y$ is the function taking x, y to y.

Next consider the atomic formula $\psi(x,y) = (x \cdot x \approx y)$. Its interpretation in our structure is the set of all pairs of reals (x,y) so that $y = x^2$. That is, a parabola.

Let $\theta(y) = (\exists x)\psi$. Its interpretation is a subset of $\mathbb{R}$, specifically all reals b for which $b = x^2$ for some x. That is, all non-negative reals.

Finally, $\varphi = (\forall y)\theta$, is false, since the interpretation of $\theta(y)$ is not the entire structure $\mathbb{R}$.

A minor headache. What if the quantified-over variable z appears among the "free looking" sequence of variables $\bar{x}$? Let us first see why this interrupts our intended interpretation.

Consider $\theta = (\exists x)(x \cdot x \approx y)$, where $\psi = (x \cdot x \approx y)$ as above. Formally, we may consider $\theta(y,x)$, as y, x is a list including all the free variables of θ (which is just y).

If we were to try and follow the above definition, the interpretation of $\theta(y,x)$ will be the projection of $\psi(y,x)$, which is a subset of $\mathbb{R}$ (rather than $\mathbb{R}^2$). However, the *intended meaning of $\theta(y,x)$ is really "y is the square of something", without any mention of x!* That is, there should be no difference between

$$\theta(y,x) = (\exists x)(x \cdot x \approx y), \text{ and } \theta'(y,x) = (\exists z)(z \cdot z \approx y).$$

Note that $\theta'(y,x)$ does fall into the "normal" category. That is, it is a formula with one free variable y, y, x is a list of two distinct variables, and *non of them is the quantified-over variable z*. So the interpretation of θ' is defined at this point. We will simply define the interpretation of θ as the one for θ'. That is, $\theta^{\mathcal{A}}(y,x)$ is the set of all (b,a) in $\mathbb{R}^2$ so that $\theta^{\mathcal{A}}(y)(b) = 1$. In this case, this is the set of all (b,a) so that $b \geq 0$ (and a in $\mathbb{R}$ arbitrary).

More formally: suppose $\varphi = (\exists z)\psi$ and $\bar{x} = x_1, ..., x_n$ is a sequence of variables which *does* contain z. For convenience, assume that $x_n = z$. Then z is *not* a free variable of φ. In particular, the list $x_1, ..., x_{n-1}$ *is* a list containing all free variables of φ, and this list does not contain the quantified-over x_n. So, as above, we have already defined the interpretation $\varphi^{\mathcal{A}}(x_1, ..., x_{n-1}) \subseteq A^{n-1}$. Finally define $\varphi^{\mathcal{A}}(x_1, ..., x_n) = \varphi^{\mathcal{A}}(x_1, ..., x_{n-1}) \times A$ (a subset of A^n, as required).

A final remark: you may avoid this nonsense by using different variables (as in θ' above).

Exercise 2.26. Consider $\varphi = (\exists x)((x < 0) \wedge (\forall x)(x \cdot x \geq 0))$. What is its truth value in $\mathbb{R}$ (with the usual interpretation of the symbols)?

Notation 2.27. Given a structure $\mathcal{A}$, a formula $\varphi(x_1, ..., x_n)$ and $a_1, ..., a_n$ from A, we will write

$$\mathcal{A} \models \varphi(a_1, ..., a_n) \iff \varphi^{\mathcal{A}}(a_1, ..., a_n) = 1,$$

in which case we say that $\varphi(a_1, ..., a_n)$ is "satisfied". (The symbol $\models$ is to be read, as a verb, "models". I think.) Of particular importance are the sentences φ, for which we may ask if they are satisfied by $\mathcal{A}$, or if "$\mathcal{A}$ models φ", or if "φ is true in $\mathcal{A}$".

2.4. **Morphisms and formulas.**

Definition 2.28. Let $\mathcal{A}$, $\mathcal{B}$ be two structures for the same signature $\mathcal{S}$. Let $f \colon A \to B$ be a function from the domain of $\mathcal{A}$ to the domain of $\mathcal{B}$.

(1) Say that f is a **homomorphism**[3] if:

[3]Warning: this term may mean different things in different sources.

- for any n-ary relation symbol R in $\mathcal{S}$, for any $a_1, ..., a_n$ from A, *if* $R^{\mathcal{A}}(a_1, ..., a_n)$ *then* $R^{\mathcal{B}}(f(a_1), ..., f(a_n))$;
- for any n-ary function symbol F in $\mathcal{S}$, for any $a_1, ..., a_n, a_{n+1}$ from A, *if* $F^{\mathcal{A}}(a_1, ..., a_n) = a_{n+1}$ *then* $F^{\mathcal{B}}(f(a_1), ..., f(a_n)) = f(a_{n+1})$. [In other words, $f(F^{\mathcal{A}}(a_1, ..., a_n)) = F^{\mathcal{B}}(f(a_1), ..., f(a_n))$.]

(2) Say that f is an **embedding** if it is one-to-one and
- for any n-ary relation symbol R in $\mathcal{S}$, for any $a_1, ..., a_n$ from A, $R^{\mathcal{A}}(a_1, ..., a_n)$ *if and only if* $R^{\mathcal{B}}(f(a_1), ..., f(a_n))$;
- for any n-ary function symbol F in $\mathcal{S}$, for any $a_1, ..., a_n, a_{n+1}$ from A, *if* $F^{\mathcal{A}}(a_1, ..., a_n) = a_{n+1}$ *then* $F^{\mathcal{B}}(f(a_1), ..., f(a_n)) = f(a_{n+1})$. [Same as for a homomorphism.]

(3) Say that f is an **isomorphism** if it is an embedding and is onto.

Remark 2.29. (1) Being one-to-one is the same as the condition for being an embedding with the relation symbol $\approx$.
(2) If there are only function symbols in the vocabulary, then a one-to-one homomorphism is an embedding. Generally, a one-to-one homomorphism may not be an embedding.

Exercise 2.30. Let $f: A \to B$ be a homomorphism from $\mathcal{A}$ to $\mathcal{B}$. Show that the following are equivalent.
- $f: A \to B$ is an isomorphism from $\mathcal{A}$ to $\mathcal{B}$.
- There is a $g: B \to A$ which is a homomorphism from $\mathcal{B}$ to $\mathcal{A}$ and $f \circ g = id_B$ and $g \circ f = id_A$. [Here id_A is the identity map from A to A.]

Lemma 2.31. Let $h: \mathcal{A} \to \mathcal{B}$ be a homomorphism between two structures of the same signature. Then for any term $t(x_1, ..., x_n)$ and any $a_1, ..., a_n$ from A,
$$h(t^{\mathcal{A}}(a_1, ..., a_n)) = t^{\mathcal{B}}(h(a_1), ..., h(a_n)).$$

Proof. We prove this by induction along the construction of the terms.

Case 1: if t is a variable x_i, then $t^{\mathcal{A}}(a_1, ..., a_n) = a_i$ and $t^{\mathcal{B}}(h(a_1), ..., h(a_n)) = h(a_i)$.

Case 2: $t = F(t_1, ..., t_k)$. Let $\bar{a} = a_1, ..., a_n$ and $\bar{b} = h(a_1), ..., h(a_n)$. Then $h(t^{\mathcal{A}}(\bar{a})) = h(F^{\mathcal{A}}(t_1^{\mathcal{A}}(\bar{a}), ..., t_k^{\mathcal{A}}(\bar{a})))$ by the definition of term evaluation. The latter is equal to $F^{\mathcal{B}}(h(t_1^{\mathcal{A}}(\bar{a})), ..., h(t_k^{\mathcal{A}}(\bar{a})))$ since h is a homomorphism. By the inductive assumption, $h(t_i^{\mathcal{A}}(\bar{a})) = t_i^{\mathcal{B}}(\bar{b})$ so the latter expression is equal to $F^{\mathcal{B}}(t_1^{\mathcal{B}}(\bar{b}), ..., t_k^{\mathcal{B}}(\bar{b}))$, which is $t^{\mathcal{B}}(\bar{b})$, again by the definition of term evaluation. $\square$

Say that two structures $\mathcal{A}$ and $\mathcal{B}$ are **isomorphic** if there is an isomorphism from $\mathcal{A}$ to $\mathcal{B}$. In this case we write $\mathcal{A} \simeq \mathcal{B}$.

Exercise 2.32. Show that "isomorphism" is an equivalence relation on structures (of a fixed signature $\mathcal{S}$). That is:
- Every structure $\mathcal{A}$ is isomorphic to itself;
- If $\mathcal{A}$ is isomorphic to $\mathcal{B}$ then $\mathcal{B}$ is isomorphic to $\mathcal{A}$;
- If $\mathcal{A}$ is isomorphic to $\mathcal{B}$ and $\mathcal{B}$ is isomorphic to $\mathcal{C}$ then $\mathcal{A}$ is isomorphic to $\mathcal{C}$.

Recall that if two structures $\mathcal{A}$ and $\mathcal{B}$ are isomorphic, that should mean they are really, essentially, the same thing. For example, given a sentence φ, it better be that $\mathcal{A}$ and $\mathcal{B}$ agree on whether φ is true or false, if they are isomorphic.

Theorem 2.33. Let $\mathcal{A}$ and $\mathcal{B}$ be structures for the same signature $\mathcal{S}$. Suppose that $f\colon A \to B$ is an isomorphism of $\mathcal{A}$ and $\mathcal{B}$. Let $\varphi(x_1, ..., x_n)$ be a formula. Then

$$(\star) \qquad \varphi^{\mathcal{A}}(a_1, ..., a_n) = 1 \iff \varphi^{\mathcal{B}}(f(a_1), ..., f(a_n)) = 1, \text{ for any } a_1, ..., a_n \text{ from } A.$$

In particular, if φ is a sentence, then

$$\mathcal{A} \models \varphi \iff \mathcal{B} \models \varphi.$$

Recall that in our first class we showed that $(\mathbb{Q}, <)$ and $(\mathbb{N}, <)$ are not isomorphic. Our proof back then can be seen as follows. The sentence $\varphi = (\exists x)(\forall y)((x \approx y) \vee x < y)$ is true in $(\mathbb{N}, <)$ but false in $(\mathbb{Q}, <)$. So they cannot be isomorphic.

Proof. Recall that we write $\varphi(x_1, ..., x_n)$ only when $x_1, ..., x_n$ is a list of distinct variables which includes all free variables of φ.

The proof will proceed inductively along the construction of formulas.

By Lemma 2.31 we already know that $f(t^{\mathcal{A}}(a_1, ..., a_n)) = t^{\mathcal{B}}(f(a_1), ..., f(a_n))$ for any $a_1, ..., a_n$ in A and any term t.

Start with atomic formulas. Let φ be of the form $R(t_1, ..., t_k)$ where $t_1, ..., t_k$ are terms and R is an k-ary relation. Let $x_1, ..., x_n$ be a list of variables including all the variables appearing in φ. Fix some $a_1, ..., a_n$ from A. Let $d_i = t_i^{\mathcal{A}}(a_1, ..., a_n)$ for $i = 1, ..., k$.
Then, by definition, $\varphi^{\mathcal{A}}(a_1, ..., a_n) = 1$ if and only if $(d_1, ..., d_k) \in R^{\mathcal{A}}$.
Similarly, let $e_i = t_i^{\mathcal{B}}(f(a_1), ..., f(a_n))$.
Then, by definition, $\varphi^{\mathcal{B}}(f(a_1), ..., f(a_n)) = 1$ if and only if $(e_1, ..., e_k) \in R^{\mathcal{B}}$.
We know that $e_i = f(d_i)$. Finally, since f is an isomorphism, $(d_1, ..., d_k) \in \mathbb{R}^{\mathcal{A}}$ if and only if $(e_1, ..., e_k) \in \mathbb{R}^{\mathcal{B}}$, which concludes the proof of $(\star)$ for the atomic formula φ.

Next we consider connectives. Assume $(\star)$ is true for ψ, and show it is true for $\varphi = \neg\psi$. Indeed, $\varphi^{\mathcal{A}}(a_1, ..., a_n) = 1$ if and only if $\psi^{\mathcal{A}}(a_1, ..., a_n) = 0$ if and only if (inductive assumption) $\psi^{\mathcal{B}}(f(a_1), ..., f(a_n)) = 0$, if and only if $\varphi^{\mathcal{B}}(f(a_1), ..., f(a_n)) = 1$. So $(\star)$ is true for φ.

Assume now that $(\star)$ is true for ψ_1 and ψ_2, and show that it is true for $\varphi = (\psi_1 \wedge \psi_2)$.
Then $\varphi^{\mathcal{A}}(a_1, ..., a_n) = 1$ iff both $\psi_1^{\mathcal{A}}(a_1, ..., a_n) = 1$ and $\psi_2^{\mathcal{A}}(a_1, ..., a_n)$.
By the inductive assumption, for each i, $\psi_i^{\mathcal{A}}(a_1, ..., a_n) = 1$ iff $\psi_i^{\mathcal{B}}(f(a_1), ..., f(a_n)) = 1$.
So $\varphi^{\mathcal{A}}(a_1, ..., a_n) = 1$ iff both $\psi_1^{\mathcal{B}}(f(a_1), ..., f(a_n)) = 1$ and $\psi_2^{\mathcal{B}}(f(a_1), ..., f(a_n)) = 1$, which is true iff $\varphi^{\mathcal{B}}(f(a_1), ..., f(a_n)) = 1$. This concludes the proof of $(\star)$ for φ.

The cases where $\varphi = (\psi_1 \to \psi_2)$ or $\varphi = (\psi_1 \vee \psi_2)$ are extremely similar, and are left for you to complete. These can also be skipped using some "logical equivalences" [see Pset 3]. For example, $(\psi_1 \vee \psi_2)$ is equivalent to $\neg(\neg\psi_1 \wedge \neg\psi_2)$, and $(\psi_1 \to \psi_2)$ is equivalent to $(\psi_2 \vee \neg\psi_1)$.

You may be a bit bored by this... Indeed not much has been going on, other than repeatedly stating our definitions. Indeed, the *major interesting case is quantification.*
Start with an existential quantifier. That is, assume $(\star)$ is true for ψ, and let φ be of the form $(\exists x)\psi$, and prove that $(\star)$ is true for φ. As usual $x_1, ..., x_n$ is a list of variables containing all free variables in φ.

Assume first that x is not one of $x_1, ..., x_n$. (The normal situation...)
If $\varphi^{\mathcal{A}}(a_1, ..., a_n) = 1$, by the definitions, there is some $a \in A$ so that $\psi^{\mathcal{A}}(a_1, ..., a_n, a) = 1$.
By $(\star)$ for ψ, we know that $\psi^{\mathcal{B}}(f(a_1), ..., f(a_n), f(a)) = 1$.
Again by definition it follows that $\varphi^{\mathcal{B}}(f(a_1), ..., f(a_n)) = 1$.

So we proved the $\implies$ of $(\star)$ for φ.
Assume now $\varphi^{\mathcal{B}}(f(a_1),...,f(a_n)) = 1$. So there is some b in B for which $\psi^{\mathcal{B}}(f(a_1),...,f(a_n),b) = 1$.
Since f is onto, there is some a in A so that $b = f(a)$. So $\psi^{\mathcal{B}}(f(a_1),...,f(a_n),f(a)) = 1$.
Now similarly by $(\star)$ for ψ we get $\psi^{\mathcal{A}}(a_1,...,a_n,a)$, and therefore $\varphi^{\mathcal{A}}(a_1,...,a_n) = 1$, concluding the $\impliedby$ of $(\star)$ for φ.

Assume now that $\varphi = (\forall x)\psi$ and we know $(\star)$ for ψ. Again this case can be avoided using the logical equivalence between $(\forall x)\psi$ and $\neg(\exists x)\neg\psi$ [see Pset 3], but let us repeat the argument for clarity.
If $\varphi^{\mathcal{A}}(a_1,...,a_n) = 1$, then for *any* $a \in A$, $\psi^{\mathcal{A}}(a_1,...,a_n,a) = 1$.
We want to show that $\varphi^{\mathcal{B}}(f(a_1),...,f(a_n)) = 1$. That is, that for any $b \in B$, $\psi^{\mathcal{B}}(f(a_1),...,f(a_n)) = 1$.
Fix some $b \in B$. Since f is onto, there is some $a \in A$ for which $f(a) = b$. As $\psi^{\mathcal{A}}(a_1,...,a_n,a) = 1$, applying $(\star)$ for ψ we conclude that $\psi^{\mathcal{B}}(f(a_1),...,f(a_n),b)$, as required.

On the other hand, assume that $\varphi^{\mathcal{B}}(f(a_1),...,f(a_n)) = 1$. We want to show that $\varphi^{\mathcal{A}}(a_1,...,a_n) = 1$.
Fix an arbitrary $a \in A$. We need to show that $\varphi^{\mathcal{A}}(a_1,...,a_n,a) = 1$. By assumption, and the definition of interpretation for $\forall$, $\psi^{\mathcal{B}}(f(a_1),...,f(a_n),f(a)) = 1$. Finally, by $(\star)$ for ψ, $\psi^{\mathcal{A}}(a_1,...,a_n,a) = 1$, as required.

Finally, let us treat the weird case, where in our list of variables $x_1,...,x_n$ we have the variable over which we just quantified. Let us do this for the existential quantifier only.

For notational convenience, assume without loss of generality that we quantified over the last variable x_n. That is, $\varphi = (\exists x_n)\psi$ and $x_1,...,x_n$ is a list of variables containing all the free variables of φ (as well is x_n which is not a free variable of φ). Assume $(\star)$ holds for ψ, and show it for φ.

Recall that in this case, we may also view φ as $\varphi(x_1,...,x_{n-1})$, and in fact we defined the interpretation of $\varphi(x_1,...,x_n)$ by: $\varphi^{\mathcal{A}}(a_1,...,a_n) = 1 \iff \varphi^{\mathcal{A}}(a_1,...,a_{n-1}) = 1$, for any structure $\mathcal{A}$. [More formally: $\varphi^{\mathcal{A}}(x_1,...,x_n)(a_1,...,a_n) = 1 \iff \varphi^{\mathcal{A}}(x_1,...,x_{n-1})(a_1,...,a_{n-1}) = 1$.]

In conclusion, for any $a_1,...,a_n$ from A, $\varphi^{\mathcal{A}}(a_1,...,a_n) = 1$ iff $\varphi^{\mathcal{A}}(a_1,...,a_{n-1}) = 1$ iff $\varphi^{\mathcal{B}}(f(a_1),...,f(a_{n-1})) = 1$ [as we proved this case above] iff $\varphi^{\mathcal{B}}(f(a_1),...,f(a_n)) = 1$ [again by the "silly quantifier definition"].

□

Definition 2.34 (Elementary equivalence)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be models for the same vocabulary. Say that $\mathcal{A}$ and $\mathcal{B}$ are **elementary equivalent**, denoted $\mathcal{A} \equiv \mathcal{B}$ if for any sentence φ, $\mathcal{A} \models \varphi \iff \mathcal{B} \models \varphi$.

Corollary 2.35. If $\mathcal{A}$ and $\mathcal{B}$ are isomorphic then they are elementary equivalent. In symbols: $\mathcal{A} \simeq \mathcal{B} \implies \mathcal{A} \equiv \mathcal{B}$.

Remark 2.36. Whenever we compare two structures in any way: either $\simeq$ or $\equiv$, or we talk about homomorphisms between them, there is always an implicit assumption that they are structures for the same signature.

3. Substructures and elementary substructures

Definition 3.1 (Substructure)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be structure for some signature $\mathcal{S}$. Say that $\mathcal{B}$ is a **substructure** of $\mathcal{A}$ if:

- $B \subseteq A$ (the universe of $\mathcal{B}$ is a subset of the universe of $\mathcal{A}$);
- given any n-ary relation symbol P in $\mathcal{S}$, for any $b_1, ..., b_n$ in B,
$$R^{\mathcal{B}}(b_1, ..., b_n) \iff R^{\mathcal{A}}(b_1, ..., b_n);$$
- given any n-ary function symbol F in $\mathcal{S}$, for any $b_1, ..., b_n$ in B,
$$F^{\mathcal{B}}(b_1, ..., b_n) = F^{\mathcal{A}}(b_1, ..., b_n).$$

Let us focus first on a *relational language*, that is, when the signature has only relation symbols, for example, linear orders. In this case, a substructure of $\mathcal{A}$ is some subset $B \subseteq A$, where we interpret the relations in $\mathcal{B}$ according to these given by $\mathcal{A}$. In particular, for *any* set $B \subseteq A$, we can interpret the relation symbols according to $\mathcal{A}$, and this will define a substructure. For example, for any subset $B \subseteq \mathbb{Q}$, we may view $(B, <)$ as a linear order, where $<$ is given according to $\mathbb{Q}$.

If there is a constant symbol c, then its interpretation in $\mathcal{B}$ and in $\mathcal{A}$ must be the same. In particular, $c^{\mathcal{A}}$ must be in B. If there are function symbols, then we also want the substructure to be *closed under* these functions. That is, if $B \subseteq A$ is a set, and for any n-ary function symbol F in the language, and any $b_1, ..., b_n$ from B, $F^{\mathcal{A}}(b_1, ..., b_n) \in B$, then we can again make B into a substructure by simply interpreting the symbols according to $\mathcal{A}$. For example, given a vector space over $\mathbb{R}$, in the language described before, a substructure will be precisely a subspace (as is commonly defined for vector spaces).

Remark 3.2. $\mathcal{B}$ is a substructure of $\mathcal{A}$ if and only if the identity map $f\colon B \to A$, defined by $f(b) = b$ for any $b \in B$, is an embedding from $\mathcal{B}$ to $\mathcal{A}$.

Remark 3.3. As always, the language matters. Given some group $(G, \cdot, e)$, a substructure is some set $H \subseteq G$ so that $e \in H$ and H is closed under multiplication. However, a subgroup is assumed to be also closed under inverses. We can view the group in an expanded language $(G, \cdot, \square^{-1}, e)$, where $\square^{-1}$ is an unary function symbol which here is interpreted as sending $g \in G$ to its inverse g^{-1}. In this language, a substructure is precisely a subgroup.

Definition 3.4 (Elementary substructure)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be structures for a signature $\mathcal{S}$. say that $\mathcal{B}$ is an **elementary substructure** of $\mathcal{A}$, denoted $\mathcal{B} \preceq \mathcal{A}$, if $B \subseteq A$ and for any formula $\varphi(x_1, ..., x_n)$ and any $b_1, ..., b_n$ from B,
$$\varphi^{\mathcal{B}}(b_1, ..., b_n) \iff \varphi^{\mathcal{A}}(b_1, ..., b_n).$$

Remark 3.5. By considering only sentences φ, it follows that if $\mathcal{A} \preceq \mathcal{B}$ then $\mathcal{A} \equiv \mathcal{B}$.

If $\mathcal{B} \preceq \mathcal{A}$ then $\mathcal{B}$ is a substructure of $\mathcal{A}$, but not the other way around. For example, $(\mathbb{Z}, <)$ is a substructure of $(\mathbb{Q}, <)$, but it is not elementary. [Why?]

It is of interest to understand when substructures are elementary. However, at the moment, an important feature for us will be merely the existence of (small) elementary substructures.

Theorem 3.6 ("Downwards Lowenheim-Skolem")**.** Fix a countable vocabulary $\mathcal{S}$. Let $\mathcal{A}$ be any structure (where A can have any size). Then there is an elementary substructure $\mathcal{B} \preceq \mathcal{A}$ so that B is countable.

Before proving so, let us discuss some consequences.

Corollary 3.7. For every model $\mathcal{A}$, for a countable signature, there is a countable model $\mathcal{B}$ so that $\mathcal{A} \equiv \mathcal{B}$.

Corollary 3.8. The structure $(\mathbb{R}, <)$ and $(\mathbb{Q}, <)$ are elementary equivalent: $(\mathbb{R}, <) \equiv (\mathbb{Q}, <)$.

Proof. By the above, there is *some* countable $B \subseteq \mathbb{R}$ so that $\mathcal{B} = (B, <)$ is an elementary substructure of $(\mathbb{R}, <)$. In particular $\mathcal{B} \equiv (\mathbb{R}, <)$, and so $\mathcal{B}$ is a DLO. Since $\mathcal{B}$ and $(\mathbb{Q}, <)$ are countable DLOs, they are isomorphic! In particular, $\mathcal{B} \equiv (\mathbb{Q}, <)$. It follows that $(\mathbb{R}, <) \equiv (\mathbb{Q}, <)$ as well. $\square$

In particular, we see that elementary equivalence does not generally imply isomorphism. Recall that $\mathbb{R}$ is not countable, and therefore $(\mathbb{R}, <)$ and $(\mathbb{Q}, <)$ are not isomorphic.

There are more subtle reasons of being non-isomorphic. We will see examples of structures that are elementary equivalent and are both countable, they they are not isomorphic. In fact there are very natural examples: vector spaces and algebraically closed fields.

In the above considerations, all we used is that "all countable DLOs are isomorphic". We will see other examples of axioms with this properties soon. Let us see a simple example here, much simpler than the case for DLOs.

Given finitely many formulas $\psi_1, ..., \psi_k$, we write as short-hand notation $\bigwedge_{i=1}^{k} \psi_i$, or $\bigwedge_{1 \leq i \leq k} \psi_i$, for the long conjunction $\psi_1 \wedge \psi_2 \wedge ... \wedge \psi_k$. [Note that more formally this should be $(...((\psi_1 \wedge \psi_2) \wedge \psi_3) \wedge ... \wedge \psi_k)$. Again $\psi_1 \wedge \psi_2 \wedge ... \psi_k$ is short-hand notation for it. Of course part of why this is a reasonable notation is because we know (check) that the order of paranthesising these conjuctions would *not* change the interpretations. For example, $(\psi_1 \wedge \psi_2) \wedge \psi_3$ and $\psi_1 \wedge (\psi_2 \wedge \psi_3)$ always interpret the same way.]

Exercise 3.9. Consider the empty vocabulary. A model is simply a set. Define the sentences $\varphi_n = \exists x_1 ... \exists x_n (\bigwedge_{1 \leq i < j \leq n} \neg(x_i \approx x_j))$ saying that "there are at least n distinct objects". Let $T = \{\varphi_n : n = 1, 2, 3, ...\}$.

(1) Show that any two countable models of T are isomorphic.
(2) Conclude that any two sets (of any size) are elementary equivalent in the empty vocabulary.

We now turn towards proving Theorem 3.6. Before, we must recall some facts about countable sets.

3.1. **Countable sets.** Understanding countable sets, and how to manipulate them, is extremely important for us (and in many parts of mathematics). We give here a detailed review.

Consider $\mathbb{N} = \{0, 1, 2, ...\}$ the counting numbers. Recall that a set X is **countable** if there is an onto map $f : \mathbb{N} \to X$. (Or if X is empty, we still say that $\emptyset$ is countable, but this will not be an important case.) We identify such map f with the sequence $\langle f(0), f(1), f(2), ... \rangle$, which we think of as an **enumeration** of X. Note that a finite set is also countable. The enumeration does not have to be one-to-one.

Fact 3.10. Suppose X is infinite (not finite) and countable. Then there exists a one-to-one and onto map between $\mathbb{N}$ and X.

Proof. Since X is countable, we may fix some enumeration $x_0, x_1, x_2, \ldots$ of all the members of X. Define $y_0, y_1, \ldots$ recursively as follows.

- $y_0 = x_0$;
- given $y_0, \ldots, y_k$, let m be the smallest natural number so that x_m does not appear in the list $y_0, \ldots, y_k$, and define $y_{k+1} = x_m$.

Note that this process never fails because X is not finite. Now $y_0, y_1, y_2, \ldots$ lists all members of X and for $i \neq j$, $y_i \neq y_j$ [check]. In other words, the map $f(n) = y_n$ is one-to-one and onto between $\mathbb{N}$ and X. $\square$

Fact 3.11. If Y is countable and $g\colon Y \to X$ is onto, then X is countable.

Proof. If $f\colon \mathbb{N} \to Y$ is onto and $g\colon Y \to X$ is onto then $g \circ f\colon \mathbb{N} \to X$ is onto. $\square$

Fact 3.12. Suppose $X \subseteq \mathbb{N}$. Then X is countable.

Proof. We use the fact that any subset of $\mathbb{N}$ has a minimal member (according to the usual ordering of $\mathbb{N}$). Define $x_0, x_1, x_2, \ldots$ as follows:

- x_0 is the minimal member of X;
- given $x_0, \ldots, x_n$, x_{n+1} is the minimal member of $X \setminus \{x_0, \ldots, x_n\}$, *if* $X \setminus \{x_0, \ldots, x_n\}$ is not empty. If $X \setminus \{x_0, \ldots, x_n\}$ is empty, $x_{n+1} = x_n$.

$\square$

Exercise 3.13. If $h\colon X \to Y$ is one-to-one and Y is countable, then X is countable. [Hint: assume first that $Y = \mathbb{N}$. Next use the fact that there is a one-to-one and onto map between Y and $\mathbb{N}$.]

Corollary 3.14. A countable subset of a countable set is countable. (If $X \subseteq Y$, then the map $h\colon X \to Y$, $h(x) = x$, is one-to-one.)

Corollary 3.15. Let X be an *infinite* set. The following are equivalent.

(1) X is countable;
(2) there is a one-to-one map from X to $\mathbb{N}$;
(3) there is a one-to-one and onto map from X to $\mathbb{N}$.

((1) and (2) are equivalent even for finite X.)

Recall that for sets A and B, their product $A \times B$ is the set of all ordered pairs (a, b) where $a \in A$ and $b \in B$.

Fact 3.16. The product $\mathbb{N} \times \mathbb{N}$ is countable.

"Proof by picture":

$$
\begin{array}{ccccccc}
0 & & 1 & & 5 & & 6 \\
0,0 & \to & 1,0 & & 2,0 & \to & 3,0 \\
2 & \swarrow & 4 & \nearrow & 7 & \swarrow & \\
0,1 & & 1,1 & & 2,1 & & \\
3\downarrow & \nearrow & 8 & \swarrow & & & \\
0,2 & & 1,2 & & 2,2 & &
\end{array}
$$

Corollary 3.17. The set $\mathbb{Q}$ is countable.

Proof. Consider the map sending (n,m) to $\frac{n}{m+1}$. It is onto from $\mathbb{N}\times\mathbb{N}$ to $\mathbb{Q}$. $\square$

Note that if X, X', Y, Y' are sets, so that there is an onto map from $X \to X'$ and an onto map from $Y \to Y'$, then there is an onto map from $X \times Y \to X' \times Y'$.

Corollary 3.18. If A and B are countable then so is $A \times B$.

Similarly, if X, X' have the same cardinality (there is a one-to-one and onto map from X to X') and Y, Y' have the same cardinality, then $X \times Y$ and $X' \times Y'$ have the same cardinality.

The following is super important.

Lemma 3.19. If $A_0, A_1, A_2, ...$ are countable sets, then $A = \bigcup_{n=0,1,...} A_n$ is also countable.

Proof. It suffices to find an onto map from $\mathbb{N}\times\mathbb{N}$ to A. [why?]

For each n, A_n is countable, so we may enumerate it by $a^n_0, a^n_1, a^n_2, ...$
Define $g\colon \mathbb{N}\times\mathbb{N} \to A$ by $g(n,m) = a^n_m$. Then g is onto.

$$\begin{array}{ccc} a^0_0 & a^0_1 & a^0_2 \\ \\ a^1_0 & a^1_1 & a^1_2 \\ \\ a^2_0 & a^2_1 & a^2_2 \end{array}$$

$\square$

Corollary 3.20. "A countable union of countable sets is countable". That is, if I is a countable set, and for each $i \in I$ we have some countable set A_i, then the set $A = \bigcup_{i\in I} A_i$ is countable.

For a set A let $A^{<\mathbb{N}} = \bigcup_{n\in\mathbb{N}} A^n$, the set of all finite tuples from A, of arbitrary finite length.

Exercise 3.21. Suppose that A is countable. Prove that $A^{<\mathbb{N}}$ is countable.

Corollary 3.22. Let $\mathcal{S}$ be a countable signature (countably many relation and function symbols). Let $\mathcal{F}$, $\mathcal{T}$ be the set of all formulas in the signature $\mathcal{S}$, and the set of all terms in the signature $\mathcal{S}$, respectively. Then $\mathcal{F}$ and $\mathcal{T}$ are countable.

Proof. Let A be $\mathcal{S}$ together with the symbols $\{(,), \to, \vee, \wedge, \exists, \forall, \bot, x_0, x_1, x_2, ...\}$. ($x_0, x_1, ...$ is supposed to represent some infinite sequence of variables that we will use for the formulas.) Any formula and any term can be identified as a finite string of symbols from A. In other words, we may identify $\mathcal{T}$ and $\mathcal{F}$ as subsets of $A^{<\mathbb{N}}$, which is countable. $\square$

In particular, if our vocabulary is finite (as for linear orders, groups, fields, graphs) then the *language* is countable.

Fact 3.23. The set of real numbers $\mathbb{R}$ is not countable.

Note that English is as well a countable language. In particular, there are real numbers which you cannot *describe* in any way.

3.2. **Existence of elementary substructures.** First, being "just a substructure" already gives us "some elementarity". By definition, if $\mathcal{B}$ is a substructure of $\mathcal{A}$, then for any atomic formula $\varphi(x_1, ..., x_n)$, for any $b_1, ..., b_n$ from B, $\varphi^{\mathcal{B}}(b_1, ..., b_n) \iff \varphi^{\mathcal{A}}(b_1, ..., b_n)$.

We may define **quantifier free** formulas as we defined formulas, omitting the quantifiers stage. That is, atomic formulas are quantifier free, negations, conjuctions, disjunctions, and implications between quantifier free formulas, are again quantifier free formulas.

Exercise 3.24. Suppose $\mathcal{B}$ is a substructure of $\mathcal{A}$. Prove that for any quantifier free formula $\varphi(x_1, ..., x_n)$, for any $b_1, ..., b_n$ in B,

$$\varphi^{\mathcal{B}}(b_1, ..., b_n) = 1 \iff \varphi^{\mathcal{A}}(b_1, ..., b_n) = 1.$$

[You proved this in Pset 3 in greater generality, for an embedding between structures.]

Being an elementary substructure is quite powerful, and therefore seemingly difficult to verify. One has to worry about all formulas and worry about the difference between how $\mathcal{A}$ and $\mathcal{B}$ may interpret them. The following gives a simpler "step by step" criterion for verifying elementarity.

Theorem 3.25 (Tarski-Vaught criterion)**.** Suppose $\mathcal{B}$ is a substructure of $\mathcal{A}$. The following are equivalent:

(1) $\mathcal{B} \preceq \mathcal{A}$;
(2) for any formula $\varphi(x_1, ..., x_n, x_{n+1})$, for any $b_1, ..., b_n$ in B, *if* there is some a in A for which $\varphi^{\mathcal{A}}(b_1, ..., b_n, a)$ holds, *then* there is some b in B for which $\varphi^{\mathcal{A}}(b_1, ..., b_n, b)$ holds.

[Why is the second line *not* simply the definition of elementarity for the formula $\exists(x_{n+1})\varphi$?]

Proof. (2) $\implies$ (1) ("The easy direction"): Assume that $\mathcal{B} \prec \mathcal{A}$, $\varphi(x_1, ..., x_n, x_{n+1}$ is a formula, $b_1, ..., b_n$ are in B and $\varphi^{\mathcal{A}}(b_1, ..., b_n, a)$ is true for some a in A. Consider the formula $\theta = (\exists x_{n+1})\varphi$. Then $\theta^{\mathcal{A}}(b_1, ..., b_n)$ is true. By the elementarity assumption, $\theta^{\mathcal{B}}(b_1, ..., b_n)$ holds as well. In particular, there is some $b \in B$ so that $\theta^{\mathcal{B}}(b_1, ..., b_n, b)$ holds. Now again by the elementarity assumption $\theta^{\mathcal{A}}(b_1, ..., b_n, b)$ holds, as required.

(2) $\implies$ (1) (the main point): We prove by induction on the construction of formulas that for any formula $\varphi(x_1, ..., x_n)$ and any $b_1, ..., b_n$ in B,

$$(\star)\ \varphi^{\mathcal{B}}(b_1, ..., b_n) \iff \varphi^{\mathcal{A}}(b_1, ..., b_n).$$

For atomic formulas, this follows from Exercise 3.24. (Here we are not using elementarity, just that $\mathcal{B}$ is a substructure of $\mathcal{A}$.)

The connectives case is similar to arguments we have seen a few times now. For example, if $\varphi = \neg\psi$ then

$$\varphi^{\mathcal{B}}(b_1, ..., b_n) = 1 \iff \psi^{\mathcal{B}}(b_1, ..., b_n) = 0 \iff \psi^{\mathcal{A}}(b_1, ..., b_n) = 0 \iff \varphi^{\mathcal{A}}(b_1, ..., b_n) = 1,$$

where the middle $\iff$ is by the inductive assumption. The other connectives are similar.

Let us focus on the case of an existential quantifier. Assume that $(\star)$ holds for ψ, and we need to prove it for $\varphi = (\exists x)\psi$. We will also focus on the case where the quantified-over variable x does not appear in the list $x_1, ..., x_n$.

If $\varphi^{\mathcal{B}}(b_1, ..., b_n)$ holds, then there is some $b \in B$ for which $\psi^{\mathcal{B}}(b_1, ..., b_n, b)$ is true. It follows that $\psi^{\mathcal{A}}(b_1, ..., b_n, b)$ is true, and therefore $\varphi^{\mathcal{A}}(b_1, ..., b_n)$ holds. [The assumption (2) was not used here.]

Finally, assume that $\varphi^{\mathcal{A}}(b_1, ..., b_n)$ holds, so there is some $a \in A$ for which $\psi^{\mathcal{A}}(b_1, ..., b_n, a)$ holds. By assumption (2), there is some $b \in B$ for which $\psi^{\mathcal{A}}(b_1, ..., b_n, b)$ holds. *By the inductive hypothesis*, $\psi^{\mathcal{B}}(b_1, ..., b_n, b)$ holds, since $b_1, ..., b_n, b$ are all from B. In turn $\varphi^{\mathcal{B}}(b_1, ..., b_n)$ is true, as required. $\square$

Finally we prove a strengthened version of the "downwards Lowenheim-Skolem" theorem stated above.

Theorem 3.26. Let $\mathcal{S}$ be a countable signature. Let $\mathcal{A}$ be a structure and $X \subseteq A$ a subset. Assume furthermore that X is countable. Then there is an elementary substructure $\mathcal{B} \preceq \mathcal{A}$ with B countable and $X \subseteq B$.

Proof. We will define a sequence of countable sets $B_0, B_1, B_2, ...$ where $B_0 = X$. At each stage we will add more "witnesses" to satisfy the Tarski-Vaught criterion. At the end, we will have "caught our tails".

Let $\mathcal{F}$ be the set of all formulas in the language. It is countable, since the vocabulary is countable.

Assume that B_k is defined and is countable. We construct B_{k+1} as follows. For each formula $\varphi(x_1, ..., x_n, x_{n+1})$ and for any parameters $b_1, ..., b_n$ in B_k, we ask: is there some a in A for which $\mathcal{A} \models \varphi(b_1, ..., b_n, a)$? If there is such an a, *choose* one. (Call it, say, $a_{\varphi(x_1,...,x_n,x_{n+1}),b_1,...,b_n}$.)

How many pairs of formula & sequence of parameters from B_k are there? $\mathcal{F} \times B_k^{<\mathbb{N}}$. Countably many! Collect all of these, together with B_k, to form B_{k+1}. So B_{k+1} is countable as a union of two countable sets.

Finally, define $B = \bigcup_{k=0,1,2,...} B_k$. B is countable as a countable union of countable sets. $X = B_0 \subseteq B$. Using the Tarski-Vaught criterion, we show that B is an elementary substructure of $\mathcal{A}$.

First we need to check that $\mathcal{B}$ is a substructure, that is, it is closed under all functions $F^{\mathcal{A}}$ for F a function symbol in $\mathcal{S}$. We could have "taken care of it directly", but it actually follows from our construction. Given $b_1, ..., b_n$ from B, there is some large enough k so that $b_1, ..., b_n$ are all in B_k [Tail = caught]. Consider the formula $\psi(x_1, ..., x_n, x_{n+1}) = (x_{n+1} \approx F(x_1, ..., x_n))$. For $a = F^{\mathcal{A}}(b_1, ..., b_n)$, $\mathcal{A} \models \psi(b_1, ..., b_n, a)$. We need to show that $a \in B$. By construction, *some* $a' \in A$ for which $\mathcal{A} \models \psi(b_1, ..., b_n, a')$ was thrown into B_{k+1}. However, since $\mathcal{A} \models \psi(b_1, ..., b_n, a')$, then $a = F^{\mathcal{A}}(b_1, ..., b_n) = a'$, so $a = a'$ is in B, as required. [Note that at this point we can actually talk about $\mathcal{B}$ as a *structure* whose domain is B.]

The argument for the Tarski-Vaught criterion is similar. Let $\varphi(x_1, ..., x_n, x_{n+1})$ be any formula, $b_1, ..., b_n$ some members of B, and assume that $\varphi^{\mathcal{A}}(b_1, ..., b_n, a)$ is true for some a in A. Fix k large enough so that $b_1, ..., b_n$ are all in B_k. Then in the construction we threw into B_{k+1} some a' so that $\varphi^{\mathcal{A}}(b_1, .., b_n, a')$ holds. In particular this a' is in B, as required in the Tarski-Vaught criterion. $\square$

Example 3.27. Consider the "field of complex numbers" $(\mathbb{C}, +, \cdot, 0, 1)$. Recall that $\mathbb{C}$ is **algebraically closed**, meaning that for any polynomial $P(x) = \sum_{i=0}^{n} a_i x^i$ with coefficients $a_i \in \mathbb{C}$ (which are not all zero), then P has a root: some $c \in \mathbb{C}$ for which $P(c) = 0$. Note that polynomials are essentially terms in this language.

What does the construction above look like for this structure, when we start with the empty set?

For example, any natural number k has a term $t_k = 1 + ... + 1$ (k-times) which "defines it". Since it is true in $\mathbb{C}$ that $(\exists x)x = t_k$, then k will be added to our first stage B_1. Similarly, you can see that every rational number $q \in \mathbb{Q}$ will be added to B_1. Moreover, as $(\exists x)(x^2 + 1 = 0)$ is true in $\mathbb{C}$, $\imath$ (the square root of -1), will be added to B_1 as well.

The elementary substructure we get is a countable algebraically closed field. In fact, we get the minimal algebraically closed field (of characteristic 0).

[Remark: in this case, in fact $B_1 = B_2 = B_3 ....$ That is, after one step we already get an algebraically closed field. This however is something unique to fields, and not general model theoretic.]

Example 3.28. Consider $\mathcal{A} = (\mathbb{N}, <)$ and $\mathcal{B} = (\mathbb{N} \setminus \{0\}, <)$. Then

- $\mathcal{B} \subseteq \mathcal{A}$ (a substructure).
- $\mathcal{B} \simeq \mathcal{A}$, the map $f(n) = n - 1$ is an isomorphism. (In particular, $\mathcal{A} \equiv \mathcal{B}$).
- However, $\mathcal{B}$ is *not* an elementary substructure of $\mathcal{A}$. Specifically, $\mathcal{B} \models \neg\varphi(1)$ and $\mathcal{A} \models \varphi(1)$, where $\varphi(x) = (\exists y)(y < x)$.

3.3. **Theories.** As always, we work with some fixed vocabulary. Say that T is a **theory** if T is just a set of sentences in the language. Say that T is **satisfied** by a model $\mathcal{A}$, denoted $\mathcal{A} \models T$, if $\mathcal{A} \models \varphi$ for any $\varphi \in T$. (We will also say that $\mathcal{A}$ is **a model of** T.) Say that T is **satisfiable** if there exists *some* model which satisfies it. (If it has a model.)

Remark 3.29. If T is a satisfiable theory (it has some model), then it has a *countable* model. It may tempt us to think that, in order to understand a theory T, we may restrict out attention to countable models. This, however, turns out not to be so.

Definition 3.30. Say that a theory T **logically implies** a sentence φ, denoted $T \models \varphi$, if for *any* model $\mathcal{A}$ of T, $\mathcal{A} \models \varphi$ as well. We will also say that φ is a **logical consequence** of T.

A sentence φ is said to be **logically valid** if it is true in *any* model. We will also denote this by $\models \varphi$. (That is, "the empty theory" logically implies φ.)

Remark 3.31. Logical implication is *the* central question here. For example, if you prove something about vector spaces, you start with some structure satisfying the vector space assumption, and you prove things in that structure. Since the structure was arbitrary, you were proving that something is a consequence of the vector space axioms.

Another central question is whether some T is satisfiable. Note that $T \models \varphi$ if and only if the theory$T \cup \{\neg\varphi\}$ is *not* satisfiable. (The latter will sometimes be denoted $T \cup \{\neg\varphi\} \models \perp$.)

Example 3.32.

- Let T be the theory of linear orders, in the signature $<$. Let φ be the sentence $(\forall x_1)(\forall x_2)(\forall x_3)(\forall x_4)(((x_1 < x_2) \wedge (x_2 < x_3) \wedge (x_3 < x_4)) \rightarrow (x_1 < x_4))$. Then T logically implies φ. On the other hand, T *does not* logically imply the (Density) axiom of DLO.
- Let T be the theory DLO of dense linear orders. Let φ be the sentence $(\forall x)(\forall y)((x < y) \rightarrow (\exists z)(\exists w)((x < z) \wedge (z < w) \wedge (w < y)))$. Then T logically implies φ.
- Let φ be the sentence $(\forall x)(x \approx x)$. Then φ is logically valid.
- Let P be an unary predicate. Let φ be $(\forall x)(P(x) \vee \neg P(x))$. Then φ is logically valid. Let ψ be the sentence $(\forall x)P(x) \vee (\forall x)\neg P(x)$. Then ψ is *not* logically valid, meaning there is *some* model in which it is false.

Definition 3.33. Fix a signature $\mathcal{S}$. A theory T (of sentences in the signature $\mathcal{S}$) is **complete** if for any sentence φ (in the signature $\mathcal{S}$) either $\varphi \in T$ or $\varphi \notin T$.

Definition 3.34. Let $\mathcal{A}$ be a structure. The **theory of** $\mathcal{A}$, denoted $\mathrm{Th}(\mathcal{A})$, is the set of all sentences φ in the language so that $\mathcal{A} \models \varphi$.

Claim 3.35. For any structure $\mathcal{A}$, $\mathrm{Th}(\mathcal{A})$ is a complete theory.

Proof. Fix a sentence φ in the language. If $\varphi^{\mathcal{A}} = 1$ (is true), then $\varphi \in \mathrm{Th}(\mathcal{A})$. Otherwise, by the definition of the interpretation of formulas in a structure, $(\neg\varphi)^{\mathcal{A}} = 0$ (is false), so $\neg\varphi \in \mathrm{Th}(\mathcal{A})$. $\square$

We will often identify a theory T with its logical consequences: the set of all sentences φ in the language for which $T \models \varphi$. (Note that a structure $\mathcal{A}$ is a model for T if and only if it is a model for the set of consequences of T.) In that spirit, we may say that T is complete if for any sentence φ, either $T \models \varphi$ or $T \models \neg\varphi$.

Recall that two models $\mathcal{A}$ and $\mathcal{B}$ (for the same vocabulary) are elementary equivalent, denoted $\mathcal{A} \equiv \mathcal{B}$, if for any sentence φ, $\mathcal{A} \models \varphi \iff \mathcal{B} \models \varphi$.

Remark 3.36. Two models $\mathcal{A}$ and $\mathcal{B}$ are elementary equivalent if and only if $\mathcal{B} \models \mathrm{Th}(\mathcal{A})$.

Being elementary equivalent means that we cannot distinguish the models using any sentence in our formal language. On the other hand, recall that being isomorphic means that the models are truly "essentially the same". We saw that elementary equivalence does not imply isomorphism. For example, $(\mathbb{R}, <)$ and $(\mathbb{Q}, <)$ are elementary equivalent, but not isomorphic, as they have different size. There could also be elementary equivalent $\mathcal{A}$, $\mathcal{B}$ of the same size, which are still not isomorphic. There are in fact many interesting such examples. For example, $\mathbb{R}$ and $\mathbb{R}^2$, viewed as vector spaces in the language discussed earlier, are not isomorphic (they have different dimensions), but they turn out to be elementary equivalent (and they have the same size). Also, any two countable algebraically closed fields of characteristic 0 are elementary equivalent, yet they are not all isomorphic to one another.

Finally, what we mentioned earlier about $(\mathbb{R}, <)$ and $(\mathbb{Q}, <)$ can be strengthened, and applies in greater generality.

Lemma 3.37. Let T be a theory (in a countable vocabulary) so that any two *countable* models of T are isomorphic. Then any two models (of any size) are elementary equivalent. In particular, T is complete: for any sentence φ, either $T \models \varphi$ or $T \models \neg\varphi$.

Proof. For any two models $\mathcal{A}$ and $\mathcal{B}$ of T, we may find countable models $\mathcal{A}'$ and $\mathcal{B}'$ so that $\mathcal{A} \equiv \mathcal{A}'$ and $\mathcal{B} \equiv \mathcal{B}'$. By assumption, $\mathcal{A}'$ and $\mathcal{B}'$ are isomorphic, so $\mathcal{A}' \equiv \mathcal{B}'$. It follows that $\mathcal{A} \equiv \mathcal{B}$.

Fix a sentence φ. Assume for contradiction that neither $T \models \varphi$ nor $T \models \neg\varphi$. The first assumption means (by definition) that there is some model $\mathcal{A}$ for T so that $A \models \neg\varphi$; the second that there is some model $\mathcal{B}$ for T so that $\mathcal{B} \models \varphi$. By the previous argument however $\mathcal{A} \equiv \mathcal{B}$, a contradiction. $\square$

Corollary 3.38. $\mathrm{Th}(\mathbb{R}, <) = \mathrm{Th}(\mathbb{Q}, <)$ and is precisely all logical consequences of the theory DLO.

Proof. If T is the set of logical consequences of DLO, then any model of DLO must satisfy T, and therefore the theory of any such model is precisely T. $\square$

Exercise 3.39. Let E be a binary relation symbol. Consider the theory T saying that E is an equivalence relation with precisely 2 equivalence classes, and each equivalence class is infinite. More precisely, T consists of the axioms

- $(\forall x)(x\ E\ x)$,
- $(\forall x)(\forall y)(x\ E\ y \to y\ E\ x)$,
- $(\forall x)(\forall y)(\forall z)((x\ E\ y \wedge y\ E\ z) \to x\ E\ z)$
- $(\exists x)(\exists y)((\forall z)(z\ E\ x \vee z\ E\ y) \wedge \neg(x\ E\ y))$,

as well as the axioms

- $(\forall x)(\exists x_1)...(\exists x_n)(\bigwedge_{i=1}^{n}(x_i\ E\ x) \wedge \bigwedge_{i \neq j} \neg(x_i \approx x_j))$,

for each natural number n. Prove that any two countable models for T are isomorphic.

4. A playful approach

Many construction in mathematics are often cast in terms of games and strategies. We will discuss some games which characterize elementary equivalence between models, and isomorphism between countable models. Several facts in this section will be stated without proof. Those will not be part of our core material in the class, but feel free to read more or reach out if you want to discuss these or further directions.

For notational convenience, let us restrict attention to *relational languages*, meaning that we have only relation symbols and no function symbols (so no constant symbols either). These ideas can be generalized to arbitrary languages without much difficulty.

Fix a relational signature $\mathcal{S}$. Fix structures $\mathcal{A}, \mathcal{B}$. Any subset $X \subseteq A$ may be viewed as a substructure. Given $X \subseteq A$ and $Y \subseteq B$, and a function $f: X \to Y$, say that f is a **partial embedding** if f is an embedding from the substructure of $\mathcal{A}$ with domain X to the substructure of $\mathcal{B}$ with domain Y. (Equivalently, this is just the definition of embedding but we only take elements from X instead of all of A.)

Consider the following two-player game, which we will denote $G(\mathcal{A}, \mathcal{B})$. We may call this the **Back-and-Forth** game. This is often called the **Ehrenfeucht-Fraïssé** game. (Extra credit for pronunciation!)

A "play" in the game looks as follows.
First, player I picks some a_0 in A.
In response, player II must pick some b_0 in B. This b_0 must satisfy that the function $f: \{a_0\} \to \{b_0\}$, $f(a_0) = b_0$, is a partial embedding.

Next, player I picks some b_1 in B.
In response, player II must pick some a_1 in A, so that the function $f: \{a_0, a_1\} \to \{b_0, b_1\}$.

... Suppose we arrived at some even stage $2k$, where $a_0, ..., a_{2k-1}$, $b_0, ..., b_{2k-1}$ were already played (according to the rules).
Player I now picks some a_{2k} in A.
In response, player II picks some b_{2k} in B, so that $f: \{a_0, ..., a_{2k}\} \to \{b_0, ..., b_{2k}\}$, $f(a_i) = b_i$, is a partial embedding.

Next, at stage $2k+1$, player I picks some b_{2k+1} from B.
In response, player II must pick some a_{2k+1} from A, so that $f: \{a_0, ..., a_{2k+1}\} \to \{b_0, ..., b_{2k+1}\}$, $f(a_i) = b_i$, is a partial embedding.

Clearly, player I is on the offense, while player II is on the defense. What is happening is that that the two players are studying the models $\mathcal{A}$ and $\mathcal{B}$. Player II really believes that they are isomorphic. Player I, however, is skeptical. At each stage player I challenges player II, by adding some members (either to the domain or the range of a "potential isomorphism"), and player II must respond by showing how such potential isomorphism will behave on these members.

This game continues indefinitely... Who wins? If at any point, player II failed to find a response, then player II loses. If player II prevails through all finite stages, always providing a response, then player II wins.

What is a **strategy** in a game? Exactly what it sounds like... A strategy for player II is some pre-determined decision of how to respond to any move of player I. [Technically, a strategy us a function τ that takes as input the information of everything that has happened in the game so far, together with the current stage of the game, and spits out the next move.]

When is a strategy a **winning strategy**? If it guarantees a victory. A strategy τ for player II is a winning strategy if *no matter* what moves player I plays, as long as player II follows this strategy, player II will win. (In this case, it simply means that player II always has a legit move to make.)

For example, if you are playing connect-4, the second player may play a "mirroring strategy", to always play the mirror image of player I's move. This, however, is a losing strategy!

Theorem 4.1. Suppose that $\mathcal{A}$ and $\mathcal{B}$ are countable structures. The following are equivalent.

(1) $\mathcal{A}$ and $\mathcal{B}$ are isomorphic.
(2) Player II has a winning strategy for $G(\mathcal{A}, \mathcal{B})$.

Remark 4.2. Consider the language consisting of one relation symbol $<$. Let $\mathcal{A}$ and $\mathcal{B}$ be DLOs. In Pset 1, Question 3, you described a winning strategy for player II in the game $G(\mathcal{A}, \mathcal{B})$.

Proof. Assume first that $\mathcal{A} \simeq \mathcal{B}$, and fix some isomorphism $f\colon A \to B$. Player II has the following strategy, to "respond according to f". That is, when player I plays a_{2k} in A, player II responds $b_{2k} = f(a_{2k})$. When player I plays b_{2k+1} in B, player II responds $a_{2k+1} = f(b_{2k+1})$. Since f is an isomorphism, its restriction to any subdomain is a partial embedding. It follow that the moves player II makes are always legit, and so player II wins.

Assume now that player II has some winning strategy τ in the game $G(\mathcal{A}, \mathcal{B})$, and construct an isomorphism between $\mathcal{A}$ and $\mathcal{B}$. This is exactly what we have done in the first week of classes.

By assumption, A and B are countable. We may fix enumerations $a'_0, a'_1, a'_2, \ldots$ and $b'_0, b'_1, b'_2, \ldots$ of A and B respectively. We define new enumerations $a_0, a_1, a_2, \ldots$ and $b_0, b_1, b_2, \ldots$ so that the map $f(a_i) = b_i$ will be an isomorphism.

While player I is in some sense "the bad one", challenging our belief that $\mathcal{A}$ and $\mathcal{B}$ are isomorphic, their skepticism turns out useful for us now. We now take the role of player I, describing their moves in the game.

First, let player I play $a_0 = a'_0$.
Player II now responds according to the strategy to provide some b_0 (which is not necessarily b'_0).
Next, let player I play $b_1 = b'_0$.
Player II responds according to the strategy to provide some a_1 in A.

At stage $2k$, player I plays $a_{2k} = a'_k$, player II responds with some b_{2k}. Then player I plays $b_{2k+1} = b'_k$, and player II responds with some a_{2k+1}.

Finally, define $f\colon A \to B$ by $f(a_i) = b_i$. *Since player II always played according to a winning strategy* the map f is in fact an isomorphism. [Why?]

[Unlike in Week 1, we did not worry here about whether, at stage $2k$, a'_k already appears as some a_i or not. Similarly for b'_k at stage $2k+1$. Does this pose a problem?] $\square$

For uncountable models, "player II having a winning strategy" does not necessarily imply isomorphism. It is however a strong and interesting notion of "similarity" between the models.

Shorter plays. [*Optional reading topic*] Note that player II's task is not easy, finding ahead of time a strategy that will last infinitely many rounds of the game. Fix a natural number n. We may consider a game $G_n(\mathcal{A}, \mathcal{B})$ which is played just as $G(\mathcal{A}, \mathcal{B})$ but is halted after n many rounds. Again player II can lose by failing to respond at any given round, and wins by prevailing until the final round.

Theorem 4.3. Let $\mathcal{A}$ and $\mathcal{B}$ be structures for a *finite* relational signature. The following are equivalent.

(1) $\mathcal{A}$ and $\mathcal{B}$ are elementary equivalent.
(2) For each n, player II has a winning strategy for the game $G_n(\mathcal{A}, \mathcal{B})$.

If you are interested, you can find a proof of this result in [Marker, Theorem 2.4.6]. I will be happy to discuss it. (The definition of G_n there is slightly different, and the result more refined, as [Marker, Lemma 2.4.9].)

Note that for $n < m$, the strategy τ_m for $G_m(\mathcal{A}, \mathcal{B})$ may (and most likely will) *not* agree with τ_n, even on the first n stages.

Note that there is no assumption on the cardinalities of A and B. They are not even assumed to be of the same cardinality! In fact, you can use Theorem 4.3 to give another proof that $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ are elementary equivalent.

Remark 4.4. It is worth noting that clause (2) in Theorem 4.3 is formulated completely in terms of the structures. (Unlike "elementary equivalence", which is purely described in terms of the formal language.) That is, the definition of being a "partial embedding" only talks directly about the structure, and does not involve the formal language.

4.1. **The random graph.** Consider the vocabulary $\{E\}$ of one binary relation. (Intended here for graphs.) Let ψ_n be the following sentence

$$\forall x_1 ... \forall x_n \forall y_1 ... \forall y_n \left((\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{n} x_i \neq y_j) \to \exists z \bigwedge_{i=1}^{n} ((x_i \mathrel{E} z) \wedge \neg(y_i \mathrel{E} z)) \right).$$

Here $\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{n} x_i \neq y_j$ is notation for $x_1 \neq y_1 \wedge x_1 \neq y_2 \wedge ... \wedge x_1 \neq y_n \wedge x_2 \neq y_1 \wedge ...$ As usual $x \neq y$ is short for $\neg \approx (x, y)$. Similarly, $\bigwedge_{i=1}^{n}$ is short for n many $\wedge$'s.

What does ψ_n say: for any two disjoint subsets of vertices $\{a_1, ..., a_n\}$ and $\{b_1, ..., b_n\}$, we may find some vertex c which has an edge with each of the a_i's and does not have an edge with any of the b_i's.

The most common example of a graph with this property is the **random graph**. This is a graph whose vertex set is $\mathbb{N}$ (some countable infinite set) and the edge relations are decided randomly by flipping a coin for each pair (n, m). You can see, at least intuitively, that such a graph will satisfy ψ_n for any n. [We can also construct the graph in a more explicit way. Feel free to ask.]

Let T be the theory containing the sentences:

- $\forall x \forall y(x\ E\ y \to y\ E\ x)$ (the graph axiom);
- $\forall x \neg(x\ E\ x)$ ("no loops");
- $\exists x \exists y(\neg x \approx y)$;
- ψ_n for each n.

(T is infinite.)

Exercise 4.5. (1) Suppose $\mathcal{A}$ and $\mathcal{B}$ are models of T. Show that player II has a winning strategy for the game $G(\mathcal{A}, \mathcal{B})$.

(2) Conclude that any two countable models of T are isomorphic.

(3) Conclude that the sentences which are true for the random graph (the theory of the random graph) are precisely the logical consequences of T.

You can find more details about this in [Marker, p. 50], and I will be happy to talk about it. You can see there also some probabilistic facts about the random graph.

In a sense, proving that the "back-and-forth" works here is *easier* than what you have done for DLOs. The theory T has infinitely many axioms, quite directly ensuring that the back-and-forth can go through. In the case of DLO there are only finitely many axioms. You had to do some more work to "extract" information about larger finite sets, using axioms which only talk about 2 or 3 objects at a time.

Fraïssé limits. [*Optional reading topic*] Both the the DLO $(\mathbb{Q}, <)$ and the random graph can be seen as examples of "Fraisse limits". The structure $(\mathbb{Q}, <)$ can be seen as the "limit of all finite linear order", and the random graph can be seen as the "limit of all finite graphs". The Fraisse limit is a construction that allows in some generality to construct a "limit structure" to a collection of finite structures.

If you are interested, you can read about it in [Hodges, Section 7.1], and I will be happy to talk about it. (You have all the tools necessary to read this section.)

Remark 4.6. If you like category theory, you will *love* Friasse limits.

5. A remark on $\mathbb{Q}, \mathbb{R}$ and elementary substructures

We saw that $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ are elementary equivalent. In fact $(\mathbb{Q}, <)$ is an elementary substructure of $(\mathbb{R}, <)$. Let us see why.

Lemma 5.1. Suppose $\mathcal{A} = (A, <)$ is countable DLO. Suppose $\bar{a} = a_1, ..., a_n$ and $\bar{b} = b_1, ..., b_n$ are from A and have the same type (as in Pset 1). Then there is an automorphism of $\mathcal{A}$, $f : A \to A$, such that $f(a_i) = b_i$.

This is essentially what we proved in Week 1. We can start with the sequences $\bar{a} = a_1, ..., a_n$, $\bar{b} = b_1, ..., b_n$, and continue the back-and-forth process $a_1, ..., a_n, a_{n+1}, ...,$ $b_1, ..., b_n, b_{n+1}$ so that both the a_n and b_n sequences enumerate all of A, and the map sending a_i to b_i is an isomorphism from $\mathcal{A}$ to $\mathcal{A}$.

Corollary 5.2. Let $\mathcal{A}$ be a countable DLO. Suppose $\bar{a} = a_1, ..., a_n$ and $\bar{b} = b_1, ..., b_n$ are from A and have the same type. Then for any formula $\varphi(x_1, ..., x_n)$,

$$\mathcal{A} \models \varphi(a_1, ..., a_n) \iff \mathcal{A} \models \varphi(b_1, ..., b_n).$$

Proof. If $f\colon A \to A$ is an automorphism of $\mathcal{A}$ sending a_i to b_i, then

$$\mathcal{A} \models \varphi(a_1, ..., a_n) \iff \mathcal{A} \models \varphi(f(a_1), ..., f(a_n)) \iff \mathcal{A} \models \varphi(b_1, ..., b_n).$$

□

Corollary 5.3. Suppose $\mathcal{B} \subseteq \mathcal{A}$ is a substructure and both are DLOs. Then $\mathcal{B} \preceq \mathcal{A}$.

Proof. Let us apply the Tarski-Vaught criterion. Let φ be a formula, $b_1, ..., b_n$ from B and a from A so that $\mathcal{A} \models \varphi(b_1, ..., b_n, a)$. We want to find b in B so that $\mathcal{A} \models \varphi(b_1, ..., b_n, b)$.

We consider the following cases:

(1) $a = a_i$ for some i;
(2) $a > a_i$ for $i = 1, ..., n$;
(3) $a < a_i$ for $i = 1, ..., n$;
(4) otherwise, we may find i, j so that $a_i < a < a_j$ where a_i is largest below a and a_j is smallest above a.

In case (1), we take $b = a = a_i$ and we are done. In all other cases, using the fact that $\mathcal{B}$ is a DLO, we may find $b \in B$ so that the two sequences $a_1, ..., a_n, a$ and $b_1, ..., b_n, b$ have the same type. (You did something similar in Pset 1). By the previous corollary, $\mathcal{A} \models \varphi(b_1, ..., b_n, b)$, as required. □

Remark 5.4. Lemma 5.1, as well as the following corollaries, are also true with $\mathcal{A}$ replaced by $(\mathbb{R}, <)$. You can prove Lemma 5.1 directly in this case (writing explicitly an automorphism), without an appeal to Cantor's isomorphism theorem.

Corollary 5.5. $(\mathbb{Q}, <)$ is an elementary substructure of $\mathbb{R}$.

Proof. By the downwards Lowenheim-Skolem theorem, there is some $\mathcal{A} \preceq (\mathbb{R}, <)$ with A countable and $\mathbb{Q} \subseteq A$. So $\mathcal{A}$ is a countable DLO, and therefore $(\mathbb{Q}, <) \preceq \mathcal{A}$. It follows that $(\mathbb{Q}, <) \preceq (\mathbb{R}, <)$. □

Remark 5.6. We just witnessed a very interesting phenomenon: where a substructure is automatically an *elementary substructure*. This is also true for algebraically closed fields: if F_1 is a subfield of F_2 and both are algebraically closed, then it is in fact an elementary substructure, with the language $+, \cdot, 0, 1,$. Similarly this is true for vector spaces, in the language we represented them above.

A related very interesting and important concept is that of *quantifier elimination*. We will not go into it in this class. You can find a (very rudimentary) example in [Enderton, p. 190]. See [Marker, Section 3.1] to learn more. (In particular, DLOs and algebraically closed fields "have quantifier elimination".) As always, feel free to talk to me about it if you are interested.

6. Constructing models

Let us now turn to the question: given a theory T (for a signature $\mathcal{S}$, can we find a model for T? (That is, a structure $\mathcal{A}$ in the signature $\mathcal{S}$ so that $\mathcal{A} \models T$.)

Of course, this is not always possible. If there is some sentence φ so that both φ and $\neg\varphi$ are in T, then there cannot be any model for T. More generally, if $T \models \varphi$ and $T \models \neg\varphi$, then there cannot be a model for T. Note that for such T, necessarily $T \models \psi$ for *any* sentence ψ. This is simply because the statement "$T \models \psi$" is of the form "*if* $\mathcal{A}$ is a model for T, *then* ...". If there are no models of T, this statement is always true...

This is indeed not an easy question, and we will be dealing with it for a while. To begin, let us make some simplifying assumption, and develop some intuition. *Assume* that (1) T is a complete theory (for any φ, either $\varphi \in T$ or $\neg\varphi \in T$), and (2) that there *is not* φ for which both $\varphi, \neg\varphi$ are in T (T does not contain any immediate contradiction). Recall that these two conditions are true for $\mathrm{Th}(\mathcal{A})$ for any structure $\mathcal{A}$. So it really looks like T is the theory of some structure... Yet it is still not clear.

Remark 6.1. We started by being generous with our logical connectives. This allowed us to more freely express things in the language. However, we don't really need all of them. As we have seen, if we just use the connectives $\wedge$ and $\neg$, and the universal quantifier $\forall$, we do not "lose any expressive power". Meaning, we can still represent any formula, which also uses $\vee, \rightarrow, \exists$, using just $\wedge, \neg, \forall$. (Up to logical equivalence of formulas.)

The choice of connectives is not too important. We can also use $\vee, \neg, \exists$ to express all the others.

When proving things by induction on the construction of formulas, it is convenient to restrict to fewer cases, say just $\wedge, \neg, \forall$, to avoid repeating the same argument.

6.1. **Henkin theory.** Fix a signature $\mathcal{S}$. Assume T is a complete theory with no contradictions as above. *Assume further* that the signature contains many (say, infinitely many) constant symbols. Then, we may try to construct a structure using these constant symbols.

Specifically: we may try to create a structure $\mathcal{A} = (A, ...)$ where A is the set of all constants in the signature. For any n-ary relation symbol R in the signature and any $c_1, ..., c_n$ from A (that is, constant symbols in the signature), we need to decide whether $R^{\mathcal{A}}(c_1, ..., c_n)$ is true or false (to define the structure $\mathcal{A}$).

Seems like we have a very natural way of making this decision: note that $\varphi = R(c_1, ..., c_n)$ is an atomic sentence in the signature $\mathcal{S}$. By our assumptions, either $\varphi \in T$, in which case we will decide that $R^{\mathcal{A}}(c_1, ..., c_n)$ is true, or $\neg\varphi \in T$, in which case we will decide that $R^{\mathcal{A}}(c_1, ..., c_n)$ is false.

How will we interpret the constant symbols in $\mathcal{A}$? Something like $c^{\mathcal{A}} = c$ seems right...

Issue: what if we have two *different* constant symbols c, d in $\mathcal{S}$ and T contains the sentence $c \approx d$. Then a model $\mathcal{A}$ for T will necessarily satisfy $\mathcal{A} \models c \approx d$, which means that $c^{\mathcal{A}} = d^{\mathcal{A}}$ (actually equal as members of the set A). So rather than interpreting $c^{\mathcal{A}} = c$ and $d^{\mathcal{A}} = d$, we will want to *identify* c, d as the same element of our domain. We will do this by taking a quotient. Also, we *do* want that $c \approx c$ will be in T, for any constant symbol c.

Ignoring this issue for now, suppose we defined a model $\mathcal{A}$ as above. Then by its definition for any atomic sentence φ in T, $\mathcal{A} \models \varphi$. We want $\mathcal{A}$ to satisfy all sentences in T. Let φ be in T. We will want to prove inductively that $\mathcal{A} \models \varphi$ for any φ in T.

One inductive instance is: $\varphi = \neg\psi$ is in T. Then by our non-contradiction assumption, ψ is not in T. We *would like* to conclude that $\psi^{\mathcal{A}} = 0$, in which case we deduce that $\varphi^{\mathcal{A}} = 1$, as desired. To do so, we need to carry out a stronger inductive hypothesis:

[Better inductive hypothesis] For any φ, $\varphi^{\mathcal{A}}$ is true if and only if φ is in T.

Another inductive instance will be: $\varphi = \psi_1 \wedge \psi_2$. If we know that ψ_1, ψ_2 are both true in $\mathcal{A}$, then we are good. But how do we know that? If ψ_1 and ψ_2 are both in T, then the inductive assumption will tell us that they are true in $\mathcal{A}$. Similarly, if $\psi_1 \wedge \psi_2 \notin T$, we would want to conclude that it is false in $\mathcal{A}$, which we could do if we knew that either ψ_1

or ψ_2 are not in T (in which case their negations are in T). In conclusion, we will want T to satisfy that:

$$\text{[Condition on } T\text{] } \psi_1 \wedge \psi_2 \text{ is in } T \text{ if and only if } \psi_1, \psi_2 \text{ are both in } T.$$

This seems reasonable... And it is true in case $T = \mathrm{Th}(\mathcal{A})$. So it *must* be true for T in order for T to have a model.

What about the quantifier stages (of the inductive construction of a formula). Suppose $\varphi = (\forall x)\psi$. If φ is in T, in order to (inductively) conclude that φ is true in $\mathcal{A}$, we would want the sentences $\psi[c]$ to be in T for each constant symbol c. [Recall from Pset 4 that $\psi[c]$ is the formula we get by replacing each free occurrence of x in ψ with the constant symbol c. It is helpful for intuition to assume that all occurrences of x are free in which case we simply replace each x with c.] If this is true, then inductively we know that $\psi[c]$ is true in $\mathcal{A}$, which means that $\mathcal{A} \models \psi(c^{\mathcal{A}})$ (by Pset 4), and so since the members of A are just these constant symbols, $(\forall x)\psi$ will be true in $\mathcal{A}$...

Finally, (essentially the most important case), for the other direction, if $(\forall x)\psi$ is not in T. (Equivalently, $\neg(\forall x)\psi$ is in T, which is an existential quantifier: $(\exists x)\neg\psi$). To conclude that $(\forall x)\psi$ fails in $\mathcal{A}$, we need some $a \in A$ so that $\psi^{\mathcal{A}}(a)$ fails. That is, *we need to have some constant symbol c for which $\psi[c] \notin T$.*

This is often referred to as Henkin's condition. Note that such condition implicitly makes us have infinitely many constant symbols in the language.

Under the conditions we just collected, we are in position to prove the existence of a model! (With still an extra simplifying assumption.)

Theorem 6.2. Let $\mathcal{S}$ be a vocabulary with no function symbols *other than constants*. (That is, the only function symbols are 0-ary function symbols.) Let T be a set of sentences in the language satisfying the following conditions.

(1) $[\neg]$ For any sentence φ: $\varphi \in T$ if and only if $\neg\varphi \notin T$.
(2) $[\wedge]$ For any sentences ψ_1, ψ_2: $\psi_1 \wedge \psi_2 \in T$ if and only if both ψ_1, ψ_2 are in T.
(3) $[\forall]$ For any sentence $\psi(x)$: $(\forall x)\psi \in T$ if and only if $\psi[c] \in T$ for any constant symbol c in $\mathcal{S}$.
(4) $[\approx]$ For any constant symbols c, d, e from $\mathcal{S}$:
    - $c \approx c \in T$;
    - if $c \approx d \in T$ then $d \approx c \in T$;
    - if $c \approx d \in T$ and $d \approx e \in T$ then $c \approx e \in T$.

    Furthermore, given any n-ary relation symbol R and constant symbols $c_1, ..., c_n$ and $d_1, ..., d_n$, if $R(c_1, ..., c_n)$ is in T and $c_i \approx d_i \in T$ for $i = 1, ..., n$, then also $R(d_1, ..., d_n) \in T$.

Then there is a structure $\mathcal{A}$ satisfying T, $\mathcal{A} \models T$. (Note that condition (1) implies that T is complete, so $\mathcal{A} \models T$ is equivalent to saying $\mathrm{Th}(\mathcal{A}) = T$.)

Remark 6.3. Recall that in Pset 4 you considered substituting a constant c in place of a variable x to turn a formula $\varphi(x)$ (with possibly x as a free variable) into a sentence $\varphi[c]$.

The same idea works with more variables. Given a formula $\varphi(x_1, ..., x_n)$ (so $x_1, ..., x_n$ is a list containing all free variables of φ), and constant symbols $c_1, ..., c_n$ (not necessarily different), let $\varphi[c_1, ..., c_n]$ be the result of replacing every *free* occurrence of x_i with c_i. Formally, this can be defined inductively along the construction of φ.

As always, it is best to assume that we "do not recycle variables". So if we think of φ as $\varphi(x_1, ..., x_n)$, we assume that $x_1, ..., x_n$ *do not* appear in any quantifier in φ (so if they do appear they are free variables). Then $\varphi[c_1, ..., c_n]$ is literally just replacing each appearance of x_i by c_i.

Similarly, we may "transform" a formula $\varphi(x_1, ..., x_n)$ into a formula with 1 variable $\varphi(x_1)[c_2, ..., c_n]$.

Remark 6.4. Henkin's condition follows from (3): if $\neg(\forall x)\psi \in T$ then there is some constant symbol c for which $\neg\psi[c] \in T$ – a witness for the existential statement.

If we were to use the existential quantifier $\exists$ instead of the universal quantifier $\forall$, we would phrase the Henkin condition as follows: if $(\exists x)\xi \in T$ then there is some constant symbol c so that $\xi[c] \in T$.

As discussed before, we define a model $\mathcal{A}$ using the constant symbols as elements. Since $\approx^{\mathcal{A}}$ must be interpreted as true equality, if $c \approx d \in T$ (so we want it to be a true statement in $\mathcal{A}$) we cannot introduce c and d as different members of A. Instead, the members of A will be equivalence classes.

Let C be the set of all constant symbols in $\mathcal{S}$. Define a relation E on C by

$$c\ E\ d \iff c \approx d \in T.$$

By condition (4), E is an equivalence relation on C. Define $A = \{[c]_E : c \in C\}$, the quotient space. Note that we have the "projection map" $C \to A$ sending $c \mapsto [c]_E$. We will often write $[c]$ instead of $[c]_E$, as E is fixed (and is the only equivalence we consider) from now until the end of the proof.

Remark 6.5. We are making *so many* assumptions, so you may ask, why not replace the "equivalence relation condition" in (4) with the simpler:

(4') For any constant symbols c, d from $\mathcal{S}$, $c \approx d \in T$ if and only if $c = d$ (they are the same symbol).

The reason is that, despite this assumption seeming much more reasonable than the many other assumptions we are making, we will in fact be able to realize all other assumptions, but in the natural way to do so (4') will fail, and (4) is the best we can hope for. (If you think of the way the "Henkin condition" works, you see it gives many many constant symbol, without paying much thought to whether or not they must be equal...)

However, for the purpose of intuition for the coming proof, you can often switch (4) to (4') to have a clearer mental picture. In this case, we do not need the quotient space, and simply $A = C$.

We now continue to define the model $\mathcal{A}$. Given an n-ary relation symbol R and $a_1, ..., a_n$ from A, we must decide whether $(a_1, ..., a_n)$ is in $R^{\mathcal{A}}$ or not. By definition, $a_i = [c_i]_E$ for some constant symbols $c_1, ..., c_n$. Define

$$(a_1, ..., a_n) \in R^{\mathcal{A}} \iff R(c_1, ..., c_n) \in T.$$

Exercise 6.6. Prove that this is "well defined". That is, if we present $a_i = [d_i]$ for constant symbols $d_1, ..., d_n$, then we get the same definition. Note that $[c_i] = a_i = [d_i]$ implies that $c_i \approx d_i \in T$, by definition of E

So we now have a well defined structure $\mathcal{A}$ for the vocabulary $\mathcal{S}$. To conclude Theorem 6.2 we are left to prove that any sentence φ, $\mathcal{A} \models \varphi \iff \varphi \in T$. More generally:

Claim 6.7. For any formula $\varphi(x_1,...,x_n)$ and any $a_1,...,a_n$ in A, if $a_i=[c_i]$ then

$$\mathcal{A}\models\varphi(a_1,...,a_n) \iff \varphi[c_1,...,c_n]\in T.$$

We prove this by induction. Begin with the atomic case.

Suppose $\varphi(x,y)=x\approx y$. Then $\varphi^{\mathcal{A}}([c_1],[c_2])=1 \iff [c_1]=[c_2]$ (actual same object), by the definition of structures. The latter is true if and only if $c_1\approx c_2\in T$, by the definition of the equivalence relation E. Note that $c_1\approx c_2$ is the sentence $\varphi[c_1,c_2]$ ($\varphi(x,y)$ with x,y substituted by c_1,c_2), so we are done. Note again that this does not depend on the choice of "representatives of the equivalence classes".

For the general atomic formula case, suppose $\varphi(x_1,...,x_n)=R(x_1,...,x_n)$, $a_1,...,a_n\in A$, $a_i=[c_i]$. By definition,

$$\varphi^{\mathcal{A}}(a_1,...,a_n)=1 \iff (a_1,...,a_n)\in R^{\mathcal{A}} \iff R(c_1,...,c_n)\in T,$$

as required (by the claim), as $R(c_1,...,c_n)=\varphi[c_1,...,c_n]$.

Next, we deal with the connectives and quantifiers. This is just like we did (informally) before. Now that we have all the assumptions in place, the proofs go through.

Suppose $\varphi=\psi_1\wedge\psi_2$, and $x_1,...,x_n$ is a sequence of variables containing all the free variables of φ (so also of each ψ_1,ψ_2). By assumption (2) in the theorem, $\varphi\in T$ if and only if both ψ_1,ψ_2 are in T. By the inductive hypotheses (the claim for ψ_1,ψ_2), the latter is true if and only if $\psi_1^{\mathcal{A}}=1$ and $\psi_2^{\mathcal{A}}=1$, which (by the definition of truth values in a structure) is true if and only if $\varphi^{\mathcal{A}}=1$. This concludes the claim for φ.

Suppose $\varphi=\neg\psi$, $x_1,...,x_n$ is a sequence of variables containing all the free variables of φ (so also of ψ). By assumption (1) of the theorem, $\varphi\in T$ if and only if $\psi\notin T$ which is true (by the claim for ψ) if and only if $\psi^{\mathcal{A}}=0$ which is true (by the definition of interpretations in a structure) if and only if $\varphi^{\mathcal{A}}=1$.

Finally, assume $\varphi=(\forall x)\psi$. Suppose first $\varphi\in T$. We want to conclude that $\varphi^{\mathcal{A}}=1$, that is, that $\psi^{\mathcal{A}}(a)=1$ for any $a\in A$. By assumption (3) in the theorem, given any constant symbol c, the sentence $\psi[c]$ is in T. Take any $a\in A$. $a=[c]$ for some constant symbol c in the language. Since $\psi[c]\in T$, by the inductive assumption (the claim for ψ) $\psi^{\mathcal{A}}(a)=1$. As required.

Suppose now $\varphi\notin T$. We want to conclude that $\varphi^{\mathcal{A}}=0$. That is, we need to find some $a\in A$ so that $\psi^{\mathcal{A}}(a)=0$. By assumption (3) in the theorem, *there is some constant symbol* c so that $\psi[c]\notin T$. Then by (1) necessarily $\neg\psi[c]\in T$. By the inductive assumption, we know that $\neg\psi[c]^{\mathcal{A}}=1$, that is, $\psi[c]^{\mathcal{A}}=0$, as required.

This concludes the proof of Theorem 6.2. Call a theory satisfying the conditions of the theorem a **Henkin theory**. These are a lot of assumptions. Nevertheless, we will be able to use this idea to find models for an arbitrary theory, as long as it is satisfiable. The vague idea:

$$\text{Some theory } T \rightsquigarrow \text{a Henkin theory } T' \text{ “extending } T\text{”} \rightsquigarrow \text{a model } \mathcal{A}.$$

The main work ahead of us is to justify the first step. In particular, as not every theory is satisfiable, we still need some way of determining when such "extention" is possible.

6.2. **Adding Henkin conditions.** Note that all the conditions tell us that *more things need to be in* T. Generally, there is no reason for T to satisfy any of the Henkin conditions. The key idea is to keep *expanding* the theory, to meet the Henkin conditions.

Let us consider the negation $\neg$ case. It says two things. One is that no formula and its negation both appear in T. This is clearly necessary in order for T to be satisfiable.

The second thing (the completeness assumption, is that either a formula or its negation is in T. Generally, a theory T may not have either $\varphi, \neg\varphi$ in it. It may even be that neither is a logical consequence of T. In that case, we'll just add one!

Claim 6.8. Suppose T is a satisfiable theory and φ is a sentence. Then one of the following two: $T \cup \{\varphi\}$, $T \cup \{\neg\varphi\}$, is satisfiable. (Maybe both.)

Proof. Let $\mathcal{A}$ be a model for T. If $\varphi^{\mathcal{A}} = 1$, then $\mathcal{A}$ is a model for $T \cup \{\varphi\}$. If $\varphi^{\mathcal{A}} = 0$, then $\mathcal{A}$ is a model for $T \cup \{\neg\}$. $\square$

So we can enforce (at least one instance of) condition $\neg$ without changing the main question: is our theory satisfiable or not.

What about the $\wedge$ case. Again it says two things: one is that if ψ_1, ψ_2 are in T, *then* $\psi_1 \wedge \psi_2$ are in T, and the other is that if $\psi_1 \wedge \psi_2$ are in T, *then* both ψ_1, ψ_2 are in T. Both are *closure conditions* for T.

Claim 6.9. Let T be a satisfiable theory.

- If ψ_1, ψ_2 are in T, then $T \cup \{\psi_1 \wedge \psi_2\}$ is satisfiable.
- If $\psi_1 \wedge \psi_2 \in T$, then $T \cup \{\psi_1, \psi_2\}$ is satisfiable.

Proof. Let $\mathcal{A}$ be a model for T. If ψ_1, ψ_2 are in T, then necessarily both are true in $\mathcal{A}$, and so $\psi_1 \wedge \psi_2$ is true in $\mathcal{A}$. If $\psi_1 \wedge \psi_2$ is in T, then it is true in $\mathcal{A}$, and so both ψ_1, ψ_2 are true in $\mathcal{A}$. $\square$

Again, by adding some formulas to T, we are able to enforce (one instance of) condition $\wedge$, without changing the satisfiability.

The $\approx$ case is also easy to deal with. For example, suppose $c \approx d$ is in T for some constant symbols c, d. Then we want $d \approx c$ to be in T, so we will simply add it. Note that a structure satisfies T if and only if it satisfies $T \cup \{d \approx c\}$, so again we do not change satisfiability.

Let us deal with the quantifier case $\exists$ now.

The key issue is the following. Suppose that $(\exists x)\varphi$ is in T. We want to conclude that there is a constant symbol c so that $\varphi[c] \in T$.[4] However, even if T is satisfiable, *it may be impossible* to find a constant symbol c in our language for which $T \cup \varphi[c]$ is also satisfiable. [There may not even be constant symbols in the language...]

The solution is to not only expand T but *also expand the language* by adding a *new constant symbol* c and then add $\varphi[c]$ to T, creating a theory in a larger vocabulary $\mathcal{S} \cup \{c\}$.

Claim 6.10. Let T be a theory for a signature $\mathcal{S}$. Assume T is satisfiable and $(\exists x)\varphi$ is in T. Let c be a constant symbol not in $\mathcal{S}$, and $\mathcal{S}' = \mathcal{S} \cup \{c\}$. Then $T \cup \{\varphi[c]\}$ is a satisfiable $\mathcal{S}'$ theory.

Proof. Let $\mathcal{A}$ be a model for T. Since $\mathcal{A} \models (\exists x)\varphi$, there is some $a \in A$ so that $\mathcal{A} \models \varphi(a)$. Expand $\mathcal{A}$ to a model $\mathcal{A}'$ for $\mathcal{S}'$ by interpreting $c^{\mathcal{A}'} = a$. Then $\mathcal{A}' \models T$ (see Pset 4, the not-for-submission-question) and $\mathcal{A}' \models \varphi[c]$ (see Pset 4 question 5 part (3)). $\square$

[4] This is the same as saying, suppose $\neg(\forall x)\xi \in T$, we want to find some c for which $\neg\xi[c] \in T$.

The other direction is easy: if $\varphi(x)$ is a formula, c a constant in the language, and $\varphi[c]$ is in T, we want to conclude that $(\exists x)\varphi$ is in T. As before, we may just add it: for any structure $\mathcal{A}$, $\mathcal{A} \models T \implies \mathcal{A} \models T \cup \{(\exists x)\varphi\}$. So, *if T is satisfiable, so is $T \cup \{(\exists x)\varphi\}$.*

Nevertheless, let us view this easier direction in another way, as a "closure rule for $\forall$". The direction $\varphi[c] \in T$ imlpies that $(\exists x)$

Claim 6.11. Let T be a theory for a signature $\mathcal{S}$ containing a constant symbol c. Assume T is satisfiable and $\neg(\exists x)\varphi$ is in T. Then $T \cup \{\neg\varphi[c]\}$ is satisfiable.

Proof. Any $\mathcal{S}$-structure $\mathcal{A}$ which is a model for T must satisfy $\varphi^{\mathcal{A}}(a) = 0$ for any $a \in A$, so in particular $\varphi^{\mathcal{A}}(c^{\mathcal{A}}) = 0$, so $\mathcal{A} \models \neg\varphi[c]$. (See Pset 4, the not-for-submission-question.) $\square$

In conclusion: assume T is satisfiable. It seems that we can expand it to some theory, in a larger signature, so that it will satisfy all the conditions for being a Henkin theory. The idea is to repeatedly apply the "one step" operations we described above, infinitely many times.

Before, let us rid of one, less significant, assumption we have made in Theorem 6.2: that there are no function symbols with arity ≥ 1.

6.3. **Dealing with function symbols.** So far, for simplicity, we assumed that there were no non-constant function symbols. First, where did we even use this assumption? At the very first step of the proof of Claim 6.7, when we dealt with atomic formulas. We considered only atomic formulas of the form $R(x_1, ..., x_n)$ where x_i is a variable (or a constant symbol). On "the other side", we were asking if $R(c_1, ..., c_n) \in T$ for some constant symbols $c_1, ..., c_n$.

Generally, an atomic formula $\varphi(x_1, ..., x_n)$ is of the form $R(t_1, ..., t_k)$ where $t_1, ..., t_k$ are terms whose variables are contained in $x_1, ..., x_n$. Similarly, if we have function symbols and constant symbols we get many terms t with no free variables (which we will call **constant terms**). For example, $t'_i = t_i[c_1, ..., c_n]$ is a term with no variable, and we would like to ask if $R(t'_1, ..., t'_k) \in T$ (this makes sense as $R(t'_1, ..., t'_k)$ is a sentence).

Remark 6.12. Recall that in Pset 4 you dealt with replacing a variable x_1 with a constant symbol c. Given a term $t(x_1, ..., x_n)$ we defined a term $t[c]$ with fewer variables $t[c](x_2, ..., x_n)$. Similarly given a formula $\varphi(x_1, ..., x_n)$ we defined $\varphi[c](x_2, ..., x_n)$. Similarly, we can replace several variables by several constants, constructing $t[c_1, ..., c_n]$ and $\varphi[c_1, ..., c_n]$ from $t(x_1, ..., x_n)$ and $\varphi(x_1, ..., x_n)$. In what follows, we will want to substitute a variable x with a *constant term* σ.

For example, in the language $+, \cdot, 0, 1$, consider the term $t(x, y) = (x + y) \cdot x$ and the formula $\varphi(x, y) = t(x, y) \approx 0$. Let $\sigma(x) = 1+1$, a constant term. How would we substitute x by σ?
- $t[\sigma](y) = ((1 + 1) + y) \cdot (1 + 1)$.
- $\varphi[\sigma](y) = t[\sigma](y) \approx 0 = ((1 + 1) + y) \cdot (1 + 1) \approx 0$.
Similarly we may define $t[\sigma_1, ..., \sigma_n]$ and $\varphi[\sigma_1, ..., \sigma_n]$ given a term $t(x_1, ..., x_n)$, a formula $\varphi(x_1, ..., x_n)$ and constant terms $\sigma_1, ..., \sigma_n$.

In order to prove that the structure we are constructing, the one whose domain is (equivalence classes of) constant symbols, we will need Claim 6.7 to still hold for any formula:

$$\mathcal{A} \models \varphi(t_1^{\mathcal{A}}(a_1, ..., a_n), ..., t_k^{\mathcal{A}}(a_1, ..., a_n)) \iff \varphi[t'_1, ..., t'_n] \in T.$$

That is, when we have function symbols, we have all sorts of constant terms which are not constant symbols, which correspond to more sentences (constant terms plugged in as variables). For these sentences, T has already "made up its mind" about it being true or false, and we need to model to agree with T.

Here are the main modifications of Theorem 6.2 in the case where we do have function symbols. (In particular, these modification define what it means to be a Henkin theory when there are function symbols.)

Condition (3) will be changed to

(3') – If $(\forall x)\psi \in T$ the for any *constant term* (not just constant symbol) c in the language, $\psi[c]$ is in T, and
– If $\neg(\forall x)\psi$ $((\exists x)\neg\psi)$ is in T, then there is some constant *symbol* c so that $\neg\psi[c] \in T$.

Conditions (1) and (2) remain the same.

Condition (4) will be replaced by

(4') For any constant *terms* c, d, e from $\mathcal{S}$:
– $c \approx c \in T$;
– if $c \approx d \in T$ then $d \approx c \in T$;
– if $c \approx d \in T$ and $d \approx e \in T$ then $c \approx e \in T$.

Furthermore, given any n-ary relation symbol R and constant *terms* $c_1, ..., c_n$ and $d_1, ..., d_n$, if $R(c_1, ..., c_n)$ is in T and $c_i \approx d_i \in T$ for $i = 1, ..., n$, then also $R(d_1, ..., d_n) \in T$.

Theorem 6.13. Let $\mathcal{S}$ be any vocabulary. Let T be a set of sentences in the language satisfying (1), (2), (3') and (4'). Then there is a model for T.

The model is constructed as in Theorem 6.2 above. The additional thing we need to do, to define the model, is to interpret the function symbols in $\mathcal{S}$.

Given an n-ary function symbol F and $a_1, ..., a_n$ in A, fix constant symbols $c_1, ..., c_n$ so that $a_i = [c_i]$, and define

$$F^{\mathcal{A}}(a_1, ..., a_n) = b,$$

if $b = [d]$ for a constant *symbol* d so that the sentence $F(c_1, ..., c_n) \approx d$ is in T.

Exercise 6.14. Show that $F^{\mathcal{A}}$ is well defined.

Note that one thing to show is that such a b exists. Why is that? Otherwise, if there is no constant symbol d such that $F(c_1, ..., c_n) \approx d$ is in T, then $(\exists x)F(c_1, ..., c_n) \approx x$ is *not* in T, so $(\forall x)\neg(F(c_1, ..., c_n \approx x))$ *is* in T, and therefore for any *constant term* d $\neg(F(c_1, ..., c_n) \approx d)$ is in T.

However, $d = F(c_1, ..., c_n)$ is itself a constant term, so we conclude that $\neg(d \approx d)$ is in T. By (4'), $d \approx d$ is in T as well. This contradicts (1), as we have φ and $\neg\varphi$ both in T for some sentence φ.

The rest of the exercise, and the proof of the theorem, is very similar to our proof of Theorem 6.2, and we skip it here.

Remark 6.15. You may at times forget about these function symbols. The following arguments will not change much due to these function symbols.

However, we will keep the "split condition (3)" as above. In a sense, it is in fact more natural to present it this way, as one "Henkin condition for $\exists$" and one "Henkin condition for $\forall$".

6.4. **Coding functions as relations.** [We didn't do this in class. This is a brief discussion on how such coding can be done and what needs to be proven to see that it in fact works to the fullest extent.]

There is a more general way in which one can reduce problems about a vocabulary with functions symbols to one without function symbols.

Let $\mathcal{S}$ be a vocabulary. For each n-ary function symbol F in $\mathcal{S}$, introduce an $n+1$-ary relation symbol R_F. Let $\mathcal{S}'$ be the vocabulary we get by replacing each function symbol F in $\mathcal{S}$ by R_F.

For each such F consider the sentence φ_F (in the language for $\mathcal{S}'$)

$$\varphi_F = \forall x_1, ..., \forall x_n \exists x_{n_1}[R(x_1, ..., x_n, x_{n+1}) \wedge \forall y(R(x_1, ..., x_n, y) \to x_{n+1} \approx y)].$$

Given a structures $\mathcal{A}$ for $\mathcal{S}$, define a structure $\mathcal{A}'$ for $\mathcal{S}'$ as follows. The universe A' of $\mathcal{A}'$ is A. If R is a relation symbol in $\mathcal{S}$, $\mathbb{R}^{\mathcal{A}} = \mathbb{R}^{\mathcal{A}'}$. If F is a function symbol in $\mathcal{S}$ then

$$R_F^{\mathcal{A}'} = \left\{(a_1, ..., a_n, a_{n_1}) : F^{\mathcal{A}}(a_1, ..., a_n) = a_{n+1}\right\}.$$

Then $\mathcal{A}'$ satisfies φ_F for each F in $\mathcal{S}$.

Similarly, given an $\mathcal{S}'$-structure $\mathcal{A}'$ satisfying φ_F for each $F \in S$, define an $\mathcal{S}$-structure $\mathcal{A}$ as follows: the universe A of $\mathcal{A}$ is A'. $R^{\mathcal{A}} = R^{\mathcal{A}'}$ for any relation symbol R in $\mathcal{S}$. Given a function symbol F in $\mathcal{S}$, define

$$F^{\mathcal{A}}(a_1, ..., a_n) = a_{n+1} \text{ if and only if } (a_1, ..., a_n, a_{n+1}) \in R_F^{\mathcal{A}'}.$$

The latter gives a well-defined function because $\mathcal{A}' \models \varphi_F$.

So there is a one-to-one correspondence between $\mathcal{S}$ structures and $\mathcal{S}'$ models for $\{\varphi_F : F \in \mathcal{S}\}$. Note that this correspondence *respects* the structures: $\mathcal{A} \simeq \mathcal{B}$ (as $\mathcal{S}$-structures) if and only if $\mathcal{A}' \simeq \mathcal{B}'$ (as $\mathcal{S}'$-structures).

This correspondences can be taken a step further. We can transform every $\mathcal{S}$ formula to an $\mathcal{S}'$ formula as follows. In $\mathcal{S}'$ there are no terms, other than the variables. We can however define for each term t in $\mathcal{S}$ a formula ψ_t in $\mathcal{S}'$ *implicitly defining* t. For example, if F is a binary function symbol, $t = F(x, y)$ for variables x, y, we define $\psi_t(x, y, z) = R_F(x, y, z)$. If we already defined ψ_{t_1}, ψ_{t_2} with variables x, y, z for $\mathcal{S}$-terms t_1, t_2, with variables x, y, and $t = F(t_1, t_2)$, then define $\psi_t(x, y, z) = \exists z_1 \exists z_2(\psi_{t_1}(x, y, z_1) \wedge \psi_{t_2}(x, y, z_2) \wedge \varphi_F(z_1, z_2, z))$. You can similar define $\psi(x_1, ..., x_n, x_{n+1}$ for any term $t(x_1, ..., x_n)$.

Suppose now we have an atomic formula in $\mathcal{S}$ of the form $\varphi(x, y) = P(t)$ for an unary relation symbol P, we will define $\varphi'(x, y) = \exists z(\psi_t(x, y, z) \wedge P(z))$. The point is that, for the correspondence described above, $\varphi(a, b)$ will be true in $\mathcal{A}$ if and only if $\varphi'(a, b)$ will be true in $\mathcal{A}'$. Similarly you can define φ' for any formula φ.

The upshot is the following: given an $\mathcal{S}$-theory T, let $T' = \{\varphi' : \varphi \in T\}$. Then $\mathcal{A} \models T$ if and only if $\mathcal{A}' \models T'$.

In particular, T is satisfiable if and only if T' is satisfiable.

This trick is very useful, and often used. We will often take the "no functions point of view" as well, just for the notational advantage of not having to deal with terms (other than variables and constants). However, this is *not* to say that function symbols should be discarded. The point of mathematical logic is *not* that things can be coded in this or that manner. The point is to study mathematical structures, and using functions in the language sometimes better represents these structures. For example, given models for the theory of vector spaces, in the language which we used to describe vector spaces, a substructure precisely coincides with a subspace in the vector-space sense. If function symbols are replaced by relation symbols, this natural correspondence fails.

7. Formal deductions

Back to our vague plan:

$$\text{Some theory } T \rightsquigarrow \text{a Henkin theory } T' \text{ ``extending } T\text{''} \rightsquigarrow \text{a model } \mathcal{A}.$$

We are now working towards the first step.

Recall the key Henkin theory conditions:

(1) [$\neg$] For any sentence φ: $\varphi \in T$ if and only if $\neg\varphi \notin T$.
(2) [$\wedge$] For any sentences ψ_1, ψ_2: $\psi_1 \wedge \psi_2 \in T$ if and only if both ψ_1, ψ_2 are in T.
(3) [$\forall$] If $\neg(\exists x)\psi \in T$ then for any *constant term* t in the language, $\neg\psi[t]$ is in T;
[$\exists$] If $(\exists x)\psi \in T$, then there is some constant *symbol* c so that $\psi[c] \in T$.
(4) [$\approx$] ...

(Note that (3) is equivalent to the previously stated (3'), using condition (1). It will be convenient to use the existential point of view now. Remember again that any formula is equivalent to one using only the logical symbols $\neg, \wedge, \exists, \approx$.)

As we discussed, we want to start with some theory T and keep *expanding* it (we will need to do so infinitely many times), with the hope of having a Henkin theory at the end.

It seems like this will work, and it will. Note however that this was all under the *assumption* that T is satisfiable. We still need to figure out *when* that is the case. We will talk about that soon.

In most cases it was clear how to extend T to satisfy (another instance of) one of the conditions. The one exception was the case for $\neg$. Assuming T is satisfiable, we don't necessarily know, and it may be difficult to determine, which of φ or $\neg\varphi$ may be added to T, while remaining satisfiable. All we know is that one works, and possibly both. This motivates us to talk about binary splitting trees.

7.1. **Trees.** [See board for pictures] We will consider finite and (countably) infinite trees. A tree is a structure of the form $(T, \sqsubset, r)$ where T is a set (the nodes of the tree), $r \in T$ is the room, and $\sqsubset$ is the relation of extension along the tree (existence of a branch between nodes). We assume that it is transitive:

- for a, b, c in T, $a \sqsubset b$ and $b \sqsubset c$ implies $a \sqsubset c$.

The tree has no loops:

- for a, b in T it is *not* the case that $a \sqsubset b$ and $b \sqsubset a$.

Everything extends the root:

- for any a in T, $r \sqsubset a$ or $r = a$.

And finally, when looking at the branch *below some given node in the tree*, the tree relation $\sqsubset$ is linear:

- for any a, b, c in T, if $a \sqsubset c$ and $b \sqsubset c$ then either $a \sqsubset b$ or $b \sqsubset a$.

Exercise 7.1. Write the axioms of being a tree in the language using one binary relation $\sqsubset$ and one constant symbol r.

Trees are studied in many ways, and it does make sense to study models for this theory. However, this will not be the point of view below, as the trees we consider are external.

Definition 7.2.
- Given $a \in T$, the **partial branch** below a is $\{b \in T : b \sqsubset a\}$.
- The **height** of a node $a \in T$ is the number of nodes below a. (So the height of the root is 0.)
- Say that $a \in T$ is a **leaf** if there is not b for which $a \sqsubset b$.
- If a is a leaf, we will call its partial branch a **branch**.
- For $a, b \in T$, say that b is an **immediate successor** of a in T if $a \sqsubset b$ and there is no c for which $a \sqsubset c \sqsubset b$.
- Say that a node $a \in T$ is a k**-splitting node** in T if it has k (distinct) immediate successors in T.
- A tree T is **finitely splitting** if every node is finitely splitting.
- A tree T is **binary splitting** if every node is k-splitting for $k \leq 2$.

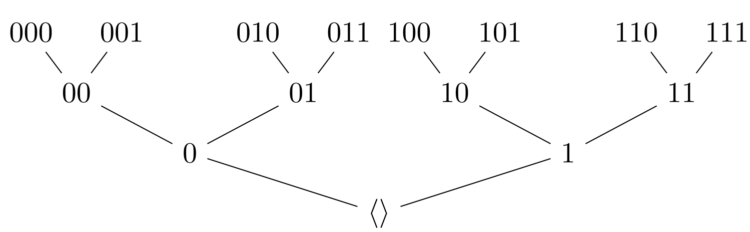
We will think of a node as some stage in a construction we are carrying. Going upwards in the tree will mean doing another step in the construction. That is, ensuring another Henkin condition. A (binary) *split* in the tree will precisely correspond to the $\neg$ case.

7.1.1. *The infinite binary tree.* Start with a root, and keep splitting all the way... A convenient way of representing this is using binary sequences. Recall that X^n is the set of all sequences $(x_1, ..., x_n)$ with x_i in X, and $X^{<\mathbb{N}} = \bigcup_{n \in \mathbb{N}} X^n$ is the set of all finite sequences from X. (A length 0 sequence is "the empty sequences" $<>$. X^0 is the set containing only the empty sequences $\{<>\}$.) The set $\{0,1\}^{<\mathbb{N}}$ is the set of all finite binary sequences.

Given two sequences σ, τ, say that τ **extends** σ, $\sigma \sqsubset \tau$, if σ is a subsequence of τ. That is, $\tau = \langle x_1, ..., x_n \rangle$ for some n, and $\sigma = \langle x_1, ..., x_m \rangle$ for some $m < n$.

For example, $< 010 >$ (strictly) extends $< 01 >$ and $< 0 >$. For the two sequences $< 010 >$ and $< 10 >$, neither one extends the other. In this case we say that they are **incomparable**.

Exercise 7.3. Check that $\{0,1\}^{<\mathbb{N}}$ with the relation $\sqsubset$ and the root $<>$, is a (binary splitting) tree.



Definition 7.4. Let $T, \sqsubset, r$ be a tree. A **subtree** is a "*downwards closed* subset of T". That is, $T' \subseteq T$ so that $r \in T'$ and for any $a \in T'$ if $b \in T$ and $b \sqsubset a$ then $b \in T'$ as well.

Example 7.5.
- The tree $\{0,1\}^{<n} = \bigcup_{k<n} \{0,1\}^k$, all binary sequences of length $< n$, is a sub-tree of the full binary tree.
- $\{0,1\}^n$ is not a sub-tree for $n > 0$.

- $\{<>, <0>, <1>, <11>, <00>, <001>, <000>\}$ is a sub-tree.
- $\{<>, <0>, <00>, <01>, <10>\}$ is not a sub-tree.

7.2. **Formal deductions.** We work with the connectives $\wedge, \neg$ and the quantifier $\exists$. Recall that the others can be expressed using these three. In particular, we think of $\neg(\exists x)\varphi$ as $(\forall x)\neg\varphi$.

Let $\mathcal{S}$ be a vocabulary and T a set of sentences. We expand $\mathcal{S}$ to $\mathcal{S}^+$ by adding infinitely many new constant symbols $c_0, c_1, c_2, \ldots$. (We assume that these symbols are *not* in $\mathcal{S}$).

Definition 7.6. A **deduction tree** for T is a finite tree Γ together with an assignment of an $\mathcal{S}^+$-sentence φ_a to each node $a \in \Gamma$, so that the following rules are satisfied. (You may think of the tree, and the assignment of sentences, as being build up recursively along the tree.)

(1) [Rule for $\neg$, "a split"] Given a node b in T, and *any* sentence φ, we can "split the tree in two" by adding nodes a, a', both immediate successors of b in the tree, where $\varphi_a = \varphi$ and $\varphi_{a'} = \neg\varphi$.

$$\begin{tikzcd} \varphi_a = \varphi & & \varphi_{a'} = \neg\varphi \\ & \text{b} \arrow[ul, dash] \arrow[ur, dash] & \end{tikzcd}$$

Otherwise, for every node a in the tree, one of the following holds:

(2) $\varphi_a \in T$ ("using an axioms").

(3) [Rule for $\wedge$]
- $\varphi_a = \psi_1 \wedge \psi_2$ where ψ_1, ψ_2 "appear below a"; that is, there are nodes b_1, b_2 below a in the tree so that $\psi_1 = \varphi_{b_1}$ and $\psi_2 = \varphi_{b_2}$;
- $\varphi_a = \psi$ where there is some $b \sqsubset a$ so that either $\varphi_b = \psi \wedge \theta$ for some θ, or $\varphi_b = \theta \wedge \psi$ for some θ.

(4) [Rule for $\exists$] $\varphi_a = \varphi[c]$ (substitution) where there is some $b \sqsubset a$ so that $\varphi_b = (\exists x)\varphi$ and c is a constant symbol in $\mathcal{S}^+ \setminus \mathcal{S}$ *which does not appear in any φ_c for $c \sqsubset a$.*

(5) [Rule for $\forall$] $\varphi_a = \neg\varphi[t]$ where there is some $b \sqsubset a$ so that $\varphi_b = \neg(\exists x)\varphi$ and t is *any constant term* (a term with no variables).

(6) [Rules for $\approx$]
- $\varphi_a = t \approx t$ where t is any constant term;
- $\varphi_a = \varphi[s]$, where there are some $b, c \sqsubset a$ for which $\varphi_b = s \approx t$ and $\varphi_c = \varphi[t]$, where s, t are constant terms. (Here φ is any formula with one free variable and $\varphi[s], \varphi[t]$ are substitutions.)

The above rules should be read as "we can assign to the node a a sentence φ_a as indicated, assuming that the following conditions are true (involving nodes below a)".

Remark 7.7. A deduction tree is a binary tree. In fact it can always be viewed as a finite subtree of the full binary tree.

Remark 7.8. The above rules are all syntactic manipulations on sentences. You can write a computer program which takes T as an input, and the program constructs a deduction tree by repeatedly applying these rules in some way.

Remark 7.9. All these rules are supposed to be "obviously true" rules of deduction.

For example, if we have some theory T and we can prove from it ψ_1 and ψ_2, then we know that $\psi_1 \wedge \psi_2$ is a consequence of the theory.

The $\neg$ rule can be seen as a "proof by conradiction". Say we want to argue that φ is true. We *split* by saying: either φ is true or is false. We then continue to argue using the assumption for contradiction $\neg\varphi$, hoping to reach a contradiction at the end, leaving us with φ as the only viable option.

The $\forall$ rule is also natural. Suppose we have a theory talking about a binary relation E representing a graph. Then $\neg(\exists x)x\ E\ x$ says no vertex is connected to itself (there are no loops). In particular, if we have a constant term t, then in any structure the interpretation of t will simply be a vertex. If all vertices are not connected to themselves, then necessarily it is true for the vertex which is the interpretation of t. So we conclude that the sentence $\neg(t\ E\ t)$ must be a consequence of $\neg(\exists x)x\ E\ x$.

Finally, the $\exists$ rule is also something we do naturally when proving: if we assume some existential statement $\exists x\varphi$, then we fix some arbitrary name for a witness.
For example, assume we have a group and we assume that "there exists an element of order 3" and "there exists an element of order 2", and we want to prove that "there exists an element of order 6".
The natural way to do this is as follows: using the first "exists assumption", *fix some element* a of order 3, using the second assumption, *fix some element* b of order 2.
Now study the element $a \cdot b$ and find out what its order has to be.

Definition 7.10. Given a deduction tree Γ, say that a branch in Γ **contains a contradiction** if there are two nodes a, b in this branch so that $\varphi_a = \neg\varphi_b$. Say that Γ is a **deduction for a contradiction** if *every* branch in Γ contains a contradiction. Equivalently: if for every leaf c in Γ, there are $a, b \sqsubseteq c$ for which $\varphi_a = \neg\varphi_b$.

Definition 7.11. Say that a theory T is **inconsistent** if there is a deduction tree for T which is a deduction tree for a contradiction. Denote this by $T \vdash \perp$.

Our goal is to prove (one form of Godel's completeness theorem):

$$T \vdash \perp \text{ if and only if } T \models \perp.$$

That is, T is inconsistent (*a syntactic condition*) if and only if T is not satisfiable (*a semantic condition*).

Let us start by talking more about deduction trees and giving some examples.

Example 7.12. Let φ be a sentence. Then the theory $T = \{\varphi, \neg\varphi\}$ is inconsistent. A contradiction deduction tree is simply:

$$\begin{array}{c} \neg\varphi \\ | \\ \varphi \end{array}$$

More precisely, we can take the tree as $\{<>, <0>\}$, $\varphi_{<>} = \varphi$ (using the "axiom rule") and $\varphi_{<0>} = \neg\varphi$ (using the "axiom rule").

Example 7.13. Let $T = \{(\exists x)\neg(x \approx x)\}$. A contradiction deduction tree from T is:

$$\begin{array}{c} c \approx c \\ | \\ \neg(c \approx c) \\ | \\ (\exists x)\neg(x \approx x) \end{array}$$

At the root we used the axiom rule. Then we used the $\exists$ rule: here c is a new constant, not appearing so far. Finally we used the $\approx$ rule to add $c \approx c$.

Note that this:

$$\begin{array}{c} \neg(c \approx c) \\ | \\ c \approx c \\ | \\ (\exists x)\neg(x \approx x) \end{array}$$

is *not* a legit deduction tree, as we cannot apply the $\exists$ rule using c, as c already appeared.

Example 7.14. Let φ be some sentence. $T = \{\varphi \wedge \neg\varphi\}$. The following a deduction of a contradiction from T.

$$\begin{array}{c} \neg\varphi \\ | \\ \varphi \\ | \\ \varphi \wedge \neg\varphi \end{array}$$

In the second and third steps we used the $\wedge$ rules, both applied to the root as the φ_b.

Example 7.15. Consider the follow variation of the $\forall$ rule (which can be seen as the "contrapositive of the $\exists$ rule):

- [$\forall'$ rule] We can write $\varphi_a = (\exists x)\varphi$ if $\varphi(x)$ is a formula with one free variable x and there is some $b \sqsubset a$ with $\varphi_b = \varphi[t]$ for some constant term t.

Then in fact, given the other rules, this rule and our $\forall$ rule are equivalent, in the following sense.

How can we "deduce" the $\forall$ rule from this rule? Suppose we have some $\varphi_b = \neg(\exists x)\varphi$ and we want to use the $\forall$ rule to add $\neg\varphi[t]$ above b. Instead, do as follows:

$$\begin{array}{ccc} (\exists x)\varphi & & \\ | & & \\ \varphi[t] & & \neg\varphi[t] \\ & \diagdown \quad \diagup & \\ & \neg(\exists x)\varphi & \end{array}$$

First we split according to the $\neg$ rule, applied for the sentence $\varphi[t]$. Then we used the $\forall'$ rule to conclude $(\exists x)\varphi$ from $\varphi[t]$. Now the left branch contains a contradiction, and we may continue with the right branch as if we used the $\forall$ rule.

On the other hand, using our standard rules, how can we use the natural looking $\forall'$ rule? Suppose we have some $\varphi_b = \varphi[t]$ for some constant term t and formula $\varphi(x)$. We want to conclude $(\exists x)\varphi$.

$$\begin{array}{ccc} & & \neg\varphi[t] \\ & & | \\ (\exists x)\varphi & & \neg(\exists x)\varphi \\ & \diagdown \quad \diagup & \\ & \varphi[t] & \end{array}$$

First we split, then we used our (usual) $\forall$ rule. The right branch contains a contradiction, and we may continue to "argue along the left branch" as if we used the $\forall'$ rule.

Definition 7.16. Let T be a theory and φ a sentence. Say that T **proves** φ (or φ is a **formal consequence** of T), denoted $T \vdash \varphi$, if $T \cup \{\neg\varphi\} \vdash \perp$. (That is, we prove φ from T by "assuming towards a contradiction" that φ fails, and reaching a contradiction.)

Example 7.17. Let ψ_1, ψ_2 be any sentences. Then $\{\psi_1, \psi_2\} \vdash \psi_1 \wedge \psi_2$. We need to construct a deduction tree from $\{\psi_1, \psi_2, \neg(\psi_1 \wedge \psi_2)\}$ so that every branch has a contradiction. We may do that as follows:

$$\begin{array}{c} \psi_1 \wedge \psi_2 \\ | \\ \psi_2 \\ | \\ \psi_1 \\ | \\ \neg(\psi_1 \wedge \psi_2) \end{array}$$

We are going towards the completeness theorem, which will tell us that if $T \models \varphi$ then in fact $T \vdash \varphi$. That is, if something is necessarily true (a semantic question) then we can formally prove it (a syntactic question).

First we note that the other directly is clearly true, since in our formal deductions we only do "obviously true" steps.

Theorem 7.18 (Soundness for $\vdash$)**.** If $T \vdash \varphi$ then $T \models \varphi$.

In particular, if $T \vdash \perp$ (T is *inconsistent*, it proves a contradiction), then $T \models \perp$: it is unsatisfiable, it has no model.

Remark 7.19. It suffices to prove the theorem for the case $\varphi = \perp$, since we can replace T with $T \cup \{\neg\varphi\}$.

Proof. It suffices to prove that if T *is* satisfiable, then $T \not\vdash \perp$, there is no proof of a contradiction from T. The key lemma is the following.

Lemma 7.20. Suppose T is satisfiable. Let Γ be a deduction tree from T. Then there is a branch (at least one) which is satisfiable. That is, there is some leaf a in Γ so that the theory $T' = \{\varphi_b : b \sqsubseteq a\}$ is satisfiable.

The proof is by induction on the construction of a deduction tree, along the allowable steps to add a node a and sentence φ_a.

Suppose Γ is a deduction tree, c is leaf in Γ so that $\{\varphi_b : b \sqsubseteq c\}$ is satisfiable, and Γ' is an extension of Γ according to one of the deduction rules.

Fix a model $\mathcal{A}$ for $\{\varphi_b : b \sqsubseteq c\}$. $\mathcal{A}$ is a model for a signature $\mathcal{S}'$ where $\mathcal{S}'$ contains $\mathcal{S}$ as well as finitely many of the constants $c_0, c_1, ...,$ those appearing in $\{\varphi_b : b \sqsubseteq c\}$.

Note that if c is a leaf in Γ' as well, then there is nothing to prove, $\{\varphi_b : b \sqsubseteq c\}$ is satisfiable, and c is still a leaf. The interesting case therefore is when we add a node above c.

We now need to consider all the cases.
One option is the split: Γ' is obtained from Γ by adding two immediate successors a, a' above c, where $\varphi_a = \theta$ and $\varphi_{a'} = \neg\theta$, for some sentence θ.
Both a, a' are now leafs. We need to show that either $\{\varphi_b : b \sqsubseteq a'\} = \{\varphi_b : b \sqsubseteq c\} \cup \{\varphi_{a'}\}$, or $\{\varphi_b : b \sqsubseteq a\} = \{\varphi_b : b \sqsubseteq c\} \cup \{\varphi_a\}$, is satisfiable.
This is precisely the content of Lemma 6.8. Specifically: either θ or $\neg\theta$ must be true in $\mathcal{A}$. For the rest of the cases in Definition 7.6, we also proved this already, in Section 6.2.

For example, if we used the $\exists$ rule: $\varphi_a = \psi[c_k]$ where $\varphi_b = (\exists x)\psi$ for some $b \sqsubseteq a$ and c_k does not appear in any of the formulas $\{\varphi_e : e \sqsubset a\} = \{\varphi_e : e \sqsubseteq c\}$.

Since $\mathcal{A} \models \varphi_b = (\exists x)\psi$, there is some $a \in A$ so that $\psi^{\mathcal{A}}(a) = 1$. Then, as we have seen, we may expand $\mathcal{A}$ to $\mathcal{A}^+$, whose signature also includes the constant symbol c_k, by defining $c_k^{\mathcal{A}+} = a$, and this way we have that $\mathcal{A}^+ \models \psi(c_k^{\mathcal{A}^+})$, and so $\mathcal{A}^+ \models \psi[c_k]$ (see Pset 4), as required.

Suppose we used the $\forall$ rule: $\varphi_a = \psi[t]$ where t is a constant term (in the signature $\mathcal{S}^+ = \mathcal{S} \cup \{c_0, c_1, ...\}$), where there is some $b \sqsubset a$ so that $\varphi_b = (\forall x)\psi$.
If t is a term in the signature of $\mathcal{A}$ (which we called $\mathcal{S}'$), then necesssarily $\mathcal{A} \models \psi[t]$, since $\mathcal{A} \models (\forall x)\psi$. (Recall Pset 4.)
If t uses additional constant symbols, not in $\mathcal{S}'$, expand $\mathcal{A}$ to $\mathcal{A}^+$ as follows: fix some $a_0 \in A$ and define $c_l^{\mathcal{A}^+} = a_0$ for any c_l which appears in t but not in $\mathcal{S}'$. Since $\mathcal{A}^+ \models (\forall x)\psi$ (recall Pset 4), then necessarily $\mathcal{A}^+ \models \psi(t^{\mathcal{A}+})$ and so $\mathcal{A}^+ \models \psi[t]$ (recall Pset 4).

The other cases are easier. For example, if $\varphi_a = \varphi_{b_1} \wedge \varphi_{b_2}$ where $b_1, b_2 \sqsubset a$, then necessarily $\mathcal{A} \models \varphi_a$.

Finally, assume that T is satisfiable, and let Γ be a deduction tree. We need to show that not every branch contains a contradiction.
By the lemma, there is some branch which is satisfiable: there is a structure $\mathcal{A}$ satisfying all sentences φ_b for b in this branch. Since a model cannot satisfy both φ and $\neg\varphi$, for any sentence φ, this branch cannot contain a contradiction. □

7.3. **Infinite trees and branches.**

Definition 7.21. Given a tree $(\Gamma, \sqsubset, r)$, an **infinite branch** is a chain $r = t_0 \sqsubset t_1 \sqsubset t_2 \sqsubset$ ..., where t_{i+1} is an immediate successor of t_i in the tree.

Example 7.22.
- If T is a finite tree, there is no infinite branch in T.
- Any infinite binary sequence, $b = \langle e_0, e_1, e_2, ...\rangle$ where $e_i \in \{0, 1\}$, corresponds to an infinite bracnh in the full binary tree: let $t_k = \langle e_0, ..., e_k\rangle$.

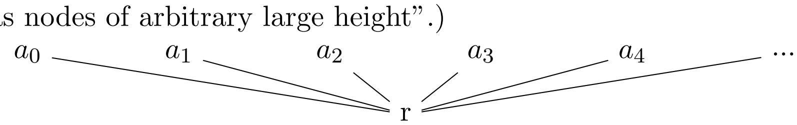
The following is an important combinatorial lemma.

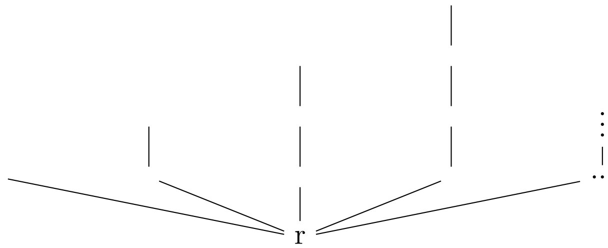
Lemma 7.23 (König's lemma)**.** Let T be a *finitely branching* tree. If T is infinite, then there is an infinite branch through T.

Remark 7.24. If T is finitely branching, the following are equivalent:

- T is infinite;
- T has nodes of arbitrary large height.

Remark 7.25. The assumption that T is finitely branching is necessary to conclude the existence of an infinite branch. (Even if the assumption "T is infinite" is replaced by "T has nodes of arbitrary large height".)





Proof of Konig's lemma. Approach: "go upwards". Find a node of height 1, then climb to a node of height 2, and continue... Can we do it? No, we may get stuck. We need to make better decisions to avoid getting stuck.

Given $a \in T$, let $T^a = \{b \in T : a \sqsubseteq b\}$. So $T^r = T$. We define the infinite branch recursively as follows.

Let $t_0 = r$ the root of the tree. Let $a_1, ..., a_k$ be the immediate successors of t_0 in the tree. (By assumption, we have only finitely many.)

Ask: is T^{a_i} finite, or infinite?

Since $T = T^r = \bigcup_{i=1}^k T^{a_i}$, it *cannot* be the case that each T^{a_i} is finite. (A finite union of finite sets is finite.)

So there must be some a_i for which T^{a_i} is infinite.

Let $t_1 = a_i$ where i is the smallest so that T^{a_i} is infinite.

Continue this way... Assume we have defined $t_0, ..., t_m$, in such a way that T^{t_l} is infinite for each $l = 0, ..., m$.

Let $a_1, ..., a_k$ be the immediate successors of t_m in T. Then $T^{t_m} = \bigcup_{i=1,...,k} T^{a_i}$.

Since T^{t_m} is infinite (the inductive assumption) then T^{a_i} must be infinite for some i.

Let $t_{m+1} = a_i$ where i is smallest so that T^{a_i} is infinite.

Note that, by definition, $t_m \sqsubset t_{m+1}$, as a_i is an immediate successor of t_m. In particular, $t_0, t_1, t_2, ...$ is an infinite branch through T. □

7.4. **Proof of the completeness theorem.**

Theorem 7.26 (Completeness for $\vdash$)**.** Let $\mathcal{S}$ be a countable signature. Suppose $T \models \varphi$. Then $T \vdash \varphi$.

That is, if something is always true (in terms of models) then we can formally prove it, using a few simple rules of deduction.

Remark 7.27. The theorem is true for any signature $\mathcal{S}$. The general proof follows similar ideas, but requires some familiarity with uncountable cardinals and ordinals.

Remark 7.28. It is enough to prove the theorem when $\varphi = \perp$, as T can be replaced with $T \cup \{\neg\varphi\}$. That is, it is enough to prove that: if T is *not satisfiable*, then there is a proof of a contradiction from T.
Equivalently: if there is not proof of contradiction from T, then T *has a model.*

Corollary 7.29. $\models \varphi$ (φ is logically valid) if and only if $\vdash \varphi$ (there is a formal proof of φ from the empty theory).

Towards the proof of the completeness theorem, fix a countable signature $\mathcal{S}$, a theory T, and let $\mathcal{S}^+$ be as above: $\mathcal{S}$ adjoined by infinitely many new constant symbols $c_0, c_1, ...$.

Remark 7.30. It will be convenient below to use the deduction rule $\forall'$ instead of the deduction rule $\forall$ rule in our formal deductions. We already saw that this does not change the provability notion $\vdash$.

Recall our Henkin conditions for a theory T: (1)-(4) at the beginning of Section 7. We may replace the $\forall$ Henkin condition in (3) with the $\forall'$ Henkin condition for T:

if $\psi[t] \in T$ for some formula $\psi(x)$ and *constant term* t, then $(\exists x)\psi \in T$ as well.

Exercise 7.31. Show that the $\forall$ and $\forall'$ Henkin rules are equivalent, given the other Henkin rules. That is: assume T satisfies (1), (2), (4) and the $\exists$ condition of (3). Show that T satisfies the $\forall$ Henkin condition if and only if T satisfies the $\forall'$ Henkin condition.

Assume that there is no proof of contradiction from T. We need to find a model for T.

The idea will be to build an ever-growing deduction tree, starting from T, attempting to find a contradiction. If we never do, at the end we will get a Henkin theory extending T, for which we can find a model.

More specifically, we will construct an *infinite* tree following our rules of deduction, so that an infinite branch will necessarily be a Henkin theory. This makes sense, as each "rule of deduction" is precisely a "closure rule for being a Henkin theory". So if we repeat these closure rules infinitely many times, we hope to have a theory which is closed under all the rules.

For example, given some sentence θ, we will want to have either θ or $\neg\theta$. If we make sure that at some level of the tree, all nodes split into $\theta, \neg\theta$, then any branch will have to make such choice.

Suppose one of the nodes in the branch is of the form $(\exists x)\varphi$. Then we would want a node above if of the form $\varphi[c]$ for some constant c.
Similarly, if somewhere along this branch ψ_1, ψ_2 appear, we will want at some point for $\psi_1 \wedge \psi_2$ to appear.
To make sure these (and other) things happen, some book-keeping needs to be done.

We split the natural numbers to 9 infinite subsets: numbers which are $0 \mod 9$, $1 \mod 9$,..., $8 \mod 9$.
We will define an increasing sequence of finite trees, $\Gamma_0, \Gamma_1, \Gamma_2, ...$, so that

- Γ_{n+1} is "an end-extension" of Γ_n: the new nodes in Γ_{n+1} are added as immediate successors of leafs of Γ_n.
- The definition of Γ_{n+1} from Γ_n will depend on $n \mod 10$. Essentially Γ_{n+1} will result by applying one of the rules of deductions to Γ_n (to each leaf).
- Each Γ_n is a deduction tree from T.

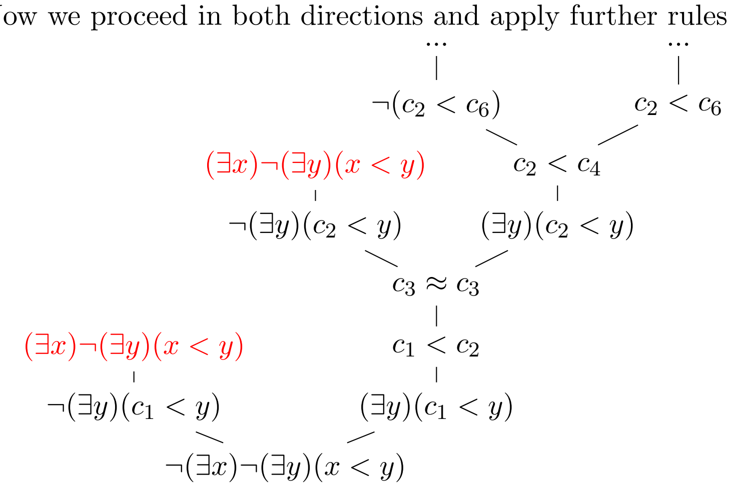
Essentially, the construction will just be to "randomly apply formal rules of deduction", without any rhyme or reason, trying to see if we can deduce a contradiction.

Another take: you provide the axioms T to a computer, and ask the computer to prove that they are contradictory. The computer does not know what the axioms are supposed to mean, so it just keeps formally applying deduction rules, checking if at any state a contradiction is reached.

Example 7.32. Here is vague sketch of how some steps of the construction may go, in case we start with the "no max" axiom for an order: $(\forall x)(\exists y)(x < y)$.

First, let us write this axiom using only $\exists, \neg, \wedge$: $\neg(\exists x)\neg(\exists y)(x < y)$.

First, we apply the axiom to the root.
This axiom says that for any x, it is not true that there is no y bigger than it. In particular, this should be true for the constant c_1 substituted for x. Indeed using the $\forall$ rules we can do this substitution. Let us instead use the $\forall'$ rule.
First we do a split using the sentence $\theta = (\exists y)(c_1 < y)$.
On the left side, we note that the formula $\neg(\exists y)(c_1 < y)$ is of the form $\psi[c_1]$ where $\psi(x) = \neg(\exists y)(x < y)$.
Using the $\forall'$ rule we conclude $(\exists x)\neg(\exists y)(x < y)$.
We now reached a contradiction on the left side (between the root and the leaf), and so we do not progress further in that direction.
On the ride side we may use the $\exists$ rule to add $c_1 < c_2$ (since c_2 has not appeared on this branch so far.
Maybe now we decide to write $c_3 \approx c_3$. Why not.
Again we can apply the $\forall$ rule to $\neg(\exists x)\neg(\exists y)(x < y)$. This is done below by using a split and the $\forall'$ rule. Note that in the $\forall$ rule we *are allowed* to substitute *any* constant term. Similarly when we split we can use *any* formula.
Next we may use the $\exists$ rule to add a witness $c_2 < c_4$. Note that we cannot use c_3 as it already appears in the branch.
Next we may want to take some θ (say, $\theta = c_2 < c_6$) and split into the two cases: $\theta, \neg\theta$.
Now we proceed in both directions and apply further rules of deduction...



Our construction proceeds as follows.
Fix an enumeration $\langle \theta_n : n = 0, 1, 2, ...\rangle$ of *all* sentences using the vocabulary $\mathcal{S}^+$.
(Recall: we proved that if $\mathcal{S}$ is countable then there are countably many formulas and terms for $\mathcal{S}$.)

Assume Γ_{n-1} is defined, is a finite tree with assignments φ_a for nodes $a \in \Gamma_{n-1}$.

- If α is a leaf in Γ_{n-1}, and the branch below α contains a contradiction, that is: there are $a, b \sqsubseteq \alpha$ with $\varphi_a = \neg\varphi_b$, then we *will not add anything above* α to Γ_n (and hereafter), so α remains a leaf.

  That is, when we reach a contradiction, along any branch, we stop the construction along that branch.

Let $\alpha_1, ..., \alpha_k$ be an enumeration of all the leaves of Γ_{n-1} for which "no contradiction was reached".

Let us deal first with the more interesting cases: $\neg$, $\exists$, and $\forall$. We will deal later with $\wedge$ and $\approx$.

- [Taking care of completeness] If $n = 1 \mod 9$, $n = 9 \cdot m + 1$ for some m, we do a "θ_m split": for each $i = 1, ..., k$, add to Γ_n two nodes a_i, a_i' which are immediate successors of α_i. Define $\varphi_{a_i} = \theta_m$ and $\varphi_{a_i'} = \neg\theta_m$.
- [Henkin witnesses for $\exists$] Suppose $n = 2 \mod 9$, $n = 9 \cdot m + 2$ for some m and that θ_m happens to be of the form $(\exists x)\psi$. Fix i and assume further that θ_m is φ_b for some $b \sqsubseteq \alpha_i$. ("$(\exists x)\psi$ appears in the branch below α_i".) Let j be the minimal natural number so that the constant symbol c_j does not appear in any formula along the branch below α_i.
  Add a node a to Γ_n, an immediate successor to α_i, and define $\varphi_a = \psi[c_j]$.
  If θ_m does not appear in the branch below α_i, define $\varphi_a = c_0 \approx c_0$.
- [$\forall'$ condition] Suppose $n = 3 \mod 9$, $n = 9 \cdot m + 3$ for some m and that θ_m happens to be of the form $\psi[t]$ where $\psi(x)$ is a formula and t is a constant term. Fix i and assume further that θ_m is φ_b for some $b \sqsubseteq \alpha_i$. ("$\psi[t]$ appears in the branch below α_i".)
  Add a node a to Γ_n, an immediate successor to α_i, and define $\varphi_a = (\exists x)\psi$.
  If θ_m does not appear in the branch below α_i, define $\varphi_a = c_0 \approx c_0$.

Next we deal with the 3 cases of the $\wedge$ condition.

- Suppose $n = 4 \mod 9$, $n = 9 \cdot m + 4$ for some m and that θ_m happens to be of the form $\psi \wedge \zeta$. Fix i.
  Add a node a to Γ_n, an immediate successor to α_i.
  If there is $b \sqsubseteq \alpha_i$ with $\varphi_b = \theta_m$, then define $\varphi_a = \psi$.
  Otherwise, define $\varphi_a = c_0 \approx c_0$.
- Suppose $n = 5 \mod 9$, $n = 9 \cdot m + 5$ for some m and that θ_m happens to be of the form $\psi \wedge \zeta$. Fix i.
  Add a node a to Γ_n, an immediate successor to α_i.
  If there is $b \sqsubseteq \alpha_i$ with $\varphi_b = \theta_m$, then define $\varphi_a = \zeta$.
  Otherwise, define $\varphi_a = c_0 \approx c_0$.
- Suppose $n = 6 \mod 9$, $n = 9 \cdot m + 6$ for some m and that θ_m happens to be of the form $\psi \wedge \zeta$. Fix i.
  Add a node a to Γ_n, an immediate successor to α_i.
  If there are $b, c \sqsubseteq \alpha_i$ with $\varphi_b = \psi$ and $\varphi_c = \zeta$, then define $\varphi_a = \psi \wedge \zeta$.
  Otherwise, define $\varphi_a = c_0 \approx c_0$.

Next we deal with the two cases for the $\approx$ condition.

- Suppose $n = 7 \mod 9$, $n = 9 \cdot m + 7$ for some m and that θ_m happens to be of the form $t \approx t$ for some constant term t. Fix i.
  Add a node a to Γ_n, an immediate successor to α_i, and define $\varphi_a = t \approx t$.
  Otherwise, define $\varphi_a = c_0 \approx c_0$.
- Suppose $n = 8 \mod 9$, $n = 9 \cdot m + 8$ for some m and that θ_m happens to be of the form $\psi[t]$ for some constant term t. Fix i.
  Add a node a to Γ_n, an immediate successor to α_i.
  If there are some $b, c \sqsubseteq \alpha_i$ with $\varphi_b = \psi[e]$ and $\varphi_c = t \approx e$, where e is a constant term, then define $\varphi_a = \psi[t]$.
  Otherwise, define $\varphi_a = c_0 \approx c_0$.

Finally, let us not forget the theory T!

- [Axiom case] Suppose $n = 0 \mod 9$, $n = 9 \cdot m$ for some m, and θ_m happens to be in T. Fix i.
  Add a node a to Γ_n, an immediate successor to α_i, and assign $\varphi_a = \theta_m$.
  Otherwise, define $\varphi_a = c_0 \approx c_0$.

For the root, we may define $\varphi_r = c_0 \approx c_0$. (Or θ_0, if it happens to be in T.)

Note that each Γ_n is a deduction tree from T. There are two options.

Case 1: the construction stops at some point. That is, there is some Γ_n for which nothing was added to Γ_{n+1}.
This can happen only if every branch of Γ_n contains a contradiction. That is, only if Γ_n is proof of contradiction from T!

We are currently assuming that this is not the case, so it must be that:

Case 2: the construction never stops. In this case let Γ be the union of all the trees Γ_n. Then Γ *must be infinite.*

Remark 7.33. The reason that one can make sense of this union tree Γ is because of this particular construction. Specifically, since each Γ_{n+1} just adds some notes "on top of" Γ_n.

The nodes of Γ are simply the nodes which appears in Γ_n for some n.
Given two nodes a, b in Γ, we ask if $a \sqsubset b$ by finding some large enough n so that a, b are nodes in Γ_n, and asking whether $a \sqsubset b$ in Γ_n.

This should be familiar from Pset 4 Question 2.

By Konig's lemma, there is an infinite branch $r = a_0 \sqsubset a_1 \sqsubset a_2 \sqsubset ...$ in Γ.
Let $T^+ = \{\varphi_{a_i} : i = 0, 1, 2, ...\}$.

Claim 7.34. T^+ satisfies all the Henkin conditions for the vocabulary $\mathcal{S}^+$. Moreover T^+ extends T.

Proof. First, we need to show that T^+ does not contain any sentence and its negation. This must be the case, for otherwise the branch would "stop growing" after finitely many steps!

Next, we want to show that for any θ, either $\theta \in T$ or $\neg\theta \in T$.
Fix m so that $\theta = \theta_m$, and let $n = 9 \cdot m + 1$.
By definition, a_{n-1} has 2 immediate successors in Γ_n, with assignments θ and $\neg\theta$.
a_n must be one of these two. So φ_{a_n} is either θ or $\neg\theta$.

To see that T^+ extends T: fix any $\theta \in T$. There is some m for which $\theta = \theta_m$. Let $n = 9 \cdot m$. The $\varphi_{a_m} = \theta_m = \theta$. So $\theta \in T^+$.

Let us look at the $\exists$ Henkin condition.
Suppose $(\exists x)\psi$ is in T^+, that is, it is φ_{a_k} for some k. We want to conclude that $\psi[c_l] \in T^+$ for some l.
We probably took care of it: in stage n, if $n = 9 \cdot m + 2$, where $\theta_m = (\exists x)\psi$.
Problem: this works only if $k < n$!

This problem is easy to fix. Instead of "fulfilling the corresponding Henkin condition" of each θ once, we will do it infinitely many times to each condition.
This will happen if $\theta = \theta_m$ *for infinitely many* m.

Lemma 7.35. Let X be a countable set. Then there exists an enumeration $x_0, x_1, x_2, ...$ of all the members of X, so that each $x \in X$ appears infinitely many times.

Proof. Since X is countable, so is $Y = X \times \mathbb{N}$.
Let $y_0, y_1, y_2, ...$ be an enumeration of Y.
For each n, if $y_n = (x, l)$ for some $l \in \mathbb{N}$, define $x_n = x$.
Pictorially, if $x'_0, x'_1, ...$ is any enumeration of X, we get the new one by:

$$\begin{array}{lllllll}
x_0 & & x_1 & & x_5 & & x_6 \\
x'_0 & \rightarrow & x'_1 & & x'_2 & \rightarrow & x'_3 \\
x_2 & \swarrow & x_4 & \nearrow & x_7 & \swarrow & \\
x'_0 & & x'_1 & & x'_2 & & \\
x_3 \downarrow & \nearrow & x_8 & \swarrow & & & \\
x'_0 & & x'_1 & & x'_2 & &
\end{array}$$

$\square$

Retroactively: let us take the enumeration $\langle \theta_n : n = 0, 1, 2, ...\rangle$ so that each $\mathcal{S}^+$-sentence θ appears infinitely many times.

Going back to the $\exists$ Henkin condition, we may now take m large enough (so that $n = 9 \cdot m + 2$ is bigger than k, and with $\theta_m = (\exists x)\psi$.

Let us now deal with the $\forall'$ Henkin condition.
Assume $\psi[t] \in T^+$ for some formula $\psi(x)$ and a constant term t. Fix k so that $\psi[t] = \varphi_{a_k}$.
Take m large enough, so that $n = 9 \cdot m + 3 > k$ and $\theta_m = \psi[t]$.
Then necessarily $\varphi_{a_n} = (\exists x)\psi$.

Exercise 7.36. Prove that the remaining Henkin conditions are satisfied for T^+.

$\square$

Finally, since T^+ is a Henkin theory, there is some model $\mathcal{A}^+$ for T^+.
The reduct $\mathcal{A}$ of $\mathcal{A}^+$ to the signature $\mathcal{S}$ is a model for T. (Recall the definitions from Pset 4. $\mathcal{A}^+$ is a structure for the signature $\mathcal{S}^+$. $\mathcal{A}$ is defined for the signature $\mathcal{S}$ by interpreting the symbols in $\mathcal{S}$ the same as $\mathcal{A}^+$.)
So T is satisfiable (has a model), concluding the proof of the completeness theorem.

Remark 7.37. If $\mathcal{S}$ is countable, and T is an $\mathcal{S}$-theory which is satisfiable, then the completeness theorem provides a countable model.
Specifically, we constructed the model using countably many constant symbols. We took a certain quotient, making that model either finite or countably infinite.

So if $\mathcal{S}$ is countable, T has a model if and only if it has a countable model. This we already know from the downwards Lowenheim-Skolem theorem.

8. Compactness

Theorem 8.1 (Compactness for $\models$)**.** Let $\mathcal{S}$ be a countable signature. Let T be a set of sentences using the signature $\mathcal{S}$. The following are equivalent.

- T is satisfiable (there is a model for T);
- for any *finite* subset $T_0 \subseteq T$, T_0 is satisfiable.

Remark 8.2. This is related to the notion of compactness from topology.

Proof. If $\mathcal{A}$ is a model for T, then $\mathcal{A}$ is a model for any subset of T.

The main point is the other direction. Suppose that every finite subset T_0 is satisfiable. We need to show that T has a model.

If T were not satisfiable, by the completeness theorem there is a proof of cotnradiction from T.
This proof is a finite deduction tree Γ. Let T_0 be all the sentences in T which are assigned to some node in Γ.
Then T_0 is finite and Γ is a proof of contradiction from T_0.
This leads to a contradiction (to the Soundness Theorem) as we assumed that T_0 is satisfiable. □

Example 8.3. Let $T = \{\psi_n : n = 1, 2, ...\}$, where ψ_n is the sentence saying "there are at least n different members": $(\exists x_1)...(\exists x_n)(\bigwedge_{i\neq j} \neg(x_i \approx x_j))$.

Each finite subset of T has a finite model. T also has a model, but it cannot be finite.

This shows that the model for T we get from the compactness theorem may have nothing to do with the models we get from the assumption that each T_0 is satisfiable.

Remark 8.4. Like the completeness theorem, the compactness theorem is true with no assumptions on $\mathcal{S}$ at all. We focus on countable languages here for simplicity.

Recall that a sentence θ cannot distinguish between countable vs uncountable models, by the downwards Lowenheim-Skolem theorem.

The theory discussed above has as its models precisely the infinite structures.

Two natural questions are:
Can a single sentence capture precisely the infinite structures?
Can a theory capture precisely the finite structures?
The answer for both is no.

Theorem 8.5. Suppose T has arbitrarily large finite models. That is, for any finite number k there is a model $\mathcal{A}$ for T with $|A| > k$. Then T has an infinite model.

Proof. Let $T^+ = T \cup \{\psi_1, \psi_2, ...\}$.

Claim 8.6. T^+ is fintiely satisfiable.

Proof. Fix a finite $T_0 \subseteq T^+$. Then there is some k so that $T_0 \subseteq T \cup \{\psi_1, ..., \psi_k\}$.

Let $\mathcal{A}$ be a model for T with $|A| \geq k$. Then $\mathcal{A} \models T$ and $\mathcal{A} \models \psi_i$ for $i = 1, ..., k$, so $\mathcal{A} \models T_0$. □

By the compactness theorem, T^+ has a model, which must be infinite. This model is also a model of T. □

Corollary 8.7. There is no theory T whose models are precisely the finite structures.

Corollary 8.8. There is no sentence θ whose models are precisely the infinite structures.

Proof. If the models of θ are precisely the infinite structures, then the models of $\neg\theta$ are precisely the finite structures. □

Recall that in Pset 3, Question 2(1), you wrote a sentence θ all of whose models are infinite. Without knowing what sentence you may write, question 2(2) asked you to prove that this particular sentence fails to characterize the infinite structures. Corollary 8.8 is precisely the reason.

8.1. **Ramsey's theorem.** Recall the (infinite) pigeon-hole principle: If X is infinite, $X = X_0 \cup ... \cup X_n$ is partitioned into finitely many pieces, then (at least) one piece X_i must be infinite.

Let us focus on the infinite set $\mathbb{N}$. Furthermore, let us restate the principle as follows, identifying a partition with a function:

$$\text{If } f\colon \mathbb{N} \to C \text{ and } C \text{ is finite, then } f^{-1}(c) \text{ is infinite for some } c \in C.$$

We will often call such a function a "coloring". That is, each $n \in \mathbb{N}$ is labelled with a color $f(n) \in C$. The conclusion is that, as there are only finitely many colors, there must be a single colored assigned to infinitely many $n \in \mathbb{N}$.

Ramsey's theorem can be seen as a "higher dimensional pigeon-hole principle".
Let us start with dimension 2.

Let $[\mathbb{N}]^2$ be the set of all *unordered* pairs $\{n, m\}$ with $n \neq m$, $n, m \in \mathbb{N}$.
A function $f\colon [\mathbb{N}]^2 \to C$ will be called a **coloring of pairs** (of natural numbers).
A set $S \subseteq \mathbb{N}$ is called **homogeneous** if there is a $c \in C$ so that for any distinct $a, b \in \mathbb{N}$, $f(\{a, b\}) = c$. That is, all pairs from S are assigned the same color. (Said another way: when restricting f to $[S]^2$, it is a constant function.)

Theorem 8.9 (Ramsey's theorem for pairs)**.** Let C be a finite set, $f\colon [\mathbb{N}]^2 \to C$. Then there is an *infinite* set $S \subseteq \mathbb{N}$ which is homogeneous for f.

Remark 8.10. Another way to view $[\mathbb{N}]^2$ is as $\{(n, m) : n < m\}$. (Upper triangle in the plane.) We can then think of our colorings as functions defined only on such pairs.

Alternatively, we can think of any such coloring as a *symmetric* function from $\mathbb{N} \times \mathbb{N} \to C$. (In this case we simply ignore the values of f on pairs of the form (n, n).)

A 2-coloring $f\colon [\mathbb{N}]^2 \to \{0, 1\}$ can be thought of as a graph (unordered, with no loops,) whose domain is $\mathbb{N}$. That is, given such f define a relation E on $\mathbb{N}$ by

$$n \mathrel{E} m \iff m \neq n \wedge f(n, m) = 1.$$

Similarly, given such graph $(\mathbb{N}, E)$ we may define $f(n, m) = 1 \iff n \mathrel{E} m$.

What is a homogeneous set $S \subseteq \mathbb{N}$ in this setting?

- S is homogeneous with fixed color 0 if *no two members of* S, $n, m \in S$, have an edge between them. In other words, looking at S as a graph, it looks like vertices with no edges. (An empty graph.)
- S is homogeneous with fixed color 1 if *any two* distinct members of S, $n, m \in S$, are connected by an edge. In other words, looking at S as a graph, it looks like infinitely many vertices which are all connected by edges. (A full graph.)

Rephrasing Ramsey's theorem in this setting:

An infinite graph either contains a copy of an infinite empty graph or contains a copy of an infinite full graph (possibly both).

This is part of "regularity phenomenon": we can always find "large trivial-looking substructures".

For a larger (finite) "color set" C, we can view a coloring $f\colon [\mathbb{N}]^2 \to C$ in graph terms as well: for any two vertices $n \neq m$ we assign an edge with some color $c \in C$. A homogeneous S set is a set so that

Remark 8.11. A similar "higher dimensional" result is also true, if we replaced $[\mathbb{N}]^2$ with $[\mathbb{N}]^m$ for some fixed m, the set of all subsets of $\mathbb{N}$ is size m.

Ramsey's original motivation was to find "infinite homogeneous subsets" of some structure $\mathcal{A}$.
Suppose $\mathcal{A}$ is a structure with domain $A = \mathbb{N}$, and $\varphi(x, y)$ is a formula. Consider the coloring $f(n, k) = 1$ if $\mathcal{A} \models \varphi(n, k)$ and $f(n, k) = 0$ if $\mathcal{A} \models \neg\varphi(n, k)$.
Using Ramsey's theorem we may find an infinite $S \subseteq \mathbb{N}$ so that the question $\mathcal{A} \models \varphi(n, k)$ has the same answer, for any $n < k$ both in S. (Either true for any $n < k$ from S or false for any $n < k$ from S.)

Similarly famous is the finite Ramsey theorem:

Theorem 8.12 (Finite Ramsey theorem)**.** Let C be a finite set. Then for any natural number h there is *some* ("large enough") natural number N, so that:
for any coloring $f\colon [N]^2 \to C$ there is some homogeneous set $S \subseteq N$ of size at least h.

That is: we want to find *large* homogeneous sets for colorings of finite graphs. We can do that, assuming the graph is large enough.

Even just formulating the finite Ramsey theorem is a little more convoluted that the infinite one. Also proving the infinite Ramsey theorem is easier, and more natural.

Back to our compactness theorem, let us see how we can deduce the finite Ramsey theorem from the infinite one (which we will proved soon).

Proof of the finite Ramsey theorem from the infinite one. For simplicity, let us work with 2 colors: $C = \{0, 1\}$.
Assume for contradiction that the finite Ramsey theorem fails in this case.
What does this mean?
There is some natural number h, so that for any natural number N, there is some coloring $f\colon [N]^2 \to \{0, 1\}$, for which there is *no* homogeneous set of size $\geq h$.
We would like to arrive at a counter example to the infinite Ramsey theorem.

Consider the language $S = \{R_0, R_1, c_0, c_1, ...\}$, where c_i are constant symbols and R_i are binary relations.
The idea is for c_k to be a "stand-in" for the number $k \in \mathbb{N}$, and the relation $R_i(c_k, c_m)$ to "represent" $f(k, m) = i$.
With this in mind, we may write axioms "saying that" this f is a counter-example to Ramsey's theorem, as follows.

(1) For each pair $k \neq m$, consider the axiom $\neg(c_k \approx c_m)$.
(2) $(\forall x)(\forall y)[(R_0(x, y) \vee R_1(x, y)) \wedge \neg(R_0(x, y) \wedge R_1(x, y))]$. Every pair is assigned one of the two colors.
(3) $(\forall x)(\forall y)(R_0(x, y) \leftrightarrow R_0(y, x))$ and $(\forall x)(\forall y)(R_0(x, y) \leftrightarrow R_0(y, x))$. We want this to correspond to a coloring of pairs as above.
(4) Given a natural number h, consider the sentence ψ_h saying that there is no homogeneous set of size h:
$$(\forall x_1)...(\forall x_h)[(\bigwedge_{1 \leq i < j \leq h} x_i \neq x_j) \to (\bigvee_{1 \leq i < j \leq h} R_0(x_i, x_j) \wedge \bigvee_{1 \leq i < j \leq h} R_1(x_i, x_j))]$$

Let T be the theory with all these sentences.

Our assumption (the failure of the finite Ramsey theorem) precisely says that any finite subset of T is satisfiable:
If $T_0 \subseteq T$ is finite, it only mentions finitely many symbols $c_0, ..., c_l$. Let h be the largest

so that ψ_h is in T_0.
By assumption, there is some N and $f\colon [N]^2 \to \{0,1\}$ with no homogeneous subset of size h.
We may assume that $N \geq l$.
(Let us view f as a symmetric function $f\colon N \times N \to \{0,1\}$.)

Define a structure $\mathcal{A}$ as follows.
$A = \{0, ..., N\}$.
$c_k = k$ for $k = 0, ..., N$. $c_k = 0$ for $k > N$.
$(k, m) \in R_0^{\mathcal{A}} \iff f(k, m) = 0$.
$(k, m) \in R_1^{\mathcal{A}} \iff f(k, m) = 0$.
$(m, m) \in R_0$ and $(m, m) \in R_1$ for all m. (We need to make some definition, but we don't really care about these values.).

Exercise 8.13. Check that $\mathcal{A} \models T_0$.

Finally, by the compactness theorem, there is some model $\mathcal{A} \models T$.
Define $f\colon [\mathbb{N}]^2 \to \{0,1\}$ as follows:
If $(c_n, c_m) \in R_0^{\mathcal{A}}$, $f(n, m) = 0$
If $(c_n, c_m) \in R_1^{\mathcal{A}}$, $f(n, m) = 1$.

The axioms in T tell us that f is well defined.
By the infinite Ramsey theorem, there is an infinite $S \subseteq \mathbb{N}$ which is homogeneous for f.
This leads to a contradiction: let $s_1, ..., s_h$ be distinct members of S.
Now $c_{s_1}^{\mathcal{A}}, ..., c_{s_h}^{\mathcal{A}}$ witness that ψ_h fails in $\mathcal{A}$. □

Proof of the infinite Ramsey theorem. Fix a finite set C and $f\colon [\mathbb{N}]^2 \to C$.
Let $a_0 = 0 \in \mathbb{N}$.
Find some $c_0 \in C$ so that $A_{c_0}^0 = \{n > 0 : f(0, n) = c_0\}$ is infinite.
There must be some such c_0, since the union $\bigcup_{c \in C} A_c$ is infinite, and C is finite. (The union of finitely many finite sets is finite.)
Let a_1 be the minimum of $A_{c_0}^0$.
Next, find some $c_1 \in C$ so that $A_{c_1}^1 = \left\{n \in A_{c_0}^0 : n > a_1 \text{ and } f(a_1, n) = c_1\right\}$.

Inductively: given c_k, an infinite $A_{c_k}^k$, let $a_{k+1} = \min A_{c_k}^k$.
Find $c_{k+1} \in C$ so that $\left\{n \in A_{c_{k+1}}^{k+1} : n > a_k \text{ and } f(a_k, n) = c_{k+1}\right\}$ is infinite.

We have a descending sequence $\mathbb{N} \supset A_{c_0}^0 \supset A_{c_1}^1 \supset ...$ with minimums $a_0 < a_1 < a_2 < ...$
Is the set $\{a_0, a_1, a_2, ...\}$ homogeneous? Not quite...
For $c \in C$, let $S_c = \{a_k : c_k = c\}$.
Since $\mathbb{N} = \bigcup_{c \in C} S_c$, we must have some c^* for which S_{c^*} is infinite.
Let $S = S_{c^*}$.

Claim 8.14. S is homogeneous for f, with color c.

Indeed: given $b < d \in S$, $b = a_k$ and $d = a_m$ where $c_k = c_m = c$.
In particular $a_m \in A_{c_k}^k$ and therefore $f(a_k, a_m) = c_k = c$. □

9. Models of a complete theory and types

As before, we focus on countable objects: $\mathcal{S}$ is a countable vocabulary, T is a theory in the language for $\mathcal{S}$, and we study countable models of T ($\mathcal{S}$-structures $\mathcal{A}$ with A countable and $\mathcal{A} \models T$).

Remark 9.1. Even when $\mathcal{S}$ is countable and T is a natural theory, such as algebraically closed fields, one can learn *a lot* by looking at the *uncountable* models. This requires some familiarity with set theoretic techniques.

We focus on the countable models as there are already a lot of interesting things we can say about those.

If $\mathrm{Con}(T)$ is not a complete theory, then there is some sentence θ and models $\mathcal{A}, \mathcal{B}$ of T with $\mathcal{A} \models \theta$ and $\mathcal{B} \models \neg\theta$. In this case, the difference between $\mathcal{A}$ and $\mathcal{B}$ is clear, and they are not isomorphic.

For example, suppose T is the theory of groups. $\mathcal{A}$ can be a group of size 5, and $\mathcal{B}$ can be a group of size 10.
If T is the theory of linear orders, $\mathcal{A}$ can be $\mathbb{Z}$ and $\mathcal{B}$ can be $\mathbb{Q}$, in which case we know that the "Density" axiom separates them.

From now on, we focus on studying the models of a complete theory T.

We saw many complete theories with "nice axiomatizations".
For example, Con(DLO), the logical consequences of the DLO axioms, is a complete theory, which is equal to $\mathrm{Th}(\mathbb{Q}, <)$, which is equal to $\mathrm{Th}(\mathbb{R}, <)$.
Similarly, the logical consequences of "DLO + $\{c_{n+1} < c_n : n = 1, 2, ...\}$", in the vocabulary $\{<, c_1, c_2, ...\}$, is a complete theory, equal to the theory of the structure $(\mathbb{Q}, <, 1, \frac{1}{2}, \frac{1}{3}, ...)$.

Similarly (though we have not proven it) the logical consequences of the theory ACF_0 - algebraically closed fields of characteristic 0 - is a complete theory which is equal to the theory of $(\mathbb{C}, \cdot, +, 0, 1)$.

Note that if T has a model, then $T = \mathrm{Th}(\mathcal{A})$ for any model $A \models T$.
Therefore, we will interchangeably work with either a complete theory, or a particular model $\mathcal{A}$, with the understanding that the theory we are studying is $\mathrm{Th}(\mathcal{A})$.
We may work with the structure $(\mathbb{Q}, <, 1, \frac{1}{2}, ...)$, meaning we are interested in its theory, and any other structure of its theory, including $(\mathbb{Q}^+, <, 1, \frac{1}{2}, ...)$.

Question: Let T be a complete theory, $\mathcal{A}, \mathcal{B}$ models of T. Must they be isomorphic?

Note that, "$\mathcal{A}$ and $\mathcal{B}$ are models of the same complete theory" is the same as $\mathcal{A} \equiv \mathcal{B}$. So an equivalent question is: given a structure $\mathcal{A}$, if $\mathcal{B} \equiv \mathcal{A}$, must they be isomorphic?)

Early on, we suspected that they should be isomorphic. We saw however that $(\mathbb{Q}, <) \equiv (\mathbb{R}, <)$, yet there is no bijective map between them, based on cardinality issues.

But now we focus on countable structures only. Still, we saw that there could be countable structures $\mathcal{A} \equiv \mathcal{B}$ yet $\mathcal{A} \not\simeq \mathcal{B}$.

We will now introduce tools "beyond sentences" to study structures and be able to distinguish between non-isomorphic ones.

Before that, let us mention the following incredible result.

Definition 9.2. Let $\mathcal{S}$ be a countable signature, T a complete theory.

Let $I(T)$ be the number of non-isomorphic countable models of T.

That is, $I(T) \geq k$ if we can find $\mathcal{A}_1, ..., \mathcal{A}_k \models T$ which are pairwise non-isomorphic. Note that $I(T)$ can be infinite as well.

Remark 9.3. Another way to view $I(T)$: let $M(T)$ be all countable models of T. The isomorphism relation $\simeq$ on $M(T)$ is an equivalence relation, and therefore partitions $M(T)$

into equivalence classes. $I(T)$ is precisely the number of equivalence classes (which can be infinite).

Example 9.4. (1) If T is unsatisfiable (inconsistent) then $I(T) = 0$.
(2) If $T = \text{Con(DLO)}$ ($S = \{<\}$), or $T = \text{Con(Random Graph)}$ ($\mathcal{S} = \{E\}$), then $I(T) = 1$.
(3) Let $\mathcal{S} = \{<, c_1, c_2, ...\}$, T be the theory of $(\mathbb{Q}, <, 1, \frac{1}{2}, ...)$. Then $I(T) = 3$.
(4) If $T = \text{Th}(\mathbb{C}, +, \cdot, 0, 1)$, $I(T)$ is infinite. Specifically, there are countable algebraically closed fields of "transendence degree n" for each $n = 0, 1, 2, 3, ...$, and they are therefore non-isomorphic to one another.

Theorem 9.5 (Vaught's theorem)**.** Let $\mathcal{S}$ be a countable signature, T a complete theory. Then

$$I(T) \neq 2.$$

Given the generality here (T is just any complete theory, meaning the theory of some structure, in an arbitrary countable signature), this is quite surprising!

Remark 9.6. For each $n = 3, 4, 5, 6, ...$, there is a complete theory T with $I(T) = n$ precisely. Such examples can be constructed in a way very similar to the example with $n = 3$ done in Pset 5.

9.1. **Types.** Given finitely many sentences $\theta_1, ..., \theta_k$, the theory $\{\theta_1, ..., \theta_k\}$ is "the same as" the single sentence theory $\{\theta_1 \wedge ... \wedge \theta_k\}$. However, as we have seen, an infinite theory can express more than any single sentence (and therefore more than any finitely many sentences) can express.

Roughly speaking, a **type** is to a formula what a theory is to a sentence.

Given a signature $\mathcal{S}$, and structure $\mathcal{A}$, and a formula $\varphi(x)$, recall that we may view the realization of $\varphi(x)$ in $\mathcal{A}$ as a subset of A, $\varphi^{\mathcal{A}}(x) \subseteq A$, which is $\{a \in A : \mathcal{A} \models \varphi(a)\}$.
$\varphi^{\mathcal{A}}(x)$ is not empty if and only if $\mathcal{A} \models (\exists x)\varphi$ (the latter is a sentence).

Suppose we have two formulas $\varphi(x), \psi(x)$, which interpret as some subsets of a model $\mathcal{A}$.
Can we find some member of $\mathcal{A}$ satisfying both?
For example, we may consider $\mathcal{A} = (\mathbb{R}, +, \cdot, 0, 1)$, $\varphi(x) = (\exists y)(y \cdot y \approx x)$ and $\psi(x) = x \cdot x \approx 1 + 1$.
There is some $a \in A$ satisfying both φ and ψ if and only if $\varphi^{\mathcal{A}}(x) \cap \psi^{\mathcal{A}}(x) \neq \emptyset$ if and only if $\mathcal{A} \models (\exists x)(\varphi \wedge \psi)$.

Recall now the structures $\mathcal{A}, \mathcal{B}, \mathcal{C}$ from Pset 5. $\mathcal{S} = \{<, c_1, c_2, ...\}$.

- $\mathcal{A} = (\mathbb{Q}, <, 1, \frac{1}{2}, ...)$.
- $\mathcal{B} = (\mathbb{Q} \setminus \{0\}, <, 1, \frac{1}{2}, ...)$.
- $\mathcal{C} = (\mathbb{Q}^+, <, 1, \frac{1}{2}, ...)$.

The key distinction between $\mathcal{C}$ and $\mathcal{A}, \mathcal{B}$ is that in $\mathcal{C}$ there is no member which is below $\frac{1}{n}$ for each n.
This cannot be expressed using any finitely many sentences.
Let $\psi_n(x)$ be the formula $x < c_n$, let $p = \{\psi_n(x) : n = 1, 2, ...\}$ be this collection of formulas. Then the question we are asking is: is there some $a \in A$ so that $\mathcal{A} \models \psi_n(a)$ for each $n = 1, 2, ...$ We will call this p a 1-type (working in the structure $\mathcal{A}$).

Definition 9.7. Fix a signature $\mathcal{S}$ and a structure $\mathcal{A}$. Let $\bar{x} = x_1, ..., x_n$ be (distinct) variables. Let p be a set of $\mathcal{S}$-formulas with free variables included in $x_1, ..., x_n$. (So each $\varphi \in p$ is thought of as $\varphi(x_1, ..., x_n)$.

Say that p is an **n-type** (for $\mathcal{A}$) if for any finite subset of p, $\theta_1(\bar{x}), ..., \theta_k(\bar{x})$ from p, there is some $\bar{a} = a_1, ..., a_n \in A$ so that $\mathcal{A} \models \theta_i(\bar{a})$ for $i = 1, ..., k$. (It is "finitely satisfiable".)

Say that p is a **complete n-type** if it is an n-type and moreover for any formula $\varphi(\bar{x})$, either $\varphi(\bar{x}) \in p$ or $\neg\varphi(\bar{x}) \in p$.

Given an n-type p and $\bar{a} = a_1, ..., a_n$ from A, say that $\bar{a}$ **realizes** p if $\mathcal{A} \models \theta(\bar{a})$ for *every* $\theta(\bar{x}) \in p$.

Say that p is **realized** (in $\mathcal{A}$) for there is some $\bar{a}$ in $\mathcal{A}$ realizing p.

Example 9.8.
1. Given a formula $\varphi(x)$, if $\mathcal{A} \models (\exists x)\varphi$ then $\{\varphi(x)\}$ is a type, which is realized in $\mathcal{A}$.
2. Let $\mathcal{C} = (\mathbb{Q}^+, <, 1, \frac{1}{2}, ...)$. Let $\psi_n(x) = x < c_n$. Then $p = \{\psi_n : n = 1, 2, ...\}$ is a 1-type. It is *not* realized in $\mathcal{C}$. p is also a 1-type in $\mathcal{A} = (\mathbb{Q}, <, 1, \frac{1}{2}, ...)$. Any $a \leq 0$ in $\mathbb{Q}$ realizes p in $\mathcal{A}$.
3. Let $\mathcal{A} = (\mathbb{R}, <, +, \cdot, 0, 1)$. Let $\psi_0(x) = 0 < x$. For $n = 1, 2, ...,$ let $\psi_n(x) = x \cdot (1 + ... + 1) < 1$ where we add n many 1's. That is, $\psi_n(x)$ says that $x < \frac{1}{n}$. Let $p = \{\psi_n : n = 0, 1, 2, ...\}$ is a type. p is not realized in $\mathcal{A}$.
4. Consider the signature for vector spaces of the field $\mathbb{Q}$. $\mathcal{S} = \{+, -, \bar{0}\} \cup \{f_q : q \in \mathbb{Q}\}$. For $q \in \mathbb{Q}$, let $\psi_q(x, y) = \neg(f_q(x) = y)$. Then $p = \{\psi_q(x, y) : q \in \mathbb{Q}\}$ says that the vectors x, y are linearly independent over $\mathbb{Q}$.
   In the vector space $\mathbb{Q}$, p is a 2-type which is not realized.
   In the vector space $\mathbb{Q}^2$, p is realized by a_1, a_2 where $a_1 = (1, 0)$ and $a_2 = (0, 1)$.
5. Let $\mathcal{A} = (\mathbb{C}, +, \cdot, 0, 1)$. We can express "$x, y$ are algebraically independent" by a 2-type $p(x, y)$.
   This would say that $P(x, y) \neq 0$ for any polynomial P with rational coefficients.
   For any such fixed polynomial, this can be done by a single formula.
6. Let $\mathcal{A} = (\mathbb{N}, <, \cdot, +, 0, 1)$. Let $p = \{\psi_n : n = 1, 2, ...\}$ where $\psi_n = (1 + ... + 1) < x$, where we add n many 1's. p is a 1-type, which is not realized in $\mathcal{A}$. A realization of p is a "non-standard natural number".

Remark 9.9. In the setting of Definition 9.7, fix $p = p(\bar{x})$. The following are equivalent:

- p is an n-type;
- for any $\theta_1(\bar{x}), ..., \theta_k(\bar{x})$ from p, $(\exists x_1)...(\exists x_n)(\theta_1 \wedge ... \wedge \theta_k) \in \mathrm{Th}(\mathcal{A})$.

In particular, the main condition for "being a type" only depends on the theory.

Corollary 9.10. If p is a type (for $\mathcal{A}$) and $\mathcal{A} \equiv \mathcal{B}$, then p is a type for $\mathcal{B}$.

We will think of a structure as large if it satisfies many types, and as small if it fails to.

Back to our examples $\mathcal{A}, \mathcal{B}, \mathcal{C}$: we see that $\mathcal{C}$ is small, compared to $\mathcal{A}, \mathcal{B}$, as it fails to realize the type $\{x < c_n : n = 1, 2, ...\}$.

What about $\mathcal{A}$ and $\mathcal{B}$. As we have seen, what distinguishes them is the 0, the "limit" of the constants c_n. While $\mathcal{A}$ has this maximal element below all the c_n's, $\mathcal{B}$ does not. *This in fact corresponds to $\mathcal{A}$ being smaller than $\mathcal{B}$!*

The idea is that $\mathcal{A}$ fails to realize the "infinitesimal" type saying $0 < x$ and $x < \frac{1}{n}$ for $n = 1, 2, ...$

More precisely, we will have to talk about **types with parameters**. This is the natural analogy of a type where all formulas are allowed to use (the same) parameters.

Definition 9.11. Fix a signature $\mathcal{S}$ and a structure $\mathcal{A}$. Let $\bar{x} = x_1, ..., x_n$ be (distinct) variables. Let $\bar{d} = d_1, ..., d_k$ be some members of A. Let p be a set of formulas of the form $\varphi(x_1, ..., x_n, y_1, ..., y_k)$.

Say that p is an **n-type** with parameter $\bar{d}$ if for any finite subset of p, $\theta_1(\bar{x}, \bar{y}), ..., \theta_l(\bar{x}, \bar{y})$ from p, there is some $\bar{a}$ from A so that $\mathcal{A} \models \theta_i(\bar{a}, \bar{d})$ for $i = 1, ..., l$. (Finitely satisfiable.)

In this case, a realization of p in $\mathcal{A}$ is $\bar{a}$ from $\mathcal{A}$ so that $\mathcal{A} \models \theta(\bar{a}, \bar{d})$ for any $\theta(\bar{x}, \bar{y})$ from p.

An equivalent way to think about types with parameters, which we will often adopt is as follows.

Remark 9.12. Let $\mathcal{S}^+$ be $\mathcal{S}$ adjoined by new constant symbols $e_1, ..., e_k$. Let $\mathcal{A}^+$ be the expansion of $\mathcal{A}$ to $\mathcal{S}^+$ by $e_i^{\mathcal{A}} = d_i$. Given a type with parameters p as above, let q be the type of all formulas $\varphi(\bar{x})$ which are of the form $\theta[\bar{e}]$ with $\theta(\bar{x}, \bar{y}) \in p$.

That is, for any $\theta(\bar{x}, \bar{y})$ in p, substitute each y_i with the constant symbol e_i, to get a formula $\varphi(\bar{x})$.

Then q is a type (with no parameters, for $\mathcal{A}^+$) if and only if p is a type (with the parameters $d_1, ..., d_k$, for $\mathcal{A}$).
Moreover, q is realized in $\mathcal{A}^+$ if and only if p is realized in $\mathcal{A}$.

Back to $\mathcal{A} = (\mathbb{Q}, <, 1, \frac{1}{2}, ...)$. Let c be a new constant symbol and expand $\mathcal{A}$ to $\mathcal{A}^+$ by $c^{\mathcal{A}^+} = 0$. Define $p = p(x)$ by $p = \{c < x\} \cup \{x < c_n : n = 1, 2, ...\}$.
Then p is a 1-type (check).
Furthermore, p is *not* realized in $\mathcal{A}$ (that is, in $\mathcal{A}^+$).

Exercise 9.13. Let $\mathcal{B} = (\mathbb{Q} \setminus \{0\}, <, 1, \frac{1}{2}, ...)$. Then for *any* expansion $\mathcal{B}^+$ for a new constant c, if p is a type for $\mathcal{B}^+$, then p *is* realized in $\mathcal{B}^+$.

Specifically, we see that $\mathcal{A}$ is "missing something" and is therefore "smaller".

We will make the notions of "big" and "small" precise soon, and see that $\mathcal{B}$ is as big is it gets, and $\mathcal{C}$ is as small as it gets.

While types may be not realized, we will see now that we can also (elementarily) extend the structure to realize them!

Recall that we defined types as something which is "locally (finitely) true". This result shows that a type can be thought of as something which is "somewhere (in some bigger universe) true".

Lemma 9.14. Fix $\mathcal{A}$ and $p = p(x_1, ..., x_n)$ an n-type. Then there is a structure $\mathcal{B}$ so that
- $A \equiv \mathcal{B}$;
- p is realized in $\mathcal{B}$.

Proof. Let $c_1, ..., c_n$ be new constant symbols. Consider the theory T, in an expanded language $\mathcal{S}^+ = \mathcal{S} \cup \{c_1, ..., c_n\}$, containing $\mathrm{Th}(\mathcal{A})$ as well as the sentences $\theta[c_1, ..., c_n]$ for each $\theta(\bar{x}) \in p$.

Claim 9.15. T is finitely satisfiable.

Proof. Let $T_0 \subseteq T$ be finite. Then there are finitely many $\theta_1(\bar{x}), ..., \theta_k(\bar{x})$ from p so that $T_0 \subseteq T \cup \{\theta_1[\bar{c}], ..., \theta_k[\bar{c}]\}$.

By assumption, there is $\bar{a}$ in A so that $\mathcal{A} \models \theta_i(\bar{a})$ for $i = 1, ..., k$.

Expand $\mathcal{A}$ to $\mathcal{A}^+$ for $\mathcal{S}^+$ by $c_i^{\mathcal{A}+} = a_i$. Then (as we have seen before) $\mathcal{A}^+ \models \theta_i[\bar{c}]$ for each i.

It follows that $A^+ \models T_0$, as required. □

By the compactness theorem, T has a model $\mathcal{B}^+$. Let $\mathcal{B}$ be its reduct to $\mathcal{S}$. This $\mathcal{B}$ satisfies $\mathrm{Th}(\mathcal{A})$ (so $\mathcal{B} \equiv \mathcal{A}$).
Moreover, if $b_i = c_i^{\mathcal{B}^+} \in B$, then $\bar{b} = b_1, ..., b_n$ realizes p in $\mathcal{B}$.

□

Definition 9.16. Given a structure $\mathcal{A}$, $\bar{a} = a_1, ..., a_n$ from A, define **the type of** $\bar{a}$, denoted $\mathrm{tp}(\bar{a})$ to be

$$\mathrm{tp}(\bar{a}) = \{\varphi(\bar{x}) : \mathcal{A} \models \varphi(\bar{a})\},$$

where $\varphi(\bar{x})$ ranges over all formulas with free variables included in $x_1, ..., x_n$.

Remark 9.17. $\mathrm{tp}(\bar{a})$ is always a complete type.

Corollary 9.18. Fix a complete theory T. For any type $p = p(\bar{x})$ (not assumed to be complete) there is a type $q = q(\bar{x})$, so that $p \subseteq q$ and q is a complete type.

Proof. Find some model $\mathcal{B}$ in which p is realized. That is, there are $\bar{a} = a_1, ..., a_n$ in B realizing p.
Let $q = \mathrm{tp}^{\mathcal{B}}(\bar{a})$. Then $p \subseteq q$ and q is a complete type. □

Next, we note that the above result can be strengthened. Specifically, we can get $\mathcal{B}$ which is not just a model of $\mathrm{Th}(\mathcal{A})$, but in fact an elementary extension of $\mathcal{A}$.

First, let us consider a slight generalization of "elementary substructure".
Recall that a substructure $\mathcal{A} \subseteq \mathcal{B}$ can be thought of as an embedding with the identity function $f(a) = a$.
Conversely, given an embedding $f : \mathcal{A} \to \mathcal{B}$, then we can think of $\mathcal{A}$ as embedded in $\mathcal{B}$.

Specifically, let $A' = \{f(a) : a \in A\}$. Then we may view A' as a substructure of $\mathcal{B}$, so that f is an isomorphism between $\mathcal{A}$ and $\mathcal{A}'$.
Since we completely identify isomorphic structure, we may identify $\mathcal{A}$ as $\mathcal{A}'$ and think of it as a substructure of $\mathcal{B}$.

Definition 9.19. Let $\mathcal{A}, \mathcal{B}$ be structures in the same vocabulary $\mathcal{S}$. a map $f : A \to B$ is an **elementary embedding** if for any formula $\varphi(x_1, ..., x_n)$, for any $a_1, ..., a_n$ from A,

$$\mathcal{A} \models \varphi(a_1, ..., a_n) \iff \mathcal{B} \models \varphi(f(a_1), ..., f(a_n)).$$

If f is the identity map, we recover the definition $\mathcal{A} \preceq \mathcal{B}$.
Furthermore, as discussed above, whenever we have an elementary embedding f from $\mathcal{A}$ to $\mathcal{B}$ we can view $\mathcal{B}$ as an elementary extension of $\mathcal{A}$ (after some renaming).

For example: given $\mathcal{A}_1, \mathcal{A}_2, ...$ and elementary embeddings $f_i : \mathcal{A}_i \to \mathcal{A}_{i+1}$, one can define a "union model" $\mathcal{A}$ and elementary embeddings from $\mathcal{A}_i$ to $\mathcal{A}$, in a natural way, similar to Pset 4.
Instead, we can (by renaming a little) view this as a chain $\mathcal{A}_1 \preceq \mathcal{A}_2 \preceq \mathcal{A}_3 \preceq ...$, and then apply the question from PSet 4 directly to get the union model $\mathcal{A}$.

Fix a structure $\mathcal{A}$ for $\mathcal{S}$. Expand $\mathcal{S}$ to $\mathcal{S}_A$ by adding a constant symbol c_a for each $a \in A$.
The **elementary diagram** of $\mathcal{A}$, $D^e_{\mathcal{A}}$, is the following $\mathcal{S}_A$-theory, which is supposed to code *all* truths in $\mathcal{A}$.
Fix *any* formula $\varphi(\bar{x})$. Fix $\bar{a} = a_1, ..., a_n \in A$. Let $\bar{c} = c_{a_1}, ..., c_{a_n}$, and $\varphi[\bar{c}]$ the result of substituting every x_i by c_i.[5]
If $\varphi^{\mathcal{A}}(\bar{a}) = 1$, then put the sentence $\varphi[\bar{c}]$ in $D^e_{\mathcal{A}}$.
If $\varphi^{\mathcal{A}}(\bar{a}) = 0$, then put the sentence $\neg\varphi[\bar{c}]$ in $D^e_{\mathcal{A}}$.

Exercise 9.20. Let $\mathcal{B}^+$ be an $\mathcal{S}_{\mathcal{A}}$-structure. Assume that $\mathcal{B}^+ \models D^e_{\mathcal{A}}$. Let $\mathcal{B}$ be the reduct to $\mathcal{S}$. Define a function $f\colon A \to B$ by $f(a) = c_a^{\mathcal{B}^+}$. Prove that f is an **elementary embedding**: that is, for any formula $\varphi(x_1, ..., x_n)$, for any $a_1, ..., a_n$ from A,

$$\mathcal{A} \models \varphi(a_1, ..., a_n) \iff \mathcal{B} \models \varphi(f(a_1), ..., f(a_n)).$$

As above, we may think of this as follows. Define $A' = \left\{c_a^{\mathcal{B}} : a \in A\right\}$, and define the structure $\mathcal{A}'$ with universe A' in the natural way. Then the map f is an isomorphism from $\mathcal{A}$ to $\mathcal{A}'$.
Furthermore, if $\mathcal{B} \models D^e_{\mathcal{A}}$ then $\mathcal{A}'$ is an elementary substructure of $\mathcal{B}$.
So, by identifying $\mathcal{A}$ with $\mathcal{A}'$, we may view $\mathcal{B}$ as an elementary *extension* of $\mathcal{A}$.

Lemma 9.21. Fix $\mathcal{A}$ and $p = p(x_1, ..., x_n)$ an n-type. Then there is a structure $\mathcal{B}$ so that

- $A \preceq \mathcal{B}$;
- p is realized in $\mathcal{B}$.

Proof. Let $\mathcal{S}^+$ be S together with the new symbols $c_1, ..., c_n$ *and* the new "symbols for $\mathcal{A}$" $\{c_a : a \in \mathcal{A}\}$.
Let T be the collection of all sentences $\theta[c_1, ..., c_n]$, for $\theta(\bar{x}) \in p$, together with $D^e_{\mathcal{A}}$. (Recall $D^e_{\mathcal{A}}$ from Pset 6.)
The same argument as in Lemma 9.14 above shows that T is finitely satisfiable. (In $\mathcal{A}^+$ above, we realize c_a as a.)
So we get a model $\mathcal{B}^+$ realizing p and satisfying $D^e_{\mathcal{A}}$. Let $\mathcal{B}$ be its reduct to $\mathcal{S}$.
Now there is an elementary embedding of $\mathcal{A}$ into $\mathcal{B}$. Specifically, the map sending a to $c_a^{\mathcal{B}^+}$ is such.
As discussed above, by renaming, we may view this $\mathcal{B}$ as an elementary extension of $\mathcal{A}$, $\mathcal{A} \preceq \mathcal{B}$. ☐

Corollary 9.22. Fix $\mathcal{A}$ and $p = p(x_1, ..., x_n)$ an n-type with parameter $\bar{d} = d_1, ..., d_k$ in A. Then there is a structure $\mathcal{B}$ so that $A \preceq \mathcal{B}$ and p is realized in $\mathcal{B}$.

Proof. Let $e_1, ..., e_k$ be new constant symbols. Expand $\mathcal{A}$ to $\mathcal{A}^+$ by $e_i^{\mathcal{A}^+} = d_i$. Now apply the previous lemma to $\mathcal{A}^+$. ☐

Let us note a simple generalization of what we have done so far.

Lemma 9.23. Fix $\mathcal{A}$ and types $p_0, p_1, ...$ (with parameters). Then there is a structure $\mathcal{B}$ so that

[5] Recall there are some subtleties with substitution when quantifiers are involved. However, we may simply assume that the variables $\bar{x}$ do not have any "quantified appearences", and then substitution is very natural.

- $\mathcal{A} \preceq \mathcal{B}$;
- for each $i = 0, 1, 2, ...$, p_i is realized in $\mathcal{B}$.

Proof. The proof is essentially the same. Here we add infinitely many constant symbols. For each i, if $p = p(x_1, ..., x_{n_i})$, we add $c^i_1, ..., c^i_{n_i}$ and add to the theory $\theta[c^i_1, ..., c^i_{n_i}]$ for each $\theta \in p_i$.
If $p = p(x_1, ..., x_{n_i})$ is a type with parameter $\bar{d} = d_1, ..., d_k$, we add $\theta[c^i_1, ..., c^i_{n_i}, c_{d_1}, ..., c_{d_k}]$ to T^+, for each $\theta(\bar{x}, \bar{y}) \in p_i$.

Again we see that this theory is finitely realizable, by the virtue of p_i being all finitely realizable, and so there is a model. In this final model, the interpretation of $c^i_1, ..., c^i_{n_i}$ is a realization of p_i. $\square$

A reformulation of our "isomorphism theorem" is the following:

Lemma 9.24. Suppose $f \colon A \to B$ is an isomorphism from $\mathcal{A}$ to $\mathcal{B}$. Fix $\bar{a} = a_1, ..., a_n$ in A, and let $\bar{b} = f(a_1), ..., f(a_n)$ in B. Then

$$\mathrm{tp}(\bar{a}) = \mathrm{tp}(\bar{b}).$$

Corollary 9.25. Suppose $\mathcal{A}$ and $\mathcal{B}$ are isomorphic. Then a type p is realized in $\mathcal{A}$ if and only if it is realized in $\mathcal{B}$.

So, the realization of types, and failure therefore, may help us to distinguish non-isomorphic structures!

Definition 9.26. Given $\mathcal{A}$, $\bar{a}$, $\bar{d}$, define

$$\mathrm{tp}(\bar{a}/\bar{d}) = \left\{\varphi(\bar{x}, \bar{y}) : \mathcal{A} \models \varphi(\bar{a}, \bar{d})\right\}.$$

(The type of $\bar{a}$ "over" $\bar{d}$.)

Equivalently, we may think of it $\mathrm{tp}(\bar{a})$ in the structure $\mathcal{A}^+$ in the language $\mathcal{S}^+$ with $c_i^{\mathcal{A}^+} = d_i$.

Again, our "isomorphism theorem" can be cast as follows:

Lemma 9.27. Suppose $f \colon A \to B$ is an isomorphism from $\mathcal{A}$ to $\mathcal{B}$. Fix $\bar{a} = a_1, ..., a_n$, $\bar{d} = d_1, ..., d_k$ in A, and let $\bar{b} = f(a_1), ..., f(a_n)$ in B and $\bar{e} = f(d_1), ..., f(d_k)$. Then

$$\mathrm{tp}(\bar{a}/\bar{d}) = \mathrm{tp}(\bar{b}/\bar{e}).$$

Equivalently, if we expand by $c_i^{\mathcal{A}^+} = d_i$ and $c_i^{\mathcal{B}^+} = f(d_i)$, then $\mathrm{tp}(\bar{a})$ (in $\mathcal{A}^+$) is equal to $\mathrm{tp}(\bar{b})$ (in $\mathcal{B}^+$).

9.2. **Large structures.**

Definition 9.28. Let $\mathcal{S}$ be a countable signature and $\mathcal{A}$ a countable structure. Say that $\mathcal{A}$ is **saturated** if any type (with parameters) is realized in $\mathcal{A}$. (Equivalently, in any expansion of $\mathcal{A}$ by finitely many constants, any type is realized.)

Theorem 9.29. Suppose $\mathcal{A}$,$\mathcal{B}$ are countable structures for a signature $\mathcal{S}$, and both are saturated. Then $\mathcal{A} \simeq \mathcal{B}$.

Proof. We have seen this in various forms, several times, since week 1.

Let us sketch a winning strategy in the game $\mathcal{G}(\mathcal{A}, \mathcal{B})$:

Suppose we have $\bar{a} = a_1, ..., a_n$ and $\bar{b} = b_1, ..., b_n$, the plays in the game so far.
So the map $a_i \mapsto b_i$ is a partial isomorphism.

Given any $a \in A$ (played by player I), how would we respond?
Let $p = \text{tp}(a/\bar{a})$.
Then $q = \left\{\varphi(x, \bar{b}) : \varphi(x, \bar{a}) \in p\right\}$ is a type in $\mathcal{B}$.
Since $\mathcal{B}$ is saturated, there is some $b \in B$ realizing q.
Now the fact that $\text{tp}(a/\bar{a}) = \text{tp}(b/\bar{b})$ means that b is a legit move (the map $a_i \mapsto b_i$ and $a \mapsto b$ is a partial isomorphism).

The other case, where player I chooses $b \in B$, is similar, using that $\mathcal{A}$ is saturated. □

A similar proof, doing only the "forth" with no "back", shows:

Theorem 9.30. Let T be a complete theory. Suppose $\mathcal{B} \models T$ is a saturated countable model. Suppose $\mathcal{A} \models T$. Then there is an elementary embedding from $\mathcal{A}$ to $\mathcal{B}$.

So a countable saturated model, if exists, is unique (up to isomorphism), and is the "largest model" in the sense that all other models appear as elementary substructures of it.

A countable saturated model does not always exists however.

Definition 9.31. Let T be a complete theory. For $n = 1, 2, ...$ define

$$S_n(T) = \{p : p \text{ is a complete n-type for } T\}. \;\; S(T) = \bigcup_n S_n.$$

(*Complete* is important here.)

Exercise 9.32. Let $\mathcal{A} = (\mathbb{Q}, <)$, $T = \text{Th}(\mathcal{A})$. Prove that there is exactly one complete 1-type.
How many complete n-types are there?

Exercise 9.33. Let $\mathcal{A} = (\mathbb{Q}, <, 1, \frac{1}{2}, ...)$, $T = \text{Th}(\mathcal{A})$. What is $S_1(T)$?

Remark 9.34. If $p(\bar{x})$ is an n-type with parameters $\bar{d} = d_1, ..., d_k$, then p is also an $(n+k)$-type (without parameters).

By "forgetting the parameters" this way, the question of *realization* may change.
Nevertheless, it shows that if we understand $S_n(T)$ for all n, then we also understand all types with parameters.

Generally speaking, each type $p \in S_n(T)$ is a set of formulas, $p \subseteq \mathcal{F}$, where $\mathcal{F}$ is the set of all formulas in the signature $\mathcal{S}$.
Let $\mathcal{P}(\mathcal{F})$ be the *posetset* of $\mathcal{F}$: $\mathcal{P}(\mathcal{F}) = \{X : X \subseteq \mathcal{F}\}$ (the set of all subsets of $\mathcal{F}$).
Recall that in our case $\mathcal{F}$ is countable.
Recall also that for a countable set $\mathcal{F}$, the powerset $\mathcal{P}(\mathcal{F})$ is *not* countable.

So generally, $S_n(T) \subseteq \mathcal{P}(\mathcal{F})$, is contained in some uncountable set.

Depending on T, $S_n(T)$ could in fact be very small (finite), could be infinite yet countable, and could be uncountable!
In either of these three cases, we learn a lot about the models of T.

Theorem 9.35. Let T be a complete theory. The following are equivalent.

(1) There is a countable model $\mathcal{A} \models T$ which is saturated.
(2) For every $n = 1, 2, ...$, $S_n(T)$ is countable.

Proof. (1) $\implies$ (2).
Assume that $\mathcal{A} \models T$ and $\mathcal{A}$ is a saturated model.
In particular, any type $p \in S_n(T)$ *is realized in* $\mathcal{A}$.

Recall that if p is a *complete* type and $\bar{a}$ in $\mathcal{A}$ is a realization of p, then necessary $\mathrm{th}(\bar{a}) = p$.
(This is analogous to "if T is a *complete* theory and $\mathcal{A}$ is a model of T, then $T = \mathrm{Th}(\mathcal{A})$".

So any $p \in S_n(T)$ is $\mathrm{tp}(\bar{a})$ for some $\bar{a} \in A^n$.
Since A is countable, A^n is countable, and so there are only countably many complete types in $S_n(T)$.

(2) $\implies$ (1).
We will simply realize more and more types, until we catch our tail. Specifically, we will build a sequence of models $\mathcal{A}_0, \mathcal{A}_1, ...$ so that

$$(\star)\ \mathcal{A}_0 \preceq \mathcal{A}_1 \preceq \mathcal{A}_2 \preceq ...$$

and so that for any type p *with parameters in* A_i, p is realized in $\mathcal{A}_{i+1}$.

Remark: (1) Recall that if $\mathcal{A} \equiv \mathcal{B}$ then a type over $\mathcal{A}$ an be viewed as a type over $\mathcal{B}$ (it depends only on the theory. (This is for a type with no parameters.)
(2) Since $\mathcal{A}_i$ is an *elementary substructure* of $\mathcal{A}_{i+1}$, then any type p *with parameters* in $\mathcal{A}_i$ can be viewed as a type in $\mathcal{A}_{i+1}$ as well. [Exercise: why is that?]

Suppose we can build a chain as in $(\star)$, where $\mathcal{A}_0 \models T$.
Let $\mathcal{A}$ be the "union model", as in Pset 4.
In particular, $\mathcal{A}_0 \preceq \mathcal{A}$, and so $\mathcal{A} \models T$ as well.

Exercise 9.36. $\mathcal{A}$ is saturated.

Note that any type in $S_n(T)$ is already realized in A_1.
To be saturated however, we need to talk about types with arbitrary finite parameters from the model $\mathcal{A}$.
The key point is that given a finite $\bar{a}$ from A, they already appear in some A_n. In this case the type can be viewed as a type with parameters in $\mathcal{A}_n$, and by construction is is realized in $\mathcal{A}_{n+1}$.

Finally, why can we find a sequence as in $(\star)$? That is, given $\mathcal{A}_i$, why can we find an elementary extension $\mathcal{A}_{i+1}$ realizing *all types with parameters in* $\mathcal{A}_i$*?*

This is true by Lemma 9.23, as there are only countable many such types!

Why are there only countably many such types?
Note that $T = \mathrm{Th}(\mathcal{A}_i)$ and by assumption $S_n(T)$ is countable for every n.
However here we also consider types with parameters.

Recall that if p is an n-type with k parameters then it is *also* an $n+k$ type.
Since $S_{n+k}(T)$ is countable, we conclude that there are only countable many n-types with k parameters in $\mathcal{A}_i$.

The set of all types with parameters in $\mathcal{A}_i$ can be written as the union over n and k of this countable set, and therefore is countable, as a countable union of countable sets is countable (twice). $\square$

Exercise 9.37. Let $T = \mathrm{Th}(\mathbb{N}, \cdot, +, 1, 0)$. Then $S_1(T)$ is not countable.

If there is no countable saturated model, we see that there are many many different non-isomorphic models. In particular in this case $I(T)$ is infinite.

Lemma 9.38. Let T be a complete theory. Suppose that $S_n(T)$ is uncountable for some n. Then $I(T)$ is infinite. In fact $I(T)$ is uncountable.

Proof. It suffices to prove the following: given countably many models of T, $\mathcal{A}_1, \mathcal{A}_2, ...$, we need to find a model $\mathcal{A} \models T$ so that $\mathcal{A}$ is *not* isomorphic to either $\mathcal{A}_1, \mathcal{A}_2, ...$

By assumption there is some k so that $S_k(T)$ is uncountable. For notational simplicity, let us assume that $S_1(T)$ is uncountable.

Recall that isomorphic models realize the same types. Since we have *so* many types, we will find $\mathcal{A}$ which realizes types which are not realized by any of $\mathcal{A}_1, \mathcal{A}_2, ...$

For $i = 1, 2, ...$, let $P_i = \{p \in S_1(T) : p \text{ is realized in } \mathcal{A}_i\}$. Since $\mathcal{A}_i$ is countable, then each P_i is countable.
In particular $\bigcup_i P_i$ is countable, and therefore not all of $S_1(T)$.
Fix some $p \in S_1(T)$ so that $p \notin \bigcup_i P_i$.

Let $\mathcal{A} \models T$ be a countable model which realizes p. Then $\mathcal{S}$ is not isomorphic to $\mathcal{A}_i$ for any i. □

Corollary 9.39. Let T be a complete theory. Suppose that there is no countable saturated model for T. Then $I(T)$ is infinite. In fact $I(T)$ is uncountable.

Proof. By the previous theorem, if there is no countable saturated model, there is some k so that $S_k(T)$ is uncountable. □

Recall that we are going towards a proof of Vaught's theorem, that if T is a complete theory in a countable signature $\mathcal{S}$ then $I(T)$ is never 2.
In particular, we may assume that T *does have* a countable saturated model (since otherwise $I(T)$ is infinite, and therefore not 2).

So, at the very least, in this case we identified one special model for T.

9.3. **Small models.** We now identify the small models of a theory as those in which types are not realized.

We need to be a little careful however. Some types are *always realized.*

Take for example $\mathcal{A} = (\mathbb{N}, <)$, $T = \mathrm{Th}(\mathcal{A})$. Let $p = \mathrm{tp}(0) = \{\varphi(x) : \mathcal{A} \models \varphi(0)\}$.
p is infinite. However, there is a single formula $\psi(x)$ that *really captures the essence of all of* p.
Specifically, let $\psi(x) = \neg(\exists y)(y < x)$, saying that x is the minimal element in the order.

Claim 9.40. For any $\varphi(x) \in p$, $T \models (\forall x)(\psi(x) \to \varphi(x))$.

Proof. T is just the theory of $\mathcal{A} = (\mathbb{N}, <)$. In this structure, the only x satisfying $\psi(x)$ is $0 \in \mathbb{N}$. Moreover, by definition of p, for any $\varphi(x) \in p$, $\mathcal{A} \models \varphi(0)$.
So $(\forall x)(\psi(x) \to \varphi(x))$ is true in $\mathcal{A}$, and therefore is in T. □

Note also that $\psi(x) \in p$, and $(\exists x)\psi \in T$.
So any model $\mathcal{B}$ of T must have some $b \in B$ satisfying $\psi(x)$, which would imply that it realizes the entire type p.

Definition 9.41. Let T be a complete theory, p an n-type, and $\psi(x)$ a formula so that $(\exists x)\psi \in T$ ("ψ is consistent with T"). Say that $\psi(\bar{x})$ **isolates** the type p if for any $\varphi(\bar{x}) \in p$,

$$(\forall x_1)...(\forall x_n)(\psi(\bar{x}) \to \varphi(\bar{x})) \in T.$$

Note that if p is a complete type (and T is not contradictory), then it must be that $\psi(\bar{x}) \in p$. (Otherwise, its negation would be in p, and we would get $(\forall \bar{x})(\psi \to \neg\psi)$ in T.)

Say that the type p is **isolated** if there is some formula isolating it.

Lemma 9.42. Let T be a complete theory, p an n-type which is isolated. Then p is realized in any model of T.

Proof. Let $\mathcal{A} \models T$. Fix $\psi(\bar{x})$ isolating p.
Since $\psi(\bar{x}) \in p$, $(\exists x_1)...(\exists x_n)\psi(\bar{x}) \in T$, so $\mathcal{A} \models (\exists x_1)...(\exists x_n)\psi(\bar{x})$.
Fix $\bar{a} = a_1, ..., a_n$ in $\mathcal{A}$ so that $\mathcal{A} \models \psi(\bar{a})$.
For any $\varphi(\bar{x}) \in p$, by assumption, $(\forall x_1)...(\forall x_n)(\psi(\bar{x}) \to \varphi(\bar{x}))$ is in T (and so true in $\mathcal{A}$).
We conclude that for any $\varphi(\bar{x}) \in p$, $\mathcal{A} \models \varphi(\bar{a})$. That is, $\bar{a}$ realizes p in $\mathcal{A}$. □

So, isolated types are always realized. We will define a model as small if these are the only types it realizes.

Definition 9.43. Let T be a complete theory, $\mathcal{A} \models T$. ($T = \mathrm{Th}(\mathcal{A})$.) Say that $\mathcal{A}$ is **atomic** if for any $\bar{a} = a_1, ..., a_n \in A$, the type $p = \mathrm{tp}(\bar{a})$ is an isolated type.

Again, such "smallest model" does not necessarily exist. If it does exist, it is unique.

Theorem 9.44. Suppose $\mathcal{A}$ and $\mathcal{B}$ are countable atomic models with $\mathcal{A} \equiv \mathcal{B}$. Then $\mathcal{A} \simeq \mathcal{B}$.

Proof. Again this is very similar to proofs we have done before.
Suppose we have $\bar{a} = a_1, ..., a_n$ in A, $\bar{b} = b_1, ..., b_n$ in B, so that $a_i \mapsto b_i$ is a "partial isomorphism". Given any $a \in A$ we want to find some $b \in B$ so that sending a to b will "extend this partial isomorphism".
We look at $p = \mathrm{tp}(a/\bar{a})$, and want to find $b \in B$ with $p = \mathrm{tp}(b/\bar{b})$.
In the saturated case, we used the fact that *all* types are realized in $\mathcal{B}$. In particular p is realized, no matter what p is.
Here it is the opposite: since $\mathcal{A}$ is small, the type p is "trivial", in the sense that it *must* be realized in any model. In particular it is realized in $\mathcal{B}$.
More precisely: let $p = \mathrm{tp}(\bar{a}^\frown a)$.
By assumption, p is isolated, since $\mathcal{A}$ is atomic.
Fix $\psi(\bar{x}, x) \in p$ isolating p.
Let $\varphi(\bar{x}) = (\exists x)\psi(\bar{x}, x)$.
Then $\mathcal{A} \models \varphi(\bar{a})$.
By assumption, $\mathcal{B} \models \varphi(\bar{b})$. (This is the step that is very similar to what we have done before. This is an inductive assumption.)
In particular, there is some $b \in B$ so that $\mathcal{B} \models \psi(\bar{b}, b)$.
Finally, recall that p is isolated by $\psi(\bar{x}, x)$. (Here it is important that $\mathcal{A} \equiv \mathcal{B}$!)
We conclude that $(\bar{b}, b)$ realizes the type p, as we wanted. □

An almost identical proof, which we skip here, gives the following:

Theorem 9.45. Let T be a complete theory. Suppose $\mathcal{A}$ is an atomic model of T. Then for *any* model $\mathcal{B}$ of T there exists an elementary embedding f from $\mathcal{A}$ to $\mathcal{B}$.

So we may think of an atomic model as some base layer "appearing" in all models of T.
Back to the question: given a theory T, when does an atomic model ("a smallest model") exists?

Suppose $\mathcal{A}$ is an atomic model. That means, for example, that for any $a \in A$, $\mathrm{tp}(a)$ is isolated.
We may suspect that *all* types in $S_1(\mathrm{Th}(\mathcal{A}))$ are isolated.
However, that is not necessarily the case. (For example, the model $\mathcal{C}$ from Pset 5 is in fact atomic, but there *is* some isolated type, which is just not realized.)

Remark 9.46. Suppose $\mathcal{A}$ is some structure, $\bar{a} = a_1, ..., a_n$ in A, and assume that $\mathrm{tp}^{\mathcal{A}}(\bar{a})$ is isolated. Given $1 \leq i_1 < ... < i_k \leq a_n$, let $b_1 = a_{i_1}, ..., b_k = a_{i_k}$, $\bar{b} = b_1, ..., b_k$. Then $\mathrm{tp}^{\mathcal{A}}(\bar{b})$ is isolated as well.

Proof. For notational simplicity, let us consider the following case. Fix $a, b \in A$ and assume that $\mathrm{tp}^{\mathcal{A}}(a, b)$ is isolated. We prove that $\mathrm{tp}^{\mathcal{A}}(a)$ is isolated as well.

By assumption, there is a formula $\psi(x, y)$ so that for any formula $\varphi(x, y)$,

$$\text{if } \mathcal{A} \models \varphi(a, b) \text{ then } \mathcal{A} \models (\forall x)(\forall y)(\psi \to \varphi).$$

The latter implies that $\mathcal{A} \models (\forall x)((\forall y)\psi \to (\forall y)\varphi)$.

What we need to do is to find a formula $\theta(x)$ so that for any formula $\zeta(x)$,

$$\text{if } \mathcal{A} \models \zeta(a) \text{ then } \mathcal{A} \models (\forall x)(\theta \to \zeta).$$

Let $\theta(x) = (\forall y)\psi$. Fix $\zeta(x)$ so that $\mathcal{A} \models \zeta(a)$. We may view ζ as $\zeta(x, y)$. Then $\mathcal{A} \models \zeta(a, b)$. In fact, $\mathcal{A} \models (\forall x)(\zeta \leftrightarrow (\forall y)\zeta)$ (the interpretation does not depend on the "dummy variable").

By the assumption, we conclude that $(\forall x)(\theta \to \zeta)$ holds in $\mathcal{A}$, as required. $\square$

There is much to say about atomic models, and non-isolated types. You can find more in [Marker, Hodges].

For now, the following will be useful to prove Vaught's theorem.

Theorem 9.47. Fix a countable signature $\mathcal{S}$ and a complete theory T. Assume that there is a countable saturated model for T. Then there is a countable atomic model for T.

The generality of this result is quite surprising. These questions, of finding a "largest countable model" (a model that every other one embeds into it), or "a smallest countable model" (a model which embeds into any other one), are quite natural, given some theory T. No matter which complete theory you are working with, if you can find a saturated model, then there is also an atomic one. (They may be isomorphic, some times.)

Remark 9.48. Another way to phrase the theorem: suppose $\mathcal{A}$ is a saturated structure. Then $\mathrm{Th}(\mathcal{A})$ has an atomic model. (May or may not be isomorphic to $\mathcal{A}$.)

Proof sketch of Theorem 9.47. We will repeat the construction of a Henkin model, with additional conditions, so that the final model is in fact atomic.

Recall that we add new constant symbols $\mathcal{S}^+ = \mathcal{S} \cup \{c_0, c_1, c_2, ...\}$ and build a theory T^+ for $\mathcal{S}^+$ which satisfies all the Henkin conditions.
We construct a model $\mathcal{A}^+$ for T^+ whose universe is precisely $\{c_1, c_2, ...\}$ (a quotient of it).
In this model, for any formula $\varphi(x_1, ..., x_n)$, $\mathcal{A}^+ \models \varphi(c_1^{\mathcal{A}^+}, ..., c_n^{\mathcal{A}^+})$ if and only if $\varphi[c_1, ..., c_n] \in T^+$.

By an earlier remark, it suffices to make sure that for arbitrary large n, $\mathrm{tp}^{\mathcal{A}^+}(c_1, ..., c_n)$ is isolated. Note that $\mathrm{tp}^{\mathcal{A}^+}(c_1, ..., c_n)$ is in $S_n(T)$.

We will make sure that these types are isolated by making sure that they are not not isolated.

Given some type $p \in S_n(T)$, *if p is a non isolated type* then we will want to find some $\psi(x_1, ..., x_n)$ in p so that ψ fails for $c_1, ..., c_n$. That is, so that $\neg\psi[c_1, ..., c_n] \in T^+$.

We will simply add this to our infinite tree construction, so that any infinite branch in the final tree will satisfy this extra assumption:

If $p \in S_n(T)$ is *not* isolated, then there is some ψ in p so that $\neg\psi[c_1, ..., c_n] \in T^+$.

There is some additional "book-keeping" to do. This book-keeping is still possible since we only have countably many types to worry about!

On top of that, it was important that all of our "add this" steps do not make the theory we are constructing so far (the finite branch) inconsistent. We need to show:

Claim 9.49. Let $\zeta_1, ..., \zeta_k$ be $\mathcal{S}^+$ sentences so that $T \cup \{\zeta_1, ..., \zeta_k\}$ is consistent (there is no proof of contradiction). Let $p \in S_n(T)$ be a non-isolated type.

Then we may find some $\psi \in p$ so that $T \cup \{\zeta_1, ..., \zeta_k\} \cup \{\neg\psi[c_1, ..., c_n]\}$ is still consistent.

Proof. We may assume that the new constant symbols appearing in $\zeta_1, ..., \zeta_k$ are contained in $c_1, ..., c_n$. (Otherwise, we may replace p with a type $q \in S_l(T)$ for a larger l, so that q extends p.)
Let $\bar{x} = x_1, ..., x_n$, $\bar{c} = c_1, ..., c_n$.
Fix formulas $\theta_1(\bar{x}), ..., \theta_k(\bar{x})$ so that $\zeta_i = \theta_i[\bar{c}]$.

What does it mean for $T \cup \{\zeta_1, ..., \zeta_k\} \cup \{\neg\psi[c_1, ..., c_n]\}$ to be *inconsistent*?
That there is a proof of contradiction $T \cup \{\theta_1[\bar{c}], ..., \theta_k[\bar{c}]\} \cup \{\neg\psi[\bar{c}]\} \vdash \perp$.
In other words, $T \cup \{\theta_1[\bar{c}] \wedge ... \wedge \theta_k[\bar{c}] \wedge \neg\psi[\bar{c}]\} \vdash \perp$.
That is, $T \vdash \neg(\theta_1[\bar{c}] \wedge ... \wedge \theta_k[\bar{c}] \wedge \neg\psi[\bar{c}])$.
Recall that $\neg(\phi_1 \wedge \neg\phi_2)$ is equivalent to $\neg\phi_1 \vee \neg\neg\phi_2$, which is equivalent to $\neg\phi_1 \vee \phi_2$, which is equivalent to $\phi_1 \rightarrow \phi_2$.

In conclusion: $T \vdash (\theta_1[\bar{c}] \wedge ... \wedge \theta_k[\bar{c}]) \rightarrow \psi[\bar{c}]$.
By the completeness theorem, this is equivalent to $T \models (\theta_1[\bar{c}] \wedge ... \wedge \theta_k[\bar{c}]) \rightarrow \psi[\bar{c}]$.
By Pset 4, question 5, this is equivalent to $T \models \forall\bar{x}((\theta_1 \wedge ... \wedge \theta_k) \rightarrow \psi)$.

So, not being able to add ψ to the theory we are building up (in $\mathcal{S}^+$), precisely corresponds to the ψ "being isolated" by the formula $\theta_1 \wedge ... \wedge \theta_k$.

Finally, since p is assumed to be *not isolated*, then can find some ψ which is "not isolated by $\theta_1 \wedge ... \wedge \theta_k$". So this ψ works for the claim. $\square$

$\square$

Having "all types isolated" does happen, for example in $(\mathbb{Q}, <)$. More generally, this happens if and only if $I(T) = 1$, in which case there is a saturated model and an atomic model, and they are in fact equal.

Theorem 9.50. Fix a countable signature $\mathcal{S}$ and let T be a (satisfiable) complete theory. The following are equivalent.

(1) $I(T) = 1$, that is, for any two models $\mathcal{A}, \mathcal{B} \models T$, $\mathcal{A}$ and $\mathcal{B}$ are isomorphic.
(2) For all n, any type in $S_n(T)$ is isolated.
(3) $S_n(T)$ is a finite set for all n.

Proof. (1) $\implies$ (2).
Assume $I(T) = 1$ (all countable models are isomorphic to one another). In particular, there exists a saturated model $\mathcal{B} \models T$. (Otherwise, $I(T)$ is infinite.)
Therefore there also exists an atomic model $\mathcal{A} \models T$.
By assumption, $\mathcal{A} \simeq \mathcal{B}$.

Since $\mathcal{B}$ is saturated, any type $p \in S_n(T)$ is realized in $\mathcal{B}$.
Therefore every type is realized in $\mathcal{A}$.
Since $\mathcal{A}$ is atomic, the types realized in $\mathcal{A}$ are all isolated.
So every type is isolated.

(2) $\implies$ (1)
Assume that every type is isolated. Let $\mathcal{A}, \mathcal{B}$ be a models of T.
By assumption, both $\mathcal{A}$ and $\mathcal{B}$ must be atomic.
By the uniqueness of an atomic model, $\mathcal{A} \simeq \mathcal{B}$.

Remark 9.51. Suppose $p_1(x), ..., p_k(x)$ are isolated types in $S_1(T)$. Let $\psi_1(x), ..., \psi_k(x)$ be formulas isolating them.

Consider the sentence $(\forall x)(\psi_1 \vee ... \vee \psi_k)$. Is it in T?
If it is in T, then $p_1, ..., p_k$ are precisely all types in $S_1(T)$! That is, $S_1(T) = \{p_1, ..., p_k\}$.

Otherwise, if $p_1, ..., p_k$ are not all 1-types, then $(\exists x)(\neg\psi_1 \wedge ... \wedge \neg\psi_k)$ is in T.

(2) $\implies$ (3).
Assume that every type is isolated.
Assum for a contradiction that $S_n(T)$ is not finite, for some n.
For notational simplicity, let us assume that $S_1(T)$ is not finite.

So we have may list S_1 as $p_1, p_2, p_3, ...$, all different complete 1-types.
Fix a formula ψ_i which isolates the type p_i.
Expand $\mathcal{S}$ by a new constant symbol, $\mathcal{S}^+ = \mathcal{S} \cup \{c\}$, and consider the theory $T^+ = T \cup \{\neg\psi_i[c] : i = 1, 2, ...\}$.

First note that T^+ is not satisfiable, since $p_1, p_2, ...$ lists *all* 1-types in $S_1(T)$.
Indeed, if $\mathcal{A}^+$ is an $\mathcal{S}^+$-structure satisfying T^+, let $\mathcal{A}$ be its reduct to $\mathcal{S}$.
Then $\mathcal{A} \models T$. In particular, for any $a \in A$, $\mathrm{tp}^{\mathcal{A}}(a)$ must be in $S_1(T)$.
Let $a = c^{\mathcal{A}^+}$. Since $\mathcal{A}^+ \models T^+$, it follows that $\psi_i(a)$ fails in $\mathcal{A}$, for each i. Therefore $\mathrm{tp}^{\mathcal{A}}(a) \neq p_i$ for each i. A contradiction.

Finally, we show that T^+ is finitely satisfiable, leading to a contradiction (by the compactness theorem).
Given a finite $T_0 \subseteq T^+$, T_0 is contained in $T \cup \{\neg\psi_1[c], ..., \neg\psi_k[c]\}$ for some finite k.
To show that this theory is satisfiable, we need to show that there is a model for T in which $(\exists x)(\neg\psi_1 \wedge ... \wedge \neg\psi_k)$ holds.
By the remark above, this is true in any model of T.

(3) $\implies$ (2).
Again for notational simplicity let us work with $S_1(T)$. Assume that $S_1(T)$ is finite. We want to show that every $p \in S_1(T)$ is isolated.

Fix $p_1, p_2, ..., p_k$ a list of *all* 1-types in $S_1(T)$.
Since they are distinct, we may find for $i < j$ a formula $\theta_{i,j}(x)$ so that $\theta_{i,j} \in p_i$ yet $\neg\theta_{i,j} \in p_j$.

For $j < i$ define $\theta_{i,j} = \neg\theta_{j,i}$. Define

$$\varphi_i(x) = \bigwedge_{j \neq i} \theta_{i,j}.$$

By asssumption, each $\theta_{i,j}$ is in p_i, so φ_i is in p_i.

We claim that φ_i isolates p_i.

For any model $\mathcal{A}$ for T, for any $a \in A$, if $\mathcal{A} \models \varphi_i(a)$, then $\mathrm{tp}^{\mathcal{A}}(a)$ is not p_j for $j \neq i$.
Since $\mathrm{tp}^{\mathcal{A}}(a)$ must be one of $p_1, ..., p_k$(by assumption), it will necessarily be p_i.

That is, for any $\mathcal{A} \models T$, for any $a \in A$, if $\mathcal{A} \models \varphi_i(a)$ then $\mathcal{A} \models \psi(a)$ for all $\psi \in p$.
In particular, for any ψ in p, $T \models (\forall x)(\varphi_i \to \psi)$. So φ_i isolates ψ.

□

Collecting what we have seen so far, here are some things we can say about the countable models of a (satisfiable) complete theory T by looking at the types $S_n(T)$:

- If $S_n(T)$ is uncountable for some n, then $I(T)$ is uncountable: there are uncountably many non-isomorphic models of T. ("Very chaotic behavior").
- If $S_n(T)$ is countable for each n, then we may identify two special countable models: a saturated one and an atomic one.
- $S_n(T)$ is in fact finite (for every n) if and only if the saturated and atomic models are isomorphic to one another. In this case $I(T) = 1$.

9.4. **Proof of Vaught's theorem.** Recall Vaught's theorem:

Fix a countable signature $\mathcal{S}$ and a complete theory T. Then $I(T) \neq 2$.
In other words, if we may find two non-isomorphic models $\mathcal{B}, \mathcal{C}$ of T, then there is a third model $\mathcal{A} \models T$ which is not isomorphic to either $\mathcal{B}$ or $\mathcal{C}$.

First, if $I(T)$ is infinite, there is nothing to prove. We may therefore assume that T has a saturated model $\mathcal{B}$ and an atomic model $\mathcal{C}$. (If either does not exists, then $I(T)$ is uncountable.)
If $\mathcal{B} \simeq \mathcal{C}$, then $I(T) = 1$, so again, we are done.

Assume that $\mathcal{B} \not\simeq \mathcal{C}$ (that is, $I(T) \neq 1$).
Then there is some n so that $S_n(T)$ is not finite, and has a non-isolated type in it.
For notational simplicity, let us assume that $S_1(T)$ is infinite, and so there is a non-isolated type p in $S_1(T)$.

Since $\mathcal{B}$ is saturated, p *is* realized in $\mathcal{B}$. Let $b \in B$ be a realization, $\mathrm{tp}^{\mathcal{B}}(b) = p$.

Let c be a new constant symbol. Expand $\mathcal{B}$ to $\mathcal{B}^+$ by $c^{\mathcal{B}^+} = b$. Let $T^+ = \mathrm{Th}(\mathcal{B}^+)$ (the $\mathcal{S}^+$-theory of $\mathcal{B}^+$, where $\mathcal{S}^+ = \mathcal{S} \cup \{c\}$).

Note that $\mathcal{B}^+$ is still a saturated model (this time for T^+). (A type with parameter $\bar{d}$ for $\mathcal{B}^+$ can be viewed as a type with parameter $(\bar{d}, b)$ for $\mathcal{B}$. Since $\mathcal{B}$ is saturated, this type is realized.)
In particular, T^+ has a saturated model, therefore it has an atomic model.

Note that $\{\varphi[c] : \varphi(x) \in p\} \subseteq T^+$. So *any* model of T^+ realizes the type p. Specifically, the interpretation of the constant symbol c always realizes the type p.

Let $\mathcal{A}^+$ be an $\mathcal{S}^+$-structure, $\mathcal{A} \models T^+$ an atomic model.
Let $\mathcal{A}$ be the reduct of $\mathcal{A}^+$ to $\mathcal{S}$.

Claim 9.52. $\mathcal{A}$ is not isomorphic to either $\mathcal{B}$ or $\mathcal{C}$.

First, the type p *is realized in* $\mathcal{A}$, by $a = c^{\mathcal{A}^+}$. So $\mathcal{A}$ is not isomorphic to $\mathcal{C}$ ($\mathcal{C}$ is atomic, and p is isolated, so p is not realized in $\mathcal{C}$).

What about $\mathcal{A}$ and $\mathcal{B}$?
The idea is that if they are isomorphic, then $I(T^+) = 1$.
However, this would mean that $S_1(T^+)$ is finite.
However, every type in $S_1(T)$ (which is infinite) extends to a type in $S_1(T^+)$. A contradiction!

More formally:
Since $S_1(T)$ is infinite, so is $S_1(T^+)$.
By the theorem, there is some type $q^+ \in S_1(T^+)$ which is not isolated.
In particular q^+ is not realized in $\mathcal{A}$.

We may view q^+ as a 1-type q *with parameter* $a = c^{\mathcal{A}^+}$ in $\mathcal{A}$. (Let $q(x) = \{\theta(x, y) : \theta(x)[c] \in q^+\}$).
So q is a 1-type with parameter, in $\mathcal{A}$, which is *not realized in* $\mathcal{A}$.
In other words, the model $\mathcal{A}$ is not saturated.
It follows that $\mathcal{A}$ is not isomorphic to $\mathcal{B}$.

Some thoughts:
- Why does the proof *not* work to find a 4'th model?
- Try to think of what the above construction looks like in our example with $\mathcal{B} = (\mathbb{Q}\setminus\{0\}, <, 1, \frac{1}{2}, ...)$ and $\mathcal{C} = (\mathbb{Q}^+, <, 1, \frac{1}{2}, ...)$.

10. What's next?

If you are particularly interested in model theory: at this point you should be able to pick up any graduate textbook on model theory[6]. Two examples are [Marker, Hodges]. In Section 11 you can find a few concrete starting points.
On our Canvas page you may find some further suggested very advanced topics. For those mostly some further reading is necessary beforehand.

10.1. **Math 141B is offered in the Fall!** The material of 141B is focused on Godel's *incompleteness theorem* and related topics, specifically notions of complexity / computability.

Ever since Section 9 (types) our setup was as follows: we considered an *arbitrary complete* theory T and studied its models. This is the very basis for "model theory". Also, to some extend this is what we do in many math classes, just for a very particular theory T.

For the majority of our concrete examples, the complete theory T was fairly well understood.

- For $T = \mathrm{Th}(\mathbb{Q}, <)$, T is in fact equal to Con(DLO), just the formal consequences of the (finitely many) DLO axioms.
- For $T = \mathrm{Th}(\mathbb{Q}, <, 1, \frac{1}{2}, ...)$, T is in fact equal to $\mathrm{Con}(\mathrm{DLO} \cup \{c_{n+1} < c_n : n = 1, 2, ...\})$. In this case, there are infinitely many "basic axioms", but this is a very "simple collection of axioms": it is clear how to determine which sentence is an axiom (if it is of the form $c_{n+1} < c_n$...).

[6] Slight caveat: some more background on infinite cardinalities will be very useful for most material beyond what we have seen in this course.

- For $T = \mathrm{Th}(\mathbb{C}, 0, 1, \cdot, +)$, it turns out that $T = \mathrm{Con}(\mathrm{ACF}_0)$, the consequences of the axioms of "an algebraically closed field of characteristic 0". (We did not prove this however.) Here there are infinitely many axioms, but very simply to describe.
- Let $T = \mathrm{Th}(\mathbb{N}, 0, 1, +, \cdot)$. This is of course a very interesting theory! Essentially for any statement in number theory, one wants to know if it is in T or not.

Question 10.1. Can we write some axioms $\theta_1, \theta_2, \ldots$, so that $T = \mathrm{Th}(\mathbb{N}, 0, 1, +, \cdot)$ is precisely the consequences of $\{\theta_1, \theta_2, \ldots\}$?

We can of course take any enumeration of T... but that doesn't seem to be meaningful. [It does not really help us understand whether any particular question in number theory is true or false... since we need to already know this in order to decide whether to put this statement in the enumeration or not.]

So the question should be: can we *describe* a *reasonable* list of axioms?
What is reasonable?
For example, finite is reasonable. But we already saw some reasonable infinite collections of axioms.

A (reasonable) intuitive definition for "reasonable" is as follows:
suppose you can write a computer program which takes as input a sentence ϕ, and outputs "YES" if this ϕ is one of our axioms, and outputs "NO" if it is not.
For example, for the list of sentences $c_{n+1} < c_n$, it is clearly doable.

In 141B we will see that it is impossible to find such "reasonable axiomatization" for number theory.

Theorem 10.2. Suppose $T_0 = \{\theta_1, \theta_2, \ldots\}$ is a "reasonable" (computable) list of axioms. Then $\mathrm{Con}(T)$ is *not* $\mathrm{Th}(\mathbb{N}, 0, 1, +, \cdot)$.

In particular the collection of "number theoretic truths" $\mathrm{Th}(\mathbb{N}, 0, 1, +, \cdot)$ itself "is not computable"!

Part of the developments in 141B will be to define what "computable" means.
The intended meaning will always "what a computer can do".
The formalization of this concept is the basics for complexity theory, (in mathematics as well as in computer science).

It is worth mentioning that when Godel developed all these things, almost a 100 years ago, there were no computers.
At the time, just arguing that there is any reasonable notion of "computable" was highly non-trivial!
The development of algorithms, computers, and computer science heavily relies on Godel's work.

Back to the topic: suppose you *try* to axiomatize number theory using a "reasonable" computable list of axioms $T_0 = \{\theta_1, \theta_2, \ldots\}$, so that each θ_i is a *true* (known) statement in number theory.
One such reasonable collection of axioms are the axioms of Peano Arithmetic (will be discussed thoroughly in 141B).
If you believe the discussion above, then $\mathrm{Con}(T_0)$ is not $\mathrm{Th}(\mathbb{N}, 0, 1, +, \cdot)$.
In this case it is necessarily a strict subset, $\mathrm{Con}(T_0) \subsetneq \mathrm{Th}(\mathbb{N}, 0, 1, +, \cdot)$.
In particular $\mathrm{Con}(T_0)$ is *not a complete theory.* This is a part of the **incompleteness**

theorem: any "reasonable list" of axioms ends up being not complete: there is always some θ so that neither θ nor $\neg\theta$ is a logical consequence of T_0.

Given a particular list of axioms, you may ask what is this θ, that refuses to be decided by our axioms? The most common phrasing of the incompleteness theorem is as follows:

Suppose T_0 is a "reasonable" list of number-theoretic axioms. Then there is a sentence θ which "says" that "T_0 is a consistent theory" (has no formal contradiction), and neither θ nor $\neg\theta$ are a logical consequence of T_0.

That is:

A theory cannot prove its own consistency!

There is much to say here. For example, how can a sentence in the language $+, \cdot, 0, 1$ talk about something being consistent or not? There is a lot of coding to do. Again this is an important part of the technical developments in 141B, and an important tool in the study of complexity / computability.

11. Topics for further reading

Below are some topics for further optional readings, depending on your interest. For these topics the notes provides a reference which you are in a position to read.

Feel free to ask about any one of these!

(Another list of more (very) advanced topics for further reading will be updated on Canvas. For some of these advanced topics, some of these optional reading topics will be a first step.)

- We discussed in Pset 6 (not for submission) that the reals $\mathbb{R}$ can be "extended by an infinitesimal".
  Much like the infinite Ramsey theorem was simpler than the finite one (omitting the need of "for all h there is N"), one can use this infinitesimal to provide simpler definitions of limits, continuity and differentiability of funcitons. (Avoiding some "alternating quantifiers". Recall Pset 2 question 3(1).)
  You can read about this "non-standard analysis" in [Enderton, Section 2.8].
- Characterizing elementary equivalence $\equiv$ in terms of winning strategy in finite length games.
  See reference in page 30.
- Fraisse limits: given a collection of finite objects (such as all finite linear orders, or all finite graphs), when can we make sense of a "limit object"? The ordered rational numbers $(\mathbb{Q}, <)$ and the random graph are examples of such limit objects.
  See reference on page 31
- Quantifier elimination: when formulas "are equivalent to" quantifier-free formulas. This happens, for example, when dealing with algebraically closed fields, or DLOs. This is another instance when substructures are necessarily elementary substructures, as we have seen for DLOs.
  See references in Remark 5.6 on page 32.

References

[Enderton] Herbert B. Enderton - A Mathematical Introduction to Logic
[Woodin-Slaman] Notes by Professor W. Hugh Woodin and Professor Theodore A. Slaman
[Marker] David Marker - Model Theory: An Introduction

[Hodges] Wilfrid Hodges - Model Theory