MATH681
Mathematical Logic

Pingbang Hu

March 27, 2023

Abstract

This is a graduate-level mathematical logic course taught by Matthew Harrison-Trainor, aiming to obtain insights into all other branches of mathematics, such as algebraic geometry, analysis, etc. Specifically, we will cover model theory beyond the basic foundational ideas of logic.

While there are no required textbooks, some books do cover part of the material in the class. For example, Marker's *Model Theory: An Introduction* [Mar02], Hodges's *A Shorter Model Theory* [HH97], and Hinman's *Fundamentals of Mathematical Logic* [Hin05].

This course is taken in Winter 2023, and the date on the covering page is the last updated time.

Contents

Chapter 1

Language, Logic, and Structures

Lecture 1: Introduction to Mathematical Logic

The goal of mathematical logic is to obtain insights into other areas of mathematics – algebra, analysis, combinatorics, and so on, by formalizing the **process** of mathematics. 5 Jan. 14:30

Remark. More concretely, there are different branches:

(a) Model Theory: Study subsets of an object defined by a formula (i.e., first-order logic).

(b) Computability Theory / Recursion Theory: Formalizing what it means to have an algorithm and studying relative computability.

(c) Set Theory: Study the structure of the mathematical universe.

(d) Proof Theory: Study the syntactic nature of proofs.

In this class, we study model theory in nature; specifically, we will cover

- basic definitions of logic:
  - What is a formula?
  - What does it mean for a formula to be true?
  - What is a proof?
- Soundness & completeness theorems:
  - Anything provable is true.
  - Anything true is provable.
- Compactness theorem:
  - Non-standard objects exist.
- Using compactness theorem for applications:
  - Chevalley's theorem.

The main theme of this course will be *syntax* v.s. *semantics*:

Syntax	v.s.	Semantics
proofs		truth
form of a formula		mathematical structures
number and type of quantifiers		isomorphisms, embeddings

And this is what this chapter aiming to address. We will turn to other topics based on these.

1.1 Languages and Structures

Let's start with the fundamental object, language.

Definition 1.1.1 (Language). A *language* $\mathcal{L}$ consists of:

- a set $\mathcal{F}$ of function symbols f with arities n_f;
- a set $\mathcal{R}$ of relation symbols R with arities n_R;
- a set $\mathcal{C}$ of constant symbols c.

A language is also sometimes called a *signature*, in which case we use σ rather than $\mathcal{L}$.

Note. A constant is the same as a 0-ary function.

Remark. Any or all sets in Definition 1.1.1 might be empty.

Example (Graph). The language of graphs, $\mathcal{L}_{\text{graph}} = \{E\}$ where E is a binary (2-ary) relation symbol.

Example (Ring). The language of rings, $\mathcal{L}_{\text{ring}} = \{0, 1, +, \cdot, -\}$, where $0, 1$ are constants, $+, \cdot$ are binary functions, and $-$ is a unary function.

Example (Ordered ring). The language of ordered rings, $\mathcal{L}_{\text{ord}} = \mathcal{L}_{\text{ring}} \cup \{\leq\}$ where $\leq$ is the binary relation for an ordered ring.

Then, given a language, we can now interpret it in the following way.

Definition 1.1.2 (Structure). Given a language $\mathcal{L}$, an *$\mathcal{L}$-structure* $\mathcal{M}$ consists of:

- a non-empty set M called the *universe*, *domain*, or *underlying set* of $\mathcal{M}$;
- for each function symbol $f \in \mathcal{F}$, a function $f^{\mathcal{M}} \colon M^{n_f} \to M$;
- for each relation symbol $R \in \mathcal{R}$, a relation $R^{\mathcal{M}} \subseteq M^{n_R}$;
- for each constant symbol $c \in \mathcal{C}$, an element $c^{\mathcal{M}} \in M$.

Notation (Interpretation). The *interpretation* of symbols f, R, c in $\mathcal{M}$ is $f^{\mathcal{M}}, R^{\mathcal{M}}, c^{\mathcal{M}}$, respectively.

Basically, a structure gives meaning to the symbols from the language, and we often write

$$\mathcal{M} = (M, f^{\mathcal{M}}, \dots, R^{\mathcal{M}}, \dots, c^{\mathcal{M}}, \dots) = (M, f^{\mathcal{M}}, R^{\mathcal{M}}, c^{\mathcal{M}} \colon f \in \mathcal{F}, R \in \mathcal{R}, c \in \mathcal{C}).$$

Notation. We usually use $\mathcal{M}, \mathcal{N}, \dots, \mathcal{A}, \mathcal{B}, \dots$ to refer to structures, and $M, N, \dots, A, B, \dots$ for the domains.[a]

[a] Some people use $|\mathcal{M}|$ for the domain of $\mathcal{M}$.

It's time to look at some examples.

Example. The rationals $\mathbb{Q}$ and integers $\mathbb{Z}$ are both $\mathcal{L}_{\text{ring}}$-structures.

Proof. Clearly, the domain is the set of rationals, and naively, we let $+^{\mathbb{Q}} = +$ in $\mathbb{Q}$, $0^{\mathbb{Q}} = 0$ in $\mathbb{Q}$, $1^{\mathbb{Q}} = 1$ in $\mathbb{Q}$, etc. In this way, $\mathbb{Q} = (\mathbb{Q}, 0, 1, +, \cdot, -)$ is an $\mathcal{L}_{ring}$-structure. Similarly, $\mathbb{Z} = (\mathbb{Z}, 0, 1, +, \cdot, -)$ is as well. ⊛

While the language we have seen are all intuitively correct with their name, e.g., $\mathcal{L}_{\text{ring}}$, $\mathcal{L}_{\text{ord}}$, and $\mathcal{L}_{\text{graph}}$, they are really just the high-level abstraction of the objects in the subscript.

Example. Nothing forces an $\mathcal{L}_{\text{ring}}$-structure to be a ring.

Proof. Since an $\mathcal{L}_{\text{ring}}$-structure is just any structure with two binary functions, a unary function, and two constants interpreting the symbols of the language; hence we can define an $\mathcal{L}_{\text{ring}}$-structure $\mathcal{M}$ as

- $\mathcal{M} = \{0, 5, 11\}$;
- $0^{\mathcal{M}} = 5$;
- $1^{\mathcal{M}} = 11$;
- $+^{\mathcal{M}}$ is the constant function 0;
- $\cdot^{\mathcal{M}}$ is the function 5;
- $-^{\mathcal{M}}$ is the identity.

This is clearly not a ring since it fails nearly every axiom of a ring. ⊛

Note. Later, we will talk about theories that let us restrict to structures we want.

1.2 Embeddings and Isomorphisms

We can now consider the relation between structures.

Definition 1.2.1 (Embedding). Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. A map $\eta\colon \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}$-*embedding* if it is one-to-one and preserves the interpretation of all symbols of $\mathcal{L}$:

(a) for each function symbol $f \in \mathcal{F}$ of arity n_f, and $a_1, \dots, a_{n_f} \in M$,

$$\eta\left(f^{\mathcal{M}}(a_1, \dots, a_{n_f})\right) = f^{\mathcal{N}}\left(\eta(a_1), \dots, \eta(a_{n_f})\right);$$

(b) for each relation symbol $R \in \mathcal{R}$ of arity n_R, and $a_1, \dots, a_{n_R} \in M$,

$$(a_1, \dots, a_{n_R}) \in R^{\mathcal{M}} \Leftrightarrow (\eta(a_1), \dots, \eta(a_{n_R})) \in R^{\mathcal{N}};$$

(c) for each constant symbol $c \in \mathcal{C}$, $c^{\mathcal{M}} = c^{\mathcal{N}}$.

From the definition, an $\mathcal{L}$-embedding is an injection, and naturally, we have the following.

Definition 1.2.2 (Isomorphism). An $\mathcal{L}$-*isomorphism* is a bijective $\mathcal{L}$-embedding.

Definition 1.2.3 (Automorphism). An $\mathcal{L}$-*automorphism* of $\mathcal{M}$ is an $\mathcal{L}$-isomorphism from $\mathcal{M}$ to $\mathcal{M}$.

Definition. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. Suppose $M \subseteq N$ and the inclusion map $\iota\colon \mathcal{M} \hookrightarrow \mathcal{N}$ is an $\mathcal{L}$-embedding.

Definition 1.2.4 (Substructure). $\mathcal{M}$ is a *substructure* of $\mathcal{N}$.

Definition 1.2.5 (Extension). $\mathcal{N}$ is an *extension* of $\mathcal{M}$.

Example. Ring embeddings are $\mathcal{L}_{\text{ring}}$-embeddings.

This generalizes the notions of embedding and isomorphism for many mathematical structures.

Remark. Asking that η be injective is the same as (b) in Definition 1.2.1 for the relation $=$ since

$$a = b \in \mathcal{M} \Leftrightarrow \eta(a) = \eta(b) \in \mathcal{N}.$$

The notion of substructure is language sensitive. For groups, there are two possible languages:

(a) $\mathcal{L}_1 = \{e, \cdot\}$;

(b) $\mathcal{L}_2 = \{e, \cdot, ^{-1}\}$, i.e., with the unary inverse operation.

While both seem valid at the first glance, we should use the second one.

To see why, if we use $\mathcal{L}_2$, the substructure of a group is the same thing as a subgroup. But if we use $\mathcal{L}_1$, then $(\mathbb{N}, +, 0)$ is a substructure of $(\mathbb{Z}, +, 0)$, while $\mathbb{N}$ is not a group for sure.[1]

Similarly, we include $-$ in $\mathcal{L}_{\text{ring}}$ for a similar reason as in the previous example.

Example. An $\mathcal{L}_{\text{ring}}$-substructure of a field will be a subring, not a subfield. If we want subfields, use $\mathcal{L}_{\text{ring}} \cup \{^{-1}\}$.[a]

[a]We can set $0^{-1} = 0$, but never use this.

Lecture 2: Formulas and First-Order Logic

We start by asking that given a function symbol f of arity n, could we replace f with an $(n+1)$-ary R relation to represent its graph? 10 Jan. 14:30

Example. Let $\mathcal{L}$ be a language with only relation symbols. Let $\mathcal{A}$ be an $\mathcal{L}$-structure. For any $B \subseteq A$, there is a substructure $\mathcal{B}$ of $\mathcal{A}$ with domain B.

Proof. For each relation symbol R, leting $R^{\mathcal{B}} = R^{\mathcal{A}} \cap B^{n_R}$ will make $\mathcal{B}$ a substructure of $\mathcal{A}$. ⊛

The above is not true for function symbols though.

Example. If $G = (\mathbb{Z}, 0, +)$, then $\mathbb{N}$ is not the domain of a subgroup. So if we took $\mathcal{L} = \{0, +, ^{-1}\}$, where 0 is the unary relation, $+$ is the ternary relation, and $^{-1}$ is the binary relation, an $\mathcal{L}$-substructure of a group might not be a subgroup.

1.3 Terms

Intuitive, an $\mathcal{L}$-formula is an expression built using the symbols in a language $\mathcal{L}$, $=$, the logical connectives $\wedge, \vee, \neg$, and variable symbols $v_1, v_2, \ldots, x, y, z$, and also quantifiers $\exists$ and $\forall$.

Definition 1.3.1 (Term). Given a language $\mathcal{L}$, the set of $\mathcal{L}$-*terms* are defined inductively by:

(a) each constant symbol is a *term*;

(b) each variable symbol $v_1, \ldots$ is a *term*;

(c) if f is a function symbol, and $t_1, \ldots, t_{n_f}$ are terms, then $f(t_1, \ldots, t_{n_f})$ is a *term*.

If $\mathcal{M}$ is an $\mathcal{L}$-structure, and t is a term involving only variables among $v_1, \ldots, v_n$, then t has an interpretation $t^{\mathcal{M}} \colon M^n \to M$ as a function as follows. On input $a_1, \ldots, a_n \in M$,

(a) if t is a constant c, $t^{\mathcal{M}}(a_1, \ldots, a_n) = c^{\mathcal{M}}$.

[1]Simply observe that both $(\mathbb{N}, 0, +), (\mathbb{Z}, 0, +)$ are $\mathcal{L}_1$-structures.

(b) if t is a variable v_i, $t^{\mathcal{M}}(a_1, \ldots, a_n) = v_i$;

(c) if t is $f(s_1, \ldots, s_k)$, then $t^{\mathcal{M}}(a_1, \ldots, a_n) = f^{\mathcal{M}}\big(s_1^{\mathcal{M}}(a_1, \ldots, a_n), \ldots, s_k^{\mathcal{M}}(a_1, \ldots, a_n)\big)$.

Intuition. We are basically substituting for variables and evaluating the expression.

Example. In $(\mathbb{R}, 0, 1, +, \cdot, -)$, a term is essentially just a polynomial with integer coefficients, assuming we interpret them in a ring. Technically, a term looks like

$$\cdot(+(1,1), +(x,y)),$$

but we will write terms the natural way, i.e.,

$$(1+1)(x+y).$$

Also, we will use $\underline{n}$ or n to represent the term $\underline{n} = \underbrace{1+1+\cdots+1}_{n \text{ times}}$. So we could write the above term as $2 \cdot (x+y)$.

1.4 Formulas

A term is just a building block of formulas, as we now see.

Definition 1.4.1 (Formula)**.** The set of $\mathcal{L}$-*formulas* is defined inductively:

(a) If s, t are terms, then $s = t$ is a *formula*.

(b) If R is a relation symbol of arity n_R and $s_1, \ldots, s_{n_R}$ are terms, then $R(s_1, \ldots, s_{n_R})$ is a *formula*.

(c) If f is a formula, then $\neg f$ is a *formula*.

(d) If φ and ψ are formulas, then $\varphi \wedge \psi$ and $\varphi \vee \psi$ are *formulas*.

(e) If φ is a formula and v_i are variables, then $\exists v_i\ \varphi$ and $\forall v_i\ \varphi$ are *formulas*.

Notation (Atomic)**.** Definition 1.4.1 (a) and (b) are called *atomic*.

Notation (Quantifier-free)**.** Definition 1.4.1 (a), (b), (c), and (d) are called *quantifier-free*.

This logic is called *first-order logic* (FO logic), since the quantifiers range over elements of the structures, but not over, e.g., subsets.

Example. We can say that an element x of a ring has a square root by $\exists y\ y^2 = x$.

Example. A group is torsion of order 2 can be said by $\forall x\ x \cdot x = e$.

Example. We can write down all the field/group/... axioms as formulas.

1.4.1 Bounded and Free Variables

Notice that for the first example, the formula $\exists y\ y^2 = x$ only has meaning if we assign what x is. In this case, we say that y is *bound* by $\exists y$. But this is local:

Example. Consider

$$y = 1 \land \exists y\ y^2 = x,$$

while the first appearance of y is free, the second appearance of y is bound by (in the scope of) $\exists y$.

While our definitions work perfectly fine with the above example, but sometimes we don't want this to happen. In such a case, we simply replace the bound instances of y with a new variable z. This idea of variables being free or bound is defined formally as follows.

Definition 1.4.2 (Free variable)**.** The *free variables* $\mathrm{FV}(\varphi)$ of a formula φ are defined inductively:

(a) $\mathrm{FV}(s = t)$ is the set of variables showing up in s or t.

(b) $\mathrm{FV}(R(s_1, \ldots, s_{n_R}))$ is the set of variables showing up in $s_1, \ldots, s_{n_R}$.

(c) $\mathrm{FV}(\neg\varphi) = \mathrm{FV}(\varphi)$.

(d) $\mathrm{FV}(\varphi \land \psi) = \mathrm{FV}(\varphi \lor \psi) = \mathrm{FV}(\varphi) \cup \mathrm{FV}(\psi)$.

(e) $\mathrm{FV}(\exists x\ \varphi) = \mathrm{FV}(\forall x\ \varphi) = \mathrm{FV}(\varphi) \setminus \{x\}$.

Example. $\mathrm{FV}(\exists y\ y^2 = x) = \{x\}$.

Example. $\mathrm{FV}(\forall x\ x \cdot x = e) = \varnothing$.

Definition 1.4.3 (Sentence)**.** A formula φ is called a *sentence* if it has no free variables.

Notation. If φ is a formula with free variables among $x_1, \ldots, x_n$ we often write $\varphi(x_1, \ldots, x_n)$.

Remark. So given $\varphi(x_1, \ldots, x_n)$, we know that φ has no other free variables than $x_1, \ldots, x_n$.

Example. It's valid to write $\varphi(x, y, z) \coloneqq x = y$.

1.5 Truths

Finally, we define the notion of truth.

Definition 1.5.1 (Truth)**.** Given an $\mathcal{L}$-structure $\mathcal{M}$, let $\varphi(x_1, \ldots, x_n)$ be an $\mathcal{L}$-formula and let $a_1, \ldots, a_n \in M$. Then we say φ is *true* of $\overline{a}$ in $\mathcal{M}$,[a] denoted as $\mathcal{M} \models \varphi(\overline{a})$, as follows:

(a) If φ is $s = t$, then $\mathcal{M} \models \varphi(\overline{a})$ if $s^{\mathcal{M}}(\overline{a}) = t^{\mathcal{M}}(\overline{a})$.

(b) If φ is $R(t_1, \ldots, t_{n_R})$, then $\mathcal{M} \models \varphi(\overline{a})$ if $\left(t_1^{\mathcal{M}}(\overline{a}), \ldots, t_{n_R}^{\mathcal{M}}(\overline{a})\right) \in R^{\mathcal{M}}$.

(c) If φ is $\neg\psi$, then $\mathcal{M} \models \varphi(\overline{a})$ if $\mathcal{M} \not\models \psi(\overline{a})$.

(d) If φ is $\psi_1 \land \psi_2$, then $\mathcal{M} \models \varphi(\overline{a})$ if $\mathcal{M} \models \psi_1(\overline{a})$ and $\mathcal{M} \models \psi_2(\overline{a})$.

(e) If φ is $\psi_1 \lor \psi_2$, then $\mathcal{M} \models \varphi(\overline{a})$ if $\mathcal{M} \models \psi_1(\overline{a})$ or $\mathcal{M} \models \psi_2(\overline{a})$.

(f) If φ is $\exists y\ \psi(\overline{x}, y)$, then $\mathcal{M} \models \varphi(\overline{a})$ if there's $b \in M$ such that $\mathcal{M} \models \psi(\overline{a}, b)$.

(g) If φ is $\forall y\ \psi(\overline{x}, y)$, then $\mathcal{M} \models \varphi(\overline{a})$ if for all $b \in M$ such that $\mathcal{M} \models \psi(\overline{a}, b)$.

[a] Or $\mathcal{M}$ satisfies $\varphi(\overline{a})$.

Remark. Every formula is true, or its negation is.

Lecture 3: Logical Consequence and Equivalence

1.5.1 Implications

12 Jan. 14:30

Notation (Material implication)**.** The *material implication* $\varphi \to \psi$ between two formulas φ, ψ is an abbreviation of $\neg\varphi \vee \psi$.

Notation. We use $\varphi \leftrightarrow \psi$ as an abbreviation of $((\varphi \to \psi) \wedge (\psi \to \varphi))$.

Essentially, $\to$ and $\leftrightarrow$ is different from $\Rightarrow$ and $\Leftrightarrow$, where the former are only shown in formula. Now, consider the language of graphs $\mathcal{L}_{\text{graph}} = \{E\}$, let's see some examples.

Example. An undirected graph can be written as

$$\forall x \forall y\ (xEy \to yEx).$$

Example. A vertex has at least three neighbors can be written as

$$\varphi(x) := \exists u \exists v \exists w\ (xEu \wedge xEv \wedge xEw \wedge u \neq v \wedge v \neq w \wedge u \neq w)$$

in non-reflexive graphs.

Example. For a vertex has exactly three neighbors,

$$\psi(x) := \exists u \exists v \exists w \forall y\ \big(xEu \wedge xEv \wedge xEw \wedge u \neq v \wedge v \neq w \wedge u \neq w \wedge (y = u \vee y = v \vee y = w \vee \neg yEx)\big).$$

Problem. Can we say that x has an even number of neighbors?

Answer. We can't. Some things are not expressible in FO logic. ⊛

Example. For a vertex x has a path of length 4 to y,

$$\Theta(x, y) := \exists u \exists v \exists w\ (xEu \wedge uEv \wedge vEw \wedge wEy).$$

We can also express that there is a path of length at most 4.

Problem. Can we say that there is a path from x to y?

Answer. We still can't! Not in FO logic (using compactness theorem). ⊛

Remark. When we prove results by induction on formulas, we only need to prove for $\neg, \wedge, \exists$, instead of for both $\wedge, \vee$, and both $\exists$ and $\forall$.

Proof. Since we can view $\varphi \vee \psi$ as an abbreviation for $\neg(\neg\varphi \wedge \neg\psi)$ and $\forall x\ \varphi$ as an abbreviation for $\neg(\exists x\ \neg\varphi)$. ⊛

Remark (Sheffer stroke)**.** In fact, we can get $\wedge, \vee, \neg$ from one logical connective, e.g., the *sheffer stroke* $\uparrow$, which is defined as

$$\varphi \uparrow \psi := \neg(\varphi \wedge \psi),$$

and we can use $\uparrow$ to define $\neg, \vee, \wedge$.

Notation. Let Φ be a (possibly infinite) set of sentences, we write $\mathcal{M} \models \Phi$ if $\mathcal{M} \models \varphi$ for all $\varphi \in \Phi$.

1.5.2 Logical Consequences and Equivalent

Definition 1.5.2 (Logical consequence). Let Φ be a set of sentences, and φ be a sentence. We say that φ is a *logical consequence* of Φ, written $\Phi \models \varphi$, if $\mathcal{M} \models \varphi$ whenever $\mathcal{M} \models \Phi$.

If $\Phi = \varnothing$ is the empty set, Definition 1.5.2 is written as $\models \varphi$, i.e., φ is true in all $\mathcal{L}$-structures.[2]

Definition 1.5.3 (Equivalent). Given two formulas φ, ψ, $\varphi(\overline{x})$ and $\psi(\overline{x})$ are *equivalent* if

$$\models \forall \overline{x} \; \big(\varphi(\overline{x}) \leftrightarrow \psi(\overline{x})\big).$$

Problem. Two sentences φ and ψ are equivalent if and only if $\varphi \models \psi$ and $\psi \models \varphi$.

DIY

As previously seen. $\mathcal{A}$ is a substructure of $\mathcal{B}$, or $\mathcal{A} \subseteq \mathcal{B}$, means that $A \subseteq B$ and $\mathrm{id}\colon A \hookrightarrow B$ is an $\mathcal{L}$-embedding.

Proposition 1.5.1. Suppose that $\mathcal{A}$ is a substructure of $\mathcal{B}$, and $\varphi(\overline{x})$ is a quantifier-free formula. Let $\overline{a} \in \mathcal{A}$,[a] then $\mathcal{A} \models \varphi(\overline{a})$ if and only if $\mathcal{B} \models \varphi(\overline{a})$.

[a] Formally, we need to write $\mathcal{A}$ to be the Cartesian product with a fixed length.

Proof. We start with terms by proving that if t is a term and $\overline{b} \in \mathcal{A}$, then $t^{\mathcal{A}}(\overline{b}) = t^{\mathcal{B}}(\overline{B})$. The proof is induction on terms.

(a) If t is a constant symbol c, then $t^{\mathcal{A}}(\overline{b}) = c^{\mathcal{A}} = c^{\mathcal{B}} = t^{\mathcal{B}}(\overline{b})$.

(b) If t is a variable x_i, then $t^{\mathcal{A}}(\overline{b}) = b_i = t^{\mathcal{B}}(\overline{b})$.

(c) If t is a function symbol $f(s_1, \dots, s_n)$ where s_i are terms, then $t^{\mathcal{A}}(\overline{b}) = f^{\mathcal{A}}\big(s_1^{\mathcal{A}}(\overline{b}), \dots, s_n^{\mathcal{A}}(\overline{b})\big)$. By the induction hypothesis, $s_i^{\mathcal{A}}(\overline{b}) = s_i^{\mathcal{B}}(\overline{b}) \in \mathcal{A}$, and hence

$$t^{\mathcal{B}}(\overline{b}) = f^{\mathcal{B}}\big(s_1^{\mathcal{B}}(\overline{b}), \dots, s_n^{\mathcal{B}}(\overline{b})\big) = f^{\mathcal{A}}\big(s_1^{\mathcal{A}}(\overline{b}), \dots, s_n^{\mathcal{A}}(\overline{b})\big) = t^{\mathcal{A}}(\mathcal{B}),$$

i.e., $f^{\mathcal{B}} \restriction_{\mathcal{A}} = f^{\mathcal{A}}$, so $t^{\mathcal{A}}(\overline{b}) = t^{\mathcal{B}}(\overline{b})$.

Now we turn to formulas, and prove that for φ quantifier-free, then $\mathcal{A} \models \varphi(\overline{a}) \Leftrightarrow \mathcal{B} \models \varphi(\overline{a})$ for $\overline{a} \in \mathcal{A}$. The proof is, again, induction on formulas.[a]

(a) If φ is $s = t$, then $s^{\mathcal{A}}(\overline{a}) = s^{\mathcal{B}}(\overline{a})$ and $t^{\mathcal{A}}(\overline{a}) = t^{\mathcal{B}}(\overline{a})$, so

$$\mathcal{A} \models \varphi(\overline{a}) \Leftrightarrow s^{\mathcal{A}}(\overline{a}) = t^{\mathcal{A}}(\overline{a}) \Leftrightarrow s^{\mathcal{B}}(\overline{a}) = t^{\mathcal{B}}(\overline{a}) \Leftrightarrow \mathcal{B} \models \varphi(\overline{a}).$$

(b) If φ is $R(s_1, \dots, s_n)$, then

$$A \models \varphi(\overline{a}) \Leftrightarrow \big(s_1^{\mathcal{A}}(\overline{a}), \dots, s_n^{\mathcal{A}}(\overline{a})\big) \in R^{\mathcal{A}} \Leftrightarrow \big(s_1^{\mathcal{B}}(\overline{a}), \dots, s_n^{\mathcal{B}}(\overline{a})\big) \in R^{\mathcal{B}} \Leftrightarrow \mathcal{B} \models \varphi(\overline{a}).$$

(c) If φ is $\neg\psi$,

$$\mathcal{A} \models \varphi(\overline{a}) \Leftrightarrow \mathcal{A} \not\models \psi(\overline{a}) \Leftrightarrow \mathcal{B} \not\models \psi(\overline{a}) \Leftrightarrow \mathcal{B} \models \varphi(\overline{a}),$$

where we use the induction hypothesis in the second $\Leftrightarrow$.

[2] Recall that we always have a language $\mathcal{L}$ implicitly.

(d) If φ is $\psi_1 \vee \psi_2$,

$$\mathcal{A} \models \varphi(\overline{a}) \Leftrightarrow \mathcal{A} \models \psi_1(\overline{a}) \text{ or } \mathcal{A} \models \psi_2(\overline{a}) \Leftrightarrow \mathcal{B} \models \psi_1(\overline{a}) \text{ or } \mathcal{B} \models \psi_2(\overline{a}) \Leftrightarrow \mathcal{B} \models \varphi(\overline{a}),$$

where we use the induction hypothesis in the second $\Leftrightarrow$.

■

[a] Recall that we only need to show one of $\vee$ or $\wedge$, and here we pick $\vee$ and treat $\wedge$ as an abbreviation.

As previously seen (Characteristic). Given a field K, the *characteristic* p of K is the number of 1 you need to add 1 in order to get 0, i.e., $\underbrace{1+1+\cdots+1}_{p} = 0$.

Example. Let L be a subfield of K, for each $p > 0$, $\varphi_p \coloneqq \underbrace{1+1+\cdots+1}_{p} = 0$, which says the characteristic p. φ_p is quantifier-free, so

$$L \models \varphi_p \Leftrightarrow K \models \varphi_p.$$

Example. Consider $\mathbb{Z} = (\mathbb{Z}, 0, 1, +, -, \cdot)$, and let $\varphi(x) \coloneqq \neg\exists y\ y + y = x$. We see that $\mathbb{Z} \models \varphi(1)$ but $\mathbb{Q} \models \neg\varphi(1)$.

Proposition 1.5.2. Suppose that $\mathcal{A}$ is a substructure of $\mathcal{B}$, and $\varphi(\overline{x}, y_1, \ldots, y_n)$ is a quantifier-free formula. Let $\overline{a} \in \mathcal{A}$, then

(a) if $\mathcal{A} \models \exists y_1 \ldots \exists y_n\ \varphi(\overline{a}, y_1, \ldots, y_n)$, then $\mathcal{B} \models \exists y_1 \ldots \exists y_n\ \varphi(\overline{a}, y_1, \ldots, y_n)$;

(b) if $\mathcal{B} \models \forall y_1 \ldots \forall y_n\ \varphi(\overline{a}, y_1, \ldots, y_n)$, then $\mathcal{A} \models \forall y_1 \ldots \forall y_n\ \varphi(\overline{a}, y_1, \ldots, y_n)$.

Proof. Suppose that $\mathcal{A} \models \exists y_1 \ldots \exists y_n\ \varphi(\overline{a}, y_1, \ldots, y_n)$, so there are $b_1, \ldots, b_n \in \mathcal{A}$ such that $\mathcal{A} \models \varphi(\overline{a}, b_1, \ldots, b_n)$. Since φ is quantifier-free, so $\mathcal{B} \models \varphi(\overline{a}, b_1, \ldots, b_n)$ from Proposition 1.5.1, and hence $\mathcal{B} \models \exists y_1 \ldots \exists y_n\ \varphi(\overline{a}, y_1, \ldots, y_n)$.

On the other hand, it's easy to see that (b) is implied by (a). ■

Notation (Existential). In Proposition 1.5.2, formulas as in (a) are called *existential* ($\exists_1$ *or* $\exists$) *formulas*.

Notation (Universal). In Proposition 1.5.2, formulas as in (b) are called *universal* ($\forall_1$ *or* $\forall$) *formulas*.

Example. Recall $\mathcal{L}_1 = \{e, \cdot\}$, $\mathcal{L}_2 = \left\{e, \cdot, ^{-1}\right\}$.

- Associativity: $\forall x \forall y \forall z\ (xy)z = x(yz)$.
- Identity: $\forall x\ ex = xe$.

These are $\forall$-formulas in either language.

- Inverses in $\mathcal{L}_1$: $\forall x \exists y\ xy = yx = e$, which is **not** an $\forall$-formula.
- Inverses in $\mathcal{L}_2$: $\forall x\ xx^{-1} = x^{-1}x = e$, which is an $\forall$-formula.

Hence, group axioms in $\mathcal{L}_1$ are not universal, but in $\mathcal{L}_2$ they are.

The above discrepancy is the reason why $\mathcal{L}_2$ is better than $\mathcal{L}_1$, i.e., $\mathcal{L}_1$-substructure might not be a group.

Problem. Show that $\forall x \exists y\ xy = yx = e$ in the above example is not equivalent to an $\forall$-formula.

Lecture 4: Theories and Axioms

17 Jan. 14:30

Example. Let $\mathcal{L}_1 = \{E\}$, where E is a binary relation representing edge relation; and $\mathcal{L}_2 = \{V, E, I\}$, where V, E are unary relations and I is a binary relation representing incidence such that $I(v, e)$ for $v \in V$, $e \in E$ means that v is a vertex on edge e. Then,

- Let G be a graph, viewed as an $\mathcal{L}_1$-structure. A substructure of G is an induced subgraph $H \subseteq G$ such that any edge in G between two vertices of H is in H.
- If we view G as an $\mathcal{L}_2$-substructure, a substructure is a subgraph H such that H has some vertices and edges from G.[a]

[a]But there might be edges in H with no vertices, which can be fixed by having two functions $I_1(e) = v$, $I_2(e) = w$ when $e\colon v \to w$.

Remark. The difference is that for $\mathcal{L}_1$, having an edge is quantifier-free, while in $\mathcal{L}_2$ is existential. To elaborate a bit further, for $\mathcal{L}_2$, vEw is quantifier-free, while in $\mathcal{L}_2$,

$$\exists\ (v \in V \wedge w \in V \wedge e \in E \wedge I(v, e) \wedge I(w, e))$$

is not quantifier-free.

Chapter 2

Soundness, Completeness, and Compactness

In this chapter, we're going to formalize proofs, including what do we mean by "having a proof" of a statement, and study properties of which.

2.1 Theories

Let's start by the notion of theory.

Definition 2.1.1 (Theory). An $\mathcal{L}$-*theory* is a set of $\mathcal{L}$-sentences.

Definition 2.1.2 (Model). $\mathcal{M}$ is a *model* of a theory T, written as $\mathcal{M} \models T$, if $\mathcal{M} \models \varphi$ for all $\varphi \in T$.

Note. Not every theory has a model, e.g., $\{\exists x\ x \neq x\}$.

The above note motivates the following.

Definition 2.1.3 (Satisfiable). A theory is *satisfiable* if it has a model.

Definition 2.1.4 (Elementary class). A class $\mathcal{K}$ of $\mathcal{L}$-structures $\mathcal{M}$ is called an *elementary class* if there is an $\mathcal{L}$-theory T such that

$$\mathcal{K} = \{\mathcal{M} \mid \mathcal{M} \models T\}.$$

One way to get an elementary class is to take an $\mathcal{L}$-structure $\mathcal{M}$ and take the full theory.

Definition 2.1.5 (Full theory). The *full theory* $\mathrm{Th}(\mathcal{M})$ of an $\mathcal{L}$-structure $\mathcal{M}$ is defined as $\mathrm{Th}(\mathcal{M}) = \{\varphi \mid \mathcal{M} \models \varphi\}$.

From the definition, $\mathcal{M} \models \mathrm{Th}(\mathcal{M})$, and $\mathrm{Th}(\mathcal{M})$ characterizes the structures satisfying the same sentences as $\mathcal{M}$.

Definition 2.1.6 (Complete). A theory T is *complete* if for any sentence φ, either $\varphi \in T$ or $\neg\varphi \in T$.

Remark. $\mathrm{Th}(\mathcal{M})$ is complete.

Definition 2.1.7 (Elementarily equivalent). $\mathcal{M}$ and $\mathcal{N}$ are *elementarily equivalent* $\mathcal{M} \equiv \mathcal{N}$ if for all sentences φ,

$$\mathcal{M} \models \varphi \Leftrightarrow \mathcal{N} \models \varphi.$$

Remark (Non-standard model of arithmetic). There are $\mathcal{N} \models \mathrm{Th}(\mathbb{N})$, but $\mathcal{N}$ is not isomorphic to $\mathbb{N}$. $\mathcal{N}$ is called a *non-standard model of arithmetic*, and $\mathcal{N}$ might have *infinite element* larger than all of $\mathcal{M}$. Here, $\mathbb{N} = (\mathbb{N}, 0, 1, +, \cdot, -)$

Example. $\mathbb{Z} \oplus \mathbb{Z} \not\equiv \mathbb{Z}$ as groups.

The other way to define a theory is to write down axioms.

Example (Infinite set). Let $\mathcal{L} = \varnothing$, and let T consist of

$$\varphi_n \coloneqq \exists x_1 \ldots \exists x_n \bigwedge_{i \neq j} x_i \neq x_j.$$

Example (Linear order). Let $\mathcal{L} = \{\leq\}$, and let T consist of the axioms of linear orders, e.g.,

$$\forall x \forall y\ (x \leq y \land y \leq x \to x = y).$$

There are other interesting theories of linear orders, e.g., dense ones.

Example (Dense linear order). Consider

$$\forall x \forall y\ (x < y \to \exists z\ x < z < y),$$

where we use $a < b$ as shorthand of saying $a \leq b \land a \neq b$.

Example (Group). In $\mathcal{L}_{\text{group}} = \{e, \cdot, ^{-1}\}$, let T be the group axioms.

Other theories of groups include Abelson group, divisible, etc.

Definition 2.1.8 (Finitely axiomatizable). A theory is *finitely axiomatizable* if it has a finite set of axioms.

Given a theory, consider $T^{\models} = \{\varphi \mid T \models \varphi\}$,[1] so $\mathcal{M} \models T$ if and only if $\mathcal{M} \models T^{\models}$. Often we think of T and $T^{\models}$ as the same. A theory T is finitely axiomatizable if there is a finite Φ such that $T^{\models} = \Phi^{\models}$.

2.2 Elementary Embeddings

Let's now consider the following notion.

Definition 2.2.1 (Elementary embedding). Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures, and $f \colon \mathcal{M} \to \mathcal{N}$ an $\mathcal{L}$-embedding. Then f is an *elementary embedding* if for any formula $\varphi(\overline{x})$ and $\overline{a} \in M$,

$$\mathcal{M} \models \varphi(\overline{a}) \Leftrightarrow \mathcal{N} \models \varphi(f(\overline{a})).$$

Definition 2.2.2 (Elementary substructure). If $f \colon \mathcal{M} \hookrightarrow \mathcal{N}$ is a elementary embedding where $\mathcal{M}$ is a substructure of $\mathcal{N}$, then $\mathcal{M}$ is an *elementary substructure* of $\mathcal{N}$, written as $\mathcal{M} \preceq \mathcal{N}$.

Example. As groups, $\mathbb{Z} \hookrightarrow \mathbb{Q}$ is not elementary. In fact, $\mathbb{Z} \not\equiv \mathbb{Q}$. Wheres, if $f \colon \mathcal{M} \hookrightarrow \mathcal{N}$ is an elementary embedding, $\mathcal{M} \equiv \mathcal{N}$.[a]

[a] And also much more is true.

[1] Recall Definition 1.5.2.

Proposition 2.2.1. Every isomorphism is an elementary embedding.

Proof. Let $f\colon \mathcal{M} \to \mathcal{N}$ be an isomorphism. We will argue by induction on formulas φ, that for all $\overline{a} \in M$,

$$\mathcal{M} \models \varphi(\overline{a}) \Leftrightarrow \mathcal{N} \models \varphi(f(\overline{a})).$$

Firstly, observe that all cases except quantifiers are the same as Proposition 1.5.1. For quantifiers, suppose that $\varphi(\overline{x})$ is $\exists y\ \psi(\overline{x}, y)$ and $\mathcal{M} \models \varphi(\overline{a})$. This means that there is $b \in M$ such that $\mathcal{M} \models \psi(\overline{a}, b)$. By the induction hypothesis, $\mathcal{N} \models \psi(f(\overline{a}), f(b))$, so $\mathcal{N} \models \varphi(f(\overline{a}))$.

Now suppose $\mathcal{N} \models \varphi(f(\overline{a}))$, then there is $c \in N$ such that $\mathcal{N} \models \psi(f(\overline{a}), c)$. Since f is an isomorphism, so there is a $b \in M$ such that $f(b) = c$. By the induction hypothesis, $\mathcal{M} \models \psi(\overline{a}, b)$, so $\mathcal{M} \models \varphi(\overline{a})$. ■

Corollary 2.2.1. If $\mathcal{M} \cong \mathcal{N}$, then $\mathcal{M} \equiv \mathcal{N}$.

2.3 Definable Sets

Consider the following.

Definition 2.3.1 (Definable)**.** Let $\mathcal{M}$ be an $\mathcal{L}$-structure, then $X \subseteq M^n$ is *definable* if there is a formula $\varphi(x_1, \ldots, x_n, \overline{y})$ and $\overline{b} \in M$ such that

$$X = \left\{\overline{a} \in M^n \mid \mathcal{M} \models \varphi(\overline{a}, \overline{b})\right\}.$$

Notation (Define)**.** We say that $\varphi(\overline{x}, \overline{b})$ *defines* X over $\overline{b}$, written as $X = \varphi(\mathcal{M}, \overline{b})$.

Notation (Parameter)**.** The tuple $\overline{b}$ is called the *parameters* when X is definable over $\overline{b}$.

Remark. Sometimes X is definable without parameters, or definable over $\varnothing$.

Example. Take $\mathbb{R} = (\mathbb{R}, 0, 1, +, \cdot, -)$ in $\mathcal{L}_{\text{ring}}$, then $\leq = \{(a, b) \colon a \leq b\}$ is definable.

Example. Let $\mathbb{Z} = (\mathbb{Z}, +, -, \cdot, 0, 1)$, then $\mathbb{N}$ is $\varnothing$-definable in $\mathbb{Z}$ by[a]

$$\mathbb{N} = \left\{z \in \mathbb{Z} \colon \exists u, v, x, y\ u^2 + v^2 + x^2 + y^2 = z\right\}.$$

[a] From the Langrange's four-square theorem, which says that every natural number is the sum of four squares.

Example. $\mathbb{Z}$ is $\varnothing$-definable in $\mathbb{Q} = (\mathbb{Q}, +, -, \cdot, 0, 1)$. This is a result of Julia Robinson [Rob49], and the formulation is very complicated.

Problem. How does one show that a set is not definable? For example, $\mathbb{R}$ is not definable in $\mathbb{C} = (\mathbb{C}, 0, 1, +, \cdot, -)$.

Lecture 5: Hilbert-Style Deductive System

We start by asking whether $\mathbb{R}$ is definable in $\mathbb{C} = (\mathbb{C}, 0, 1, +, \cdot, -)$? 19 Jan. 14:30

Proposition 2.3.1. Let $\mathcal{M}$ be an $\mathcal{L}$-structure, and let $X \subseteq M^n$ be a set which is definable over $\overline{a}$.

Then any automorphism of $\mathcal{M}$ that fixes $\bar{a}$ pointwise[a] fixes X setwise.[b]

[a] If $\bar{a} = (a_1, \ldots, a_m)$, then $f(a_i) = a_i$.
[b] If $b \in X$, then $f(b) \in X$.

Proof. Let f be an automorphism of $\mathcal{M}$ fixing $\bar{a}$ pointwise, and $X = \{\bar{b} \in M^n : \mathcal{M} \models \varphi(\bar{b}, \bar{a})\}$. Fix $\bar{b}$, and suppose $\bar{b} \in X$, so $\mathcal{M} \models \varphi(\bar{b}, \bar{a})$. Because f is an elementary embedding from Proposition 2.2.1,

$$\mathcal{M} \models \varphi(f(\bar{b}), f(\bar{a})) \Rightarrow \mathcal{M} \models \varphi(f(\bar{b}), \bar{a}),$$

hence $f(\bar{b}) \in X$. Similarly, if $\bar{b} \notin X$, $\mathcal{M} \models \neg\varphi(\bar{b}, \bar{a}) \Rightarrow \mathcal{M} \models \neg\varphi(f(\bar{b}, \bar{a}))$, so $f(\bar{b}) \notin X$. ■

Remark. If X is $\varnothing$-definable, it is fixed setwise by any automorphism.

Example. $\mathbb{N}$ is fixed setwise by any automorphism of the ring $\mathbb{Z}$. In fact, the only automorphism of $\mathbb{Z}$ is the identity.

Example. $\mathbb{N}$ is not $\varnothing$-definable in $\mathbb{Z} = (\mathbb{Z}, 0, +)$.

Proof. Consider an automorphism $f(x) = -x$ of the group $\mathbb{Z}$, which does not fix $\mathbb{N}$ setwise. ⊛

Problem. Is $\mathbb{N}$ definable in $\mathbb{Z} = (\mathbb{Z}, 0, +)$ over some parameters $\bar{a}$?

Answer. For example, if $\bar{a} = (1)$, then f does not fix 1. In fact, any automorphism fixing 1 also fixes all of $\mathbb{Z}$, but $\mathbb{N}$ is not definable in $(\mathbb{Z}, 0, +)$. To prove this we need compactness. ⊛

As previously seen. Given a field F, then $F(a) \cong F(b)$ if a and b have the same minimal polynomial over F or if both do not satisfy any polynomial over F.

Example. $\mathbb{Q}(\pi) \cong \mathbb{Q}(e)$ because π and e are both transcendental.

We now return to the big question: is $\mathbb{R}$ definable in $\mathbb{C} = (\mathbb{C}, 0, 1, +, \cdot, -)$? If $f \colon \mathbb{Q}(a) \to \mathbb{Q}(b)$ such that $a \mapsto b$, then there is an automorphism $\hat{f} \colon \mathbb{C} \to \mathbb{C}$ such that $a \mapsto b$, i.e., $\hat{f}$ extends f. In other words, we need to find such an f with $a \in \mathbb{R}$ and $b \notin \mathbb{R}$.

Example. $a = \pi$, $b = i\pi$ are both transcendental.

Example. a is a real $\sqrt[4]{2}$, b is a complex $\sqrt[4]{2}$.

The above two examples show that $\mathbb{R}$ is not $\varnothing$-definable in $\mathbb{C}$. In fact, $\mathbb{R}$ is not definable over any $\bar{a}$ because there are elements of $\mathbb{R}$ and $\mathbb{C} \setminus \mathbb{R}$ transcendental over any $\bar{a}$.

Intuition. There are so many a, b such that given any $\bar{a}$, we can still find a pair that works.

2.4 Proofs

There are all sorts of different proof systems, and the one we use is the so-called Hilbert-style deductive system. Before that, we first see some common notions.

Notation (Schema)**.** A *schema* is written in symbols for formulas, variables, etc.

Example. $\varphi \to (\psi \to \varphi)$ is a schema, i.e., an infinite set with all possible choices of φ and ψ.

Specifically, every logical axiom is written in schema, meaning that any instance of a symbol for a formula, e.g., φ, can be replaced by any formula.

Definition 2.4.1 (Generalization)**.** A formula φ is a *generalization* of a formula ψ if φ is $\forall x_1 \dots \forall x_n\ \psi$ where $x_1, \dots, x_n$ are variables.

Notation (Hypothesis)**.** *Hypotheses* are formulas that we may assume in a proof.

Definition 2.4.2 (Proof)**.** A *proof* is a sequence of formulas $\{\varphi_i\}_{i=1}^n$ such that φ_n is the conclusion, and each formula is either an axiom or is obtained from the previous formulas by a rule of inference.

Moreover, for a proof based on a set of hypotheses Γ, then in addition to a logical axiom, we can assert a formula $\varphi \in \Gamma$. If we prove ψ using Γ as hypotheses, we write $\Gamma \vdash \psi$.

Definition 2.4.3 (Valid)**.** If we prove ψ without hypotheses, we write $\vdash \psi$ and say ψ is *valid*.

Definition 2.4.4 (Logical axioms)**.** The *logical axioms* are the following formulas written in schema, as well as all of their generalizations:

Definition 2.4.5 (Propositional axioms)**.** The *propositional axioms* are

(A1) $\varphi \to (\psi \to \varphi)$.

(A2) $(\varphi \to (\psi \to \theta)) \to ((\varphi \to \psi) \to (\varphi \to \theta))$.

(A3) $(\neg\varphi \to \neg\psi) \to ((\neg\varphi \to \psi) \to \varphi)$.

(A4) $\forall x\ \varphi(x, \dots) \to \varphi(t, \dots)$ where t is any term.

(A5) $[\forall x\ (\varphi \to \psi)] \to [(\forall x\ \varphi) \to (\forall x\ \psi)]$.

(A6) $\varphi \to \forall x\ \varphi$, where x is not free in φ.

Definition 2.4.6 (Axioms for equality)**.** The *axioms for equality* is

(A7) for any terms $t, u, v, \dots$, function symbols f, and relation symbols R,

(a) $t = t$.
(b) $t = u \to u = t$.
(c) $(t = u \land u = v) \to (t = v)$.
(d) $(u_1 = t_1 \land \dots \land u_{n_f} = t_{n_f}) \to f(u_1, \dots, u_{n_f}) = f(t_1, \dots, t_{n_f})$.
(e) $(u_1 = t_1 \land \dots \land u_{n_R} = t_{n_R}) \to (R(u_1, \dots, u_{n_R}) \leftrightarrow R(t_1, \dots, t_{n_R}))$.

Definition 2.4.7 (Rule of inference)**.** From φ and $\varphi \to \psi$, deduces ψ.[a]

[a]This is called modus ponens.

These formulas might have free variables.

Example. A proof from calculus of a limit, e.g., $\forall \epsilon \exists \delta \dots$. And we start by stating

1. let $\epsilon > 0$,
2. choose $\delta = \epsilon$,

⋮

n. $|f(x) - f(y)| < \epsilon$.

We should interpret free variables as anything.

As previously seen (Propositional logic). $(p \wedge q) \vee (r \wedge \neg q)$.

Remark. We can check whether the propositional axioms are true with a truth table.

Definition 2.4.8 (Propositional tautology). A *propositional tautology* is a boolean combination $\vee, \wedge, \neg$ of formulas $\varphi_1, \ldots, \varphi_n$ which is true via a truth table assigning true or false to each of $\varphi_1, \ldots, \varphi_n$.

So instead of using propositional axioms, we could instead allow as logical axioms any propositional tautology. To prove completeness, we will need 5 propositional tautologies. We will prove some of these, but take others on faith.

Remark. Propositional axioms are enough to prove all propositional tautologies.

Notation. We write $\Gamma \vdash_{\mathcal{L}} \varphi$ if there is a proof of φ from Γ in the language $\mathcal{L}$.

Note. Passing to a larger language will not let you prove more, so we can just write $\vdash$.

Lecture 6: Soundness Theorem

To see why propositional axioms are enough to prove all propositional tautologies, we see one example. 24 Jan. 14:30

Problem. Prove $\varphi \to \varphi$ using propositional axioms.

Answer. We see that

1. $\varphi \to ((\psi \to \varphi) \to \varphi)$ from (A1), where ψ is any formula (possibly $\psi = \varphi$).
2. $[\varphi \to ((\psi \to \varphi) \to \varphi)] \to [(\varphi \to (\psi \to \varphi)) \to (\varphi \to \varphi)]$ from (A2).
3. $(\varphi \to (\psi \to \varphi)) \to (\varphi \to \varphi)$ from (MP) and the two above.
4. $\varphi \to (\psi \to \varphi)$ from (A1).
5. $\varphi \to \varphi$ from (MP) and the two above.

⊛

In general, we can prove

(a) $\varphi \to \varphi$;

(b) $\varphi \to \neg\neg\varphi$;

(c) $\neg\neg\varphi \to \varphi$;

(d) $(\varphi \to \psi) \to ((\neg\varphi \to \psi) \to \psi)$;

(e) $\varphi \to (\psi \to (\varphi \to \psi))$,

and so on.

Note. As we said, we may replace propositional axioms by every propositional tautologies.

Some proof system also have a second rule about universal quantifiers, but in our system, we have built this into the axioms. We can prove, as a theorem, what the other proof systems take as a rule.

Theorem 2.4.1. If $\Gamma \vdash \varphi$, and x does not occur freely in Γ, then $\Gamma \vdash \forall x\ \varphi$.

Proof. Fix Γ and x, we use *induction on proofs*. Consider the set $\{\varphi \mid \Gamma \vdash \forall x\ \varphi\}$, we will show that this set contains all the logical axioms, formulas from Γ, and is closed under modus ponens.[a]

(a) If φ is a logical axiom, so is its generalization $\forall x\ \varphi$, so $\Gamma \vdash \forall x\ \varphi$.

(b) If $\varphi \in \Gamma$, then x is not free in φ, so from (A6), $\varphi \to \forall x\ \varphi$, and from (MP), $\forall x\ \varphi$. The above are based on Γ, hence $\Gamma \vdash \forall x\ \varphi$.

(c) Suppose $\Gamma \vdash \forall x\ \varphi$ and $\Gamma \vdash \forall x\ (\varphi \to \psi)$, we want to show that $\Gamma \vdash \forall x\ \psi$.

1. By (A5), $\forall x\ (\varphi \to \psi) \to (\forall x\ \varphi \to \forall x\ \psi)$, Γ proves this.
2. By (MP), $\Gamma \vdash \forall x\ \varphi \to \forall x\ \psi$.
3. By (MP) again, $\Gamma \vdash \forall x\ \psi$.

■

[a] Thus, if $\Gamma \vdash \theta$, then $\theta \in \{\varphi \mid \Gamma \vdash \forall x\ \varphi\}$.

Corollary 2.4.1. If $\vdash \varphi$, then $\vdash \forall x\ \varphi$. So the generalization of anything valid is also valid.

We now ask a critical question: is our proof system a good one?

2.5 Soundness Theorem

The first thing we should check is whether our proofs are sound.

Definition 2.5.1 (Sound). A proof system is *sound* if any provable sentence φ is true.

The idea is that if an $\mathcal{L}$-sentence φ is provable, then it is true in all $\mathcal{L}$-structures, i.e., every thing we prove should be true, in other words, we can't prove wrong things.

Lemma 2.5.1 (Soundness). If Γ is a set of $\mathcal{L}$-sentences and φ is a sentence, and $\Gamma \vdash_{\mathcal{L}} \varphi$, then $\Gamma \models \varphi$.

Proof. Suppose that $\Gamma \vdash \varphi$, let $\psi_1, \psi_2, \ldots, \psi_n = \varphi$ be such a proof.[a] Let $\overline{x} = (x_1, \ldots, x_m)$ be the free variable that appears in the ψ_i. Let $\mathcal{M}$ be an $\mathcal{L}$-structure, $\mathcal{M} \models \Gamma$. To show $\mathcal{M} \models \varphi$, we show that by induction on i, for all $\overline{a} \in M^m$, $\mathcal{M} \models \psi_i(\overline{a})$. For ψ_i, we have three cases.

(a) If $\psi_i \in \Gamma$, then $\mathcal{M} \models \Gamma$ so $\mathcal{M} \models \psi_i$.

(b) If ψ_i is a (generalization of) a logical axiom, then we can check that $\mathcal{M} \models \psi_i(\overline{a})$. For example, if ψ_i is (A1), $\theta \to (\gamma \to \theta)$, it's easy to check that

$$\mathcal{M} \models \theta(\overline{a}) \to (\gamma(\overline{a}) \to \theta(\overline{a})).$$

(c) If there are $j, k < i$ such that ψ_k is $\psi_j \to \psi_i$, from inductive hypothesis, for all $\overline{a}$, $\mathcal{M} \models \psi_j(\overline{a}), \mathcal{M} \models \psi_k(\overline{a})$, then $\mathcal{M} \models \psi_j(\overline{a}) \to \psi_i(\overline{a})$. Checking our definition of truth, we get $\mathcal{M} \models \psi_i(\overline{a})$.

■

[a] Some ψ_i might be formulas, but φ should be a sentence.

There are remarks to make about some obvious properties of $\vdash_{\mathcal{L}}$.

Remark. If $\varphi \in \Gamma$, then $\Gamma \vdash \varphi$.

Remark. If $\Delta \subseteq \Gamma$, and $\Delta \vdash \varphi$, then $\Gamma \vdash \varphi$.

Remark. If $\Gamma \vdash_{\mathcal{L}} \varphi$, and $\mathcal{L}^+ \supseteq \mathcal{L}$, then $\Gamma \vdash_{\mathcal{L}^+} \varphi$.

Remark. If $\Gamma \vdash \varphi$, then there is a finite $\Delta \subseteq \Gamma$ such that $\Delta \vdash \varphi$.

We can prove the following.

Theorem 2.5.1 (Deduction theorem). For any set of formulas Γ, formulas θ and ψ,

$$\Gamma \cup \{\theta\} \vdash \psi \Leftrightarrow \Gamma \vdash \theta \to \psi.$$

Proof. The backward direction is easier. Suppose $\Gamma \vdash \theta \to \psi$, then $\Gamma \cup \{\theta\} \vdash \psi$ since we can have a proof like:

1. θ

   $\vdots$ (the proof of $\Gamma \vdash \theta \to \psi$)

n. $\theta \to \psi$

$n+1$. ψ.

Now, suppose that $\Gamma \cup \{\theta\} \vdash \psi$, then there is a proof $\psi_1, \dots, \psi_n = \psi$ from $\Gamma \cup \{\theta\}$. We argue inductively that $\Gamma \vdash \theta \to \psi_i$. For i, we have three cases.

(a) If $\psi_i \in \Gamma$ or it is a logical axiom. By (A1), $\psi_i \to (\theta \to \psi_i)$, so $\Gamma \vdash \theta \to \psi_i$.

(b) If $\psi_i = \theta$. Then $\Gamma \vdash \theta \to \theta$ by (A1) and (A2) from here, hence $\Gamma \vdash \theta \to \psi_i$.

(c) If ψ_i follows from ψ_j, $\psi_k = \psi_j \to \psi_i$, using (MP) with $j, k < i$.

   1. From the induction hypothesis, $\Gamma \vdash \theta \to \psi_j$ and $\Gamma \vdash \theta \to (\psi_j \to \psi_i)$.
   2. By (A2), $\Gamma \vdash [\theta \to (\psi_j \to \psi_i)] \to [(\theta \to \psi_j) \to (\theta \to \psi_i)]$.
   3. By (MP), $\Gamma \vdash (\theta \to \psi_j) \to (\theta \to \psi_i)$.
   4. By (MP), $\Gamma \vdash \theta \to \psi_i$.

■

Lecture 7: Soundness, Completeness, and Compactness

26 Jan. 14:30

Proposition 2.5.1 (Contraposition). If $\Gamma \cup \{\varphi\} \vdash \neg\psi$, then $\Gamma \cup \{\psi\} \vdash \neg\varphi$.

Proof. Suppose $\Gamma \cup \{\varphi\} \vdash \neg\psi$, by the deduction theorem says that $\Gamma \vdash \varphi \to \neg\psi$. From (A1), (A2), and (A3), we can prove $(\varphi \to \neg\psi) \to (\psi \to \neg\varphi)$. By (MP), $\Gamma \vdash \psi \to \neg\varphi$, then from the deduction theorem, $\Gamma \cup \{\psi\} \vdash \neg\varphi$. ■

Now we introduce an important notion.

Definition 2.5.2 (Consistent). A theory T is *consistent* if for all φ, it is not the case that $T \vdash \varphi$ and $T \vdash \neg\varphi$.

Definition 2.5.3 (Inconsistent). If a theory T is not consistent, then it's *inconsistent*.

We could make the same definition for a set of formulas.

Proposition 2.5.2 (Proof by contradiction). If $\Gamma \cup \{\varphi\}$ is inconsistent, then $\Gamma \vdash \neg\varphi$.

Proof. There is ψ such that $\Gamma \cup \{\varphi\} \vdash \psi$ and $\Gamma \cup \{\varphi\} \vdash \psi$, so $\Gamma \vdash \varphi \to \psi$ and $\Gamma \vdash \varphi \to \neg\psi$ by the deduction theorem. Using (A1), (A2), and (A3), we can prove that

$$(\varphi \to \psi) \to ((\varphi \to \neg\psi) \to \neg\varphi).$$

By (MP), $\Gamma \vdash (\varphi \to \neg\psi) \to \neg\varphi$, and by (MP) again, we have $\Gamma \vdash \neg\varphi$. ■

Proposition 2.5.3. If a theory T is consistent, and φ is a sentence, then either $T \cup \{\varphi\}$ or $T \cup \{\neg\varphi\}$ is consistent.

Proof. If they were both inconsistent, $T \vdash \neg\varphi$ and $T \vdash \neg\neg\varphi$, so T would be inconsistent $\lightning$ ■

Note. The above is also true for formulas.

Remark. If T is inconsistent, then $T \vdash \varphi$ for any φ.

Proof. If T is inconsistent, then $T \cup \{\neg\varphi\}$ is inconsistent for all φ. Hence, from proof by contradiction, $T \vdash \neg\neg\varphi$ for all φ, which is just $T \vdash \varphi$. ⊛

Definition 2.5.4 (Maximal). A theory T is *maximal* if it is consistent and for all sentences φ, either $\varphi \in T$ or $\neg\varphi \in T$.

In particular, if $T \vdash \varphi$, then $\varphi \in T$.

Intuition. Basically, a maximal consistent theory has opinion on everything.

Now, we want to see that given a consistent theory, whether we can extend it to a maximal one. To do this, we need the following.

Definition. Let $(P, \leq)$ be a partially ordered set.

Definition 2.5.5 (Chain). A *chain* is a set $C \subseteq P$ such that for every $p, q \in C$, either $p \leq q$ or $q \leq p$.

Definition 2.5.6 (Upper bound). If $X \subseteq P$ is a set, an *upper bound* for X is an element $p \in P$ such that $p \geq q$ for all $q \in X$.

Definition 2.5.7 (Maximal). An element $p \in P$ is *maximal* if there is no $q \in P$ with $q > p$.

Note. Note that a maximal element might not be greater than everything, there is just nothing greater than it.

Theorem 2.5.2 (Zorn's lemma). Let $(P, \leq)$ be a partially ordered set. If every non-empty chain in P has an upper bound, then P has a maximal element.

Theorem 2.5.3. Any consistent theory T can be extended to a maximal consistent theory $T' \supseteq T$.

Proof. We first consider the case that T is countable by considering $\mathcal{L}$ is countable since if $\mathcal{L}$ is countable, then there are only countable many formulas since there are only countable many formulas of each length.

Claim. The result holds for $\mathcal{L}$ being countable.

Proof. Firstly, list out all sentences $\varphi_1, \varphi_2, \ldots$, start with $T_0 = T$. Given T_i consistent, one of $T_i \cup \{\varphi_i\}$ or $T_i \cup \{\neg\varphi_i\}$ is consistent from Proposition 2.5.3. Let T_{i+1} be one of these that is consistent. Let $T^* = \bigcup_i T_i$, which is maximal, and we now show that T^* is consistent.

Suppose not, then $T^* \vdash \theta$ and $T^* \vdash \neg\theta$ for some θ. In this case, there is some T_i such that $T_i \vdash \theta$ and $T_i \vdash \neg\theta$ because proofs are finite, with T_i being consistent, a contradiction ↯ ⊛

Claim. The result holds for arbitrary $\mathcal{L}$.

Proof. For arbitrary $\mathcal{L}$, let $(P, \leq)$ be the set of consistent theories extending T_i ordered by inclusion. Let C be a non-empty chain, and let $T^* = \bigcup_{T' \in C} T' \supseteq T$.

We see that T^* is consistent because if $T^* \vdash \theta$ and $T^* \vdash \neg\theta$, there are finitely many formulas used in those proofs, from, say, $T_1, \ldots, T_n \in C$. Because C is a chain, by reordering, we may assume that $T_1 \subseteq \cdots \subseteq T_n$. So $T_n \vdash \theta$ and $T_n \vdash \neg\theta$, contradicting the consistency of T_n, so T^* is consistent, i.e., $T^* \in P$. Furthermore, T^* is an upper bound on C,[a] so $(P, \leq)$ has a maximal consistent theory $T^* \supseteq T$ from Zorn's lemma.

If T^* is not maximal, then there is φ where $\varphi \notin T^*$, $\neg\varphi \notin T^*$. From Proposition 2.5.3, one of $T^* \cup \{\varphi\}$ or $T^* \cup \{\neg\varphi\}$ is consistent, hence in P, contradicting to T^* being maximal ↯ ⊛

[a] Note that C is arbitrary.

■

Remark. We can do that same proof for any $\mathcal{L}$ using transfinite recursion for the uncountable case.

Motivated by Lemma 2.5.1 and Theorem 2.5.3, we close this section with the following.

Theorem 2.5.4 (Soundness). Let T be a theory and φ be a sentence.

(a) If $T \vdash \varphi$, then $T \models \varphi$.

(b) If T is satisfiable, then it is consistent.

Proof. (a) is exactly Theorem 2.5.4. For (b), let $\mathcal{M} \models T$, suppose that T was inconsistent, then $T \vdash \varphi$ and $T \vdash \neg\varphi$ for some φ. By (a), $T \models \varphi$ and $T \models \neg\varphi$, so $\mathcal{M} \models \varphi$ and $\mathcal{M} \models \neg\varphi$. But $\mathcal{M} \models \neg\varphi$ means $\mathcal{M} \not\models \varphi$, so this is a contradiction, hence T is consistent. ■

2.6 Completeness and Compactness Theorems

After knowing our proof system is sound, we now ask the converse: is our proof system complete?

Definition 2.6.1 (Complete). A proof system is *complete* if any true sentence φ is provable.

And indeed, this is the case.

Theorem 2.6.1 (Completeness). Let T be a theory and φ be a sentence.

(a) If $T \models \varphi$, then $T \vdash \varphi$.

(b) If T is consistent, then it is satisfiable.

(b) implies (a) is easy to see. Suppose that $T \models \varphi$, so $T \cup \{\neg\varphi\}$ is not satisfiable. By (b), $T \cup \{\neg\varphi\}$ is inconsistent. By proof by contradiction, $T \vdash \varphi$. One important consequence of the completeness theorem is the compactness theorem.

Theorem 2.6.2 (Compactness). Let T be a theory and φ be a sentence.

(a) If $T \models \varphi$, then there is a finite $T_0 \subseteq T$ such that $T_0 \models \varphi$.

(b) T is satisfiable if and only if every finite subset of T is satisfiable.

Proof. Consider the following.

(a*) If $T \vdash \varphi$, then there is a finite $T_0 \subseteq T$ such that $T_0 \vdash \varphi$.

(b*) If T is consistent if and only if every finite subset of T is consistent.

We see that (a*) and (b*) are true because proofs are finite, and soundness and completeness translate directly between (a) and (a*) (and (b) and (b*)). ∎

Remark. The compactness theorem does have something to do with topological compactness; consider the topological space of complete satisfiable theories, with the basic open sets being the sets

$$U_\varphi \coloneqq \{T \colon T \models \varphi\},$$

then this topological space is compact.

Let's see one cool example using compactness.

Example (Construction of non-standard model of arithmetic). Let $\mathcal{L} = \{0, 1, +, \cdot, -, <\}$, and $\mathcal{L}^* = \mathcal{L} \cup \{c\}$, where c is a new constant symbol. Then

$$T = \mathrm{Th}_{\mathcal{L}}(\mathbb{N}) \cup \{c > \underline{n} \mid n \in \mathbb{N}\},$$

is finitely satisfiable.

Proof. Given $T_0 \subseteq T$ finite, $T_0 \subseteq \mathrm{Th}_{\mathcal{L}}(\mathbb{N}) \cup \{c > \underline{n}, \ldots, c > \underline{n_\ell}\}$, and may assume they are equal and show that T_0 is satisfiable. Let $\mathcal{N}$ be the $\mathcal{L} \cup \{c\}$-structure which is the expansion of the $\mathcal{L}$-structure $\mathbb{N}$, with

$$c^{\mathcal{N}} = 1 + \max(n_1, \ldots, n_\ell),$$

then $\mathcal{N} \models T_0$, and T_0 is satisfiable. By compactness, T is satisfiable, say $\mathcal{A} \models T$. Then $\mathcal{A} \equiv \mathbb{N}$ and $\mathcal{A}$ contains an element $c^{\mathcal{A}}$ bigger than 1, $1+1$, $1+1+1$, ..., but $\mathcal{A} \not\cong \mathbb{N}$, so $\mathcal{A}$ is a non-standard model of arithmetic. ⊛

We now start a long journey toward proving completeness theorem, specifically (b).

Lecture 8: Henkin Constants

2.6.1 Henkin Construction

31 Jan. 14:30

To prove Theorem 2.6.1 (b), we need an additional definition and a technical lemma due to Henkin.

Definition 2.6.2 (Henkin constant). An $\mathcal{L}^*$-theory T^* has *Henkin constants* if for each formula $\varphi(x)$ with one free variable, there is a constant symbol $c \in \mathcal{L}^*$ such that

$$\big(\exists x\ \varphi(x)\big) \to \varphi(c) \text{ is in } T^*.$$

We see that the above is equivalent to

$$\big(\neg\forall x\ \varphi(x)\big) \to \neg\varphi(c) \text{ is in } T^*,$$

and we will use this version ($\forall$) and view $\exists$ being a shorthand for $\neg\forall\neg$; also, we will use $\to$ and $\neg$ as primitive, and $\wedge, \vee$ are shorthand.

Lemma 2.6.1. If $\Gamma \vdash \varphi(c)$, and c does not occur in Γ or in $\varphi(x)$, then there is a variable y not appearing in $\varphi(x)$, such that $\Gamma \vdash \forall y\ \varphi(y)$. Moreover, there is a proof of $\forall y\ \varphi(y)$ in which c does not appear.

Proof. Let $\alpha_1(c), \ldots, \alpha_n(c) = \varphi(c)$ be a proof of $\varphi(c)$ from Γ, and let y be a variable not appearing in this proof. We claim that $\alpha_1(y), \ldots, \alpha_n(y) = \varphi(y)$ is still a valid proof of $\varphi(y)$. There are three cases to consider (for each $i = 1, \ldots, n$):

(a) If $\alpha_i(c)$ is in Γ, then c does not actually occur in $\alpha_i(c)$ because it does not appear in Γ. So $\alpha_i(y)$ is the same as $\alpha_i(c)$, hence in Γ.

(b) If $\alpha_i(c)$ is a logical axiom, then $\alpha_i(y)$ is a logical axiom as well. For most of these it is easy to check, but for (A6), i.e., $\varphi \to \forall x\ \varphi$ if x is not free in φ, there is a little more. But y did not appear in $\alpha_i(c)$, so $y \neq x$, and substituting y for c will not stop x from being not free.

(c) If $\alpha_i(c)$ follows by (MP) from $\alpha_j(c)$ and $\alpha_k(c) = \alpha_j(c) \to \alpha_i(c)$ for $j, k < i$, then $\alpha_i(y)$ follows by (MP) from $\alpha_j(y)$ and $\alpha_k(y) = \alpha_j(y) \to \alpha_i(y)$.

So $\Gamma \vdash \varphi(y)$ and the proof does not involve c. Let $\Phi \subseteq \Gamma$ be the subset of Γ that was used in the proof, so y does not appear in Φ, hence $\Phi \vdash \varphi(y)$ and $\Phi \vdash \forall y\ \varphi(y)$, so $\Gamma \vdash \forall y\ \varphi(y)$. ■

So Lemma 2.6.1 implies that we have $\Gamma \vdash \varphi(y)$ and the proof does not involve c. And sometimes, we want to be able to choose the variable y from above.

Corollary 2.6.1. If $\Gamma \vdash \varphi(c)$, and c does not occur in Γ or in $\varphi(x)$, then $\Gamma \vdash \forall x\ \varphi(x)$. Moreover, there is a proof of $\forall x\ \varphi(x)$ not involving c.[a]

[a] Here, x is any variable that does not appear in $\varphi(c)$.

Proof. We know that for some y, $\Gamma \vdash \forall y\ \varphi(y)$, (A4) says $\forall y\ \varphi(y) \to \varphi(x)$. So $\forall y\ \varphi(y) \vdash \varphi(x)$ since x does not appear in $\forall y\ \varphi(y)$, $\forall y\ \varphi(y) \vdash \forall x\ \varphi(x)$. ■

Note. x might appear in Γ.

Theorem 2.6.3. Let T be a consistent $\mathcal{L}$-theory. There is a language $\mathcal{L}^* \supseteq \mathcal{L}$ and $T^* \supseteq T$ a consistent $\mathcal{L}^*$-theory such that T^* has Henkin constants. We can choose $\mathcal{L}^*$ such that $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$, and all new symbols in $\mathcal{L}^*$ are constants.

Proof. Let $\mathcal{L}_0 = \mathcal{L}$ and $T_0 = T$. Let $\mathcal{L}_1$ be the expansion of $\mathcal{L}_0$ by adding a new constant symbol c_φ for each $\mathcal{L}_0$-formula φ w.r.t. the Henkin construction. First, we show that after this procedure, T_0 is still a consistent $\mathcal{L}_1$-theory.

Remark. Technically, $\vdash$ is really $\vdash_{\mathcal{L}}$, so this is a key step for seeing that it does not matter.

Claim. T_0 is still a consistent $\mathcal{L}_1$-theory after the expansion of $\mathcal{L}_0$.

Proof. If not, there is a proof of a contradiction from T_0, and which uses only finitely many of the new constants symbols. By Corollary 2.6.1, we can replace these constants one-by-one by variables, e.g., if the original contradiction was $\varphi(c_1, \ldots, c_n)$ and $\neg\varphi(c_1, \ldots, c_n)$, then T_0 proves $\forall x_1, \ldots, \forall x_n\ \varphi(x_1, \ldots, x_n)$ and $\forall x_1, \ldots, \forall x_n\ \neg\varphi(x_1, \ldots, x_n)$. Moreover, these proofs take place in $\mathcal{L}_0$, so by (A4), $T_0 \vdash_{\mathcal{L}_0} \varphi(x_1, \ldots, x_n)$, and $T_0 \vdash_{\mathcal{L}_0} \neg\varphi(x_1, \ldots, x_n)$ $\lightning$ ⊛

To construct T_1 w.r.t. the Henkin construction, it's natural to consider the following: if φ is of the form $\neg\forall x\ \psi(x)$, then let

$$\theta_\varphi \coloneqq (\neg\forall x\ \psi(x)) \to \neg\psi(c_\varphi),\ \text{ i.e., } (\exists x\ \neg\psi(x)) \to \neg\psi(c_\varphi),$$

otherwise, let $\theta_\varphi \coloneqq \forall x\ (x = x)$ (trivially true). Let $\Theta = \{\theta_\varphi \mid \varphi \text{ an } \mathcal{L}_0\text{-formula}\}$, and we let that $T_1 = T_0 \cup \Theta$. We claim that T_1 is still consistent.

Claim. $T_1 = T_0 \cup \Theta$ is a consistent $\mathcal{L}_1$-language after the expansion of $\mathcal{L}_0$.

Proof. If not, then there are $\varphi_1, \ldots, \varphi_{m+1}$ such that $T_0 \cup \{\theta_{\varphi_1}, \ldots, \theta_{\varphi_m}, \theta_{\varphi_{m+1}}\}$ is inconsistent. Taking m to be as small as possible, $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m$ is consistent, so $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m \vdash \neg\theta_{\varphi_{m+1}}$ with φ_{m+1} being of the form $\neg\forall x\ \psi(x)$, $\theta_{\varphi_{m+1}}$ is $\neg\forall x\ \psi(x) \to \neg\psi(c_\varphi)$. By (A1), (A2), (A3),

$$T_0 \cup \{\theta_{\varphi_1}, \ldots, \theta_{\varphi_m}\} \vdash \neg\forall x\ \psi(x) \text{ and } T_0 \cup \{\theta_{\varphi_1}, \ldots, \theta_{\varphi_m}\} \vdash \psi(c_{\varphi_{m+1}}).$$

Since $c_{\varphi_{m+1}}$ does not appear in $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m$, so $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m \vdash \forall x\ \psi(x)$, i.e., $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m$ is inconsistent, contradicting to the fact that m is the smallest choice $\lightning$ [a] ⊛

[a] If $m = 0$, then we violate the consistency of T_0.

It might be that T_1 does not have Henkin constants since there are new $\mathcal{L}_1$-formulas which are not $\mathcal{L}_0$-formulas. But we know that T_1 does have Henkin constants for $\mathcal{L}_0$-formulas, hence we can repeat that process and keep fixing things. In general, given T_i and $\mathcal{L}_i$, define a T_{i+1} and $\mathcal{L}_{i+1}$ in the above way. Since each T_i is consistent, so $T^* = \bigcup T_i$ is an $\mathcal{L}^* = \bigcup \mathcal{L}_i$-theory. Note that T^* is consistent as a nested union of consistent theories, and T^* has Henkin constants because every $\mathcal{L}^*$-formula φ is an $\mathcal{L}_i$-formula for some i, and $\theta_\varphi \in T_{i+1} \subseteq T^*$.

Intuition. This is like "chasing its own tail," which basically fixes new errors introduced every time and then takes the union in the end.

Finally, we want to show that $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$. Given $\mathcal{L}_i$, we define $\mathcal{L}_{i+1}$ to be $\mathcal{L}_i$ plus new constants c_φ for φ on $\mathcal{L}_i$-formula. Then, we have

$$|\mathcal{L}_{i+1}| \leq |\mathcal{L}_i| + \underbrace{|\mathcal{L}_i| + \aleph_0}_{\#\text{ of } \mathcal{L}_i\text{-formulas}} = |\mathcal{L}_i| + \aleph_0.$$

So for all i, $|\mathcal{L}_i| \leq |\mathcal{L}| + \aleph_0$, and $\mathcal{L}^* = \bigcup_i \mathcal{L}_i$ is a countable union, so $|\mathcal{L}^*| \leq |\mathcal{L}| + \aleph_0$, and in fact, $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$. ■

2.6.2 Proof of Completeness Theorem

After proving Theorem 2.6.3, we see that to prove Theorem 2.6.1 (b), we can proceed by:

1. extend T^* to a maximal theory T^{**};[2]
2. turn T^{**} into a model. The elements of the model are constant symbols from $\mathcal{L}^*$, modulo the equivalence relation $c \sim d$ if $c = d$ is in T^{**}, i.e., $T^{**} \vdash c = d$.

Thankfully, the first step is easy from Theorem 2.5.3, so we just need to show the second step, and we're done.

Lecture 9: Proving the Completeness Theorem

To finish the proof of Theorem 2.6.1 (b), we follow the plan mentioned last lecture, and prove the following. 2 Feb. 14:30

Theorem 2.6.4. If T is a maximal consistent $\mathcal{L}$-theory with Henkin constants, then T has a model.

Proof. The model we build is called a "canonical model." Let $\mathcal{C}$ be the set of constants in $\mathcal{L}$, and define an equivalence relation $\sim$ on $\mathcal{C}$ by $c \sim d$ if and only if $c = d$ is in T.

Claim. The relation $\sim$ on $\mathcal{C}$ defined by $c \sim d \Leftrightarrow c = d \in T$ is an equivalence relation.

[2] Which still has Henkin constants.

Proof. We check the axioms for being an equivalence relation.

(a) $c \sim c$ because $c = c$ is in T by (A7) (a).[a]

(b) If $c \sim d$, then $c = d$ is in T so $d = c$ is in T by (A7) (b), i.e., $d \sim c$.

(c) If $c \sim d$ and $d \sim e$, then $c = d$ and $d = e \in T \Rightarrow c = e \in T$ by (A7) (c), so $c \sim e$.

⊛

[a]Otherwise, $c \neq c$ is in T from the maximality, so $T \vdash c \neq c$ with $T \vdash c = c$, so T would be inconsistent.

Let $[c]$ be the equivalence class of c. Define an $\mathcal{L}$-structure $\mathcal{M}$ with domain $M = \mathcal{C} / \sim = \{[c] \mid c \in \mathcal{C}\}$, with functions, relations, and constants defined as follows:

(a) $c^{\mathcal{M}} = [c]$.

(b) $R^{\mathcal{M}}([c_1], \ldots, [c_n])$ true if $R(c_1, \ldots, c_n)$ is in T. This is well-defined by (A7) (e).

(c) $f^{\mathcal{M}}([c_1], \ldots, [c_n]) = [d]$ if $f(c_1, \ldots, c_n) = d$ is in T. Such a d exists because $\exists x\ f(c_1, \ldots, c_n) = x$, i.e., $\neg \forall x\ f(c_1, \ldots, c_n) \neq x$, is in T.[b] If this is in T, then there is a Henkin constant d with $f(c_1, \ldots, c_n) = d$ in T. To show that this is well-defined, from (A7) (d), i.e.,

$$(t_1 = u_1 \wedge \cdots \wedge t_n = u_n) \rightarrow f(t_1, \ldots, t_n) = f(u_1, \ldots, u_n).$$

So if $[c_1] = [d_1], \ldots, [c_n] = [d_n]$, then $c_1 = d_1, \ldots, c_n = d_n$ are in T. So $f(c_1, \ldots, c_n) = f(d_1, \ldots, d_n)$ is in T by (A7) (d). If a and b are constants such that $f(c_1, \ldots, c_n) = a$ and $f(d_1, \ldots, d_n) = b$ are in T, so $a = b$ is in T by (A7) (c), i.e., the transitivity of $=$.

Now we need to show that $\mathcal{M} \models T$, i.e., we claim that

$$\mathcal{M} \models \varphi([c_1], \ldots, [c_n]) \Leftrightarrow \varphi(c_1, \ldots, c_n) \text{ is in } T.$$

We prove this by induction on terms and then formulas.

1. Terms: Show that $t^{\mathcal{M}}([c_1], \ldots, [c_n]) = [d]$ if and only if $t(c_1, \ldots, c_n) = d$ is in T.

   (a) If t is a constant e, $t^{\mathcal{M}}([c_1], \ldots, c_n) = e^{\mathcal{M}} = [e]$, and

   $$[e] = t^{\mathcal{M}}([c_1], \ldots, [c_n]) = [d] \Leftrightarrow [e] = [d] \Leftrightarrow e = d \text{ is in } T.$$

   (b) If t is x_i, $t^{\mathcal{M}}([c_1], \ldots [c_n]) = [c_i]$. This is equal to $[d]$ if and only if $c_i = d$ is in T.

   (c) Suppose that $t(x_1, \ldots, x_n) = f(s_i(x_1, \ldots, x_n), \ldots, s_m(x_1, \ldots, x_n))$. Let

   $$[d_i] = s_i^{\mathcal{M}}([c_1], \ldots, [c_n]),$$

   by the inductive hypothesis, $d_i = s_i(c_1, \ldots, c_n)$ is in T. Let $[e] = f^{\mathcal{M}}([d_1], \ldots, [d_m]) = t^{\mathcal{M}}([c_1], \ldots, [c_n])$. By the definition of f, $e = f(d_1, \ldots, d_m)$ is in T. By (A7) (d),

   $$e = f(s_1(c_1, \ldots, c_n), \ldots, s_m(c_1, \ldots, c_n))$$

   is in T. This is the direction $(\Rightarrow)$.
   Now suppose that $t(c_1, \ldots, c_n) = e'$ is in T. We want to show that $[e] = [e']$, i.e., $e = e'$ is in T. Since $e = t(c_1, \ldots, c_n)$ is in T, and $e' = t(c_1, \ldots, c_n)$ is in T. By (A7) (c), $e = e'$ is in T, so $[e'] = [e] \ = t^{\mathcal{M}}([c_1], \ldots, [c_n])$. This is $(\Leftarrow)$.

2. Formulas: Show that $\mathcal{M} \models \varphi([c_1], \ldots, [c_n])$ if and only if $\varphi(c_1, \ldots, c_n)$ is in T.[c]

   (a) If φ is $s(x_1, \ldots, x_n) = t(x_1, \ldots, x_n)$:

($\Rightarrow$) If $\mathcal{M} \models s([c_1], \ldots, [c_n]) = t([c_1], \ldots, [c_n])$,

$$s^{\mathcal{M}}([c_1], \ldots, [c_n]) = t^{\mathcal{M}}([c_1], \ldots, [c_n]).$$

Let $[d]$ be this element equal to the above, so $d = s(c_1, \ldots, c_n)$ and $d = t(c_1, \ldots, c_n)$ are in T so $\underbrace{s(c_1, \ldots, c_n) = t(c_1, \ldots, c_n)}_{\varphi(c_1, \ldots, c_n)}$ is in T by (A7) (c).

($\Leftarrow$) If $s(c_1, \ldots, c_n) = t(c_1, \ldots, c_n)$ is in T, let

$$[d] = s^{\mathcal{M}}([c_1], \ldots, [c_n]) \text{ and } [e] = t^{\mathcal{M}}([c_1], \ldots, [c_n]),$$

so $d = s(c_1, \ldots, c_n)$ and $e = t(c_1, \ldots, c_n)$ are in t, so $d = e$ is in t, and $[e] = [d]$.

(b) If φ is $R(t_1(\overline{x}), \ldots, t_m(\overline{x}))$: Let $[d_i] = t_i^{\mathcal{M}}([c_1], \ldots, [c_n])$,

$$\begin{array}{ccc}
R^{\mathcal{M}}([d_1], \ldots, [d_m]) \text{ is true} & \Longleftrightarrow & R(d_1, \ldots, d_m) \text{ is in } T \\
\Updownarrow & & \Updownarrow \\
R^{\mathcal{M}}(t_1^{\mathcal{M}}[\overline{c}], \ldots, t_m^{\mathcal{M}}[\overline{c}]) \text{ is true} & & R(t_1(\overline{c}), \ldots, t_m(\overline{c})) \text{ is in } T \\
\Updownarrow & & \\
\mathcal{M} \models \varphi([c_1], \ldots, [c_n]) & &
\end{array}$$

(c) If φ is $\neg\psi$: Then

$$\mathcal{M} \models \varphi(\overline{c}) \Leftrightarrow \mathcal{M} \not\models \psi([\overline{c}]) \Leftrightarrow \psi(\overline{c}) \text{ is not in } T \Leftrightarrow \varphi(\overline{c}) \text{ is in } T$$

where the last $\Leftrightarrow$ follows from the fact that T is maximal and consistent.

(d) If φ is $\psi \to \theta$:

- If $\psi(\overline{c}) \to \theta(\overline{c})$ is in T: then if $\psi(\overline{c})$ is in T, then $\theta(\overline{c})$ is in T by (MP).then by the induction hypotheses, if $\mathcal{M} \models \psi([\overline{c}])$, then $\mathcal{M} \models \theta([\overline{c}])$.
- If $\mathcal{M} \models \psi([\overline{c}]) \to \theta([\overline{c}])$: then either $\mathcal{M} \models \theta([\overline{c}])$ or $\mathcal{M} \models \neg\psi([\overline{c}])$. So either
  i. $\theta(\overline{c})$ is in T: by (A1), $\theta(\overline{c}) \to (\psi(\overline{c}) \to \theta(\overline{c}))$, so $\psi(\overline{c}) \to \theta(\overline{c})$ is in T.
  ii. $\neg\psi(\overline{c})$ is in T: $T \cup \{\psi(\overline{c})\}$ is now inconsistent, so $T \cup \{\psi(\overline{c})\} \vdash \theta(\overline{c})$. From the deductive theorem, $T \vdash \psi(\overline{c}) \to \theta(\overline{c})$. Because T is maximal and consistent, $\psi(\overline{c}) \to \theta(\overline{c})$ is in T.

[b] Otherwise, $\forall x\ f(c_1, \ldots, c_n) \neq x$ is in T. By (A4), $f(c_1, \ldots, c_n) \neq f(c_1, \ldots, c_n)$ is in T, contradicts to (A7) (a).
[c] In particular, for a sentence φ, $\mathcal{M} \models \varphi \Leftrightarrow \varphi$ is in T, and so $\mathcal{M} \models T$.

Lecture 10: Introduction to Model Theory

Let's start by finishing the proof of Theorem 2.6.4. 7 Feb. 14:30

Proof of Theorem 2.6.4 (Continued). There's one final case left:

(e) If φ is $\forall x\ \psi(x, \overline{y})$: Because T has Henkin constants, there is d such that $\neg\forall x\ \psi(x, \overline{c}) \to \neg\psi(d, \overline{c})$ is in T.

- If $\varphi(c_1, \ldots, c_n)$ is not in T, i.e., $\forall x\ \psi(x, \overline{c})$ is in T, then since T is maximal, $\neg\forall x\ \psi(x, \overline{c})$ is in T. So by (MP), $\neg\psi(d, \overline{c})$ is in T. So, $\mathcal{M} \models \neg\psi([d], [\overline{c}])$ by induction hypotheses, hence $\mathcal{M} \models \neg\forall x\ \psi(x, [\overline{c}])$, i.e., $\mathcal{M} \not\models \varphi([\overline{c}])$.
- If $\mathcal{M} \not\models \varphi([\overline{c}])$, then $\mathcal{M}\neg \models \forall x\ \varphi(x, [\overline{c}])$, so there is $[e]$ such that $\mathcal{M} \models \neg\psi([e], [\overline{c}])$. Hence, $\neg\psi(e, \overline{c})$ is in T. Suppose for a contradiction that $\varphi(\overline{c})$, i.e., $\forall x\ \psi(x, \overline{c})$ is in T, by (A4), $\forall x\ \psi(x, \overline{c}) \to \psi(e, \overline{c})$, so $\psi(e, \overline{c})$ is in T by maximality and by consistency. But then T is inconsistent, a contradiction $\lightning$ Hence $\varphi(\overline{c})$ is not in T.

Thus, $\mathcal{M} \models T$, so T is satisfiable, proving the theorem. ■

Remark. We see that when proving the above, when we talk about $\mathcal{M}$, the witness comes for free, while for T, we need Henkin constants for getting a witness.

Now, we can complete the proof of completeness theorem by putting everything together.

Claim. The completeness theorem (b) holds.

Proof. We see that

1. Theorem 2.6.3: There is a consistent $T^* \supseteq T$ and $\mathcal{L}^*$-theory (with $\mathcal{L}^* \supseteq \mathcal{L}$) and T^* has Henkin constants.
2. Theorem 2.5.3: There is a maximal consistent $\mathcal{L}^*$-theory $T^{**} \supseteq T^*$, where T^{**} still has Henkin constants.
3. Theorem 2.6.4: T^{**} has a model $\mathcal{M}^*$ an $\mathcal{L}^*$-structure. Let $\mathcal{M}$ be the reduct of $\mathcal{M}^*$ to an $\mathcal{L}$-structure.

Hence, $\mathcal{M} \models T$. ⊛

As previously seen (Problem set 1). Let $\mathcal{L}^* \supseteq \mathcal{L}$. If $\mathcal{M}^*$ is an $\mathcal{L}^*$-structure, then by ignoring the interpretation of the symbols in $\mathcal{L}^* - \mathcal{L}$, we get an $\mathcal{L}$-structure $\mathcal{M}$.

Notation (Reduct). $\mathcal{M}$ is a *reduct* of $\mathcal{M}^*$.

Notation (Expansion). $\mathcal{M}^*$ is an *expansion* of $\mathcal{M}$.

Remark. We see that $\vdash$ and $\models$ are the same.

2.6.3 Consequences of Completeness Theorem

Size of Models

Now, let's step back and look at the proof of the completeness theorem, and ask the following.

Problem. When we did the Henkin construction of $\mathcal{M}^* \models T^{**}$, how big was M?

This can be answered by the following.

Theorem 2.6.5. If T is a satisfiable $\mathcal{L}$-theory, then it has a model of size at most $|\mathcal{L}| + \aleph_0$.

Proof. Since $|M| \leq |\mathcal{L}^*|$ since $\mathcal{M} = \mathcal{C} / \sim$, and in step one, $|\mathcal{L}^*| \leq |\mathcal{L}| + \aleph_0$, so $|M| \leq |\mathcal{L}| + \aleph_0$. ■

Example. Let $\mathcal{L} = \{f\}$, T says that f is injective but not surjective.

Example. Let $\mathcal{L} = \{\leq\}$, T says that $\leq$ is a linear order with no greatest element.

Example. Let $\mathcal{L} = \varnothing$, T says that there are at least n elements for each n.

Single Stroke and Double Stroke Style Deduction

As previously seen. $\vdash$ and $\models$ are actually $\vdash_{\mathcal{L}}$[a] and $\models_{\mathcal{L}}$[b]

[a]Proofs can only use $\mathcal{L}$-formulas.
[b]Only looking at $\mathcal{L}$.

Remark. Suppose $\mathcal{L} \supseteq \mathcal{L}_0$, and Γ a set of $\mathcal{L}_0$-sentences, φ on $\mathcal{L}_0$-sentence.

(a) $\Gamma \models_{\mathcal{L}_0} \varphi \Leftrightarrow \Gamma \models_{\mathcal{L}_1} \varphi$.

(b) $\Gamma \vdash_{\mathcal{L}_0} \varphi \Leftrightarrow \Gamma \vdash_{\mathcal{L}_1} \varphi$.

Proof. (a) and (b) are equivalent by the completeness theorem, and we prove (a).

Suppose $\Gamma \models_{\mathcal{L}_0} \varphi$. Suppose $\mathcal{M}_1$ is an $\mathcal{L}_1$-structure such that $\mathcal{M}_1 \models \Gamma$. Let $\mathcal{M}_0$ be the reduct of $\mathcal{M}_1$ to $\mathcal{L}_0$ and $\mathcal{M}_0 \models \Gamma$, so $\mathcal{M}_0 \models \varphi$, then $\mathcal{M}_1 \models \varphi$, thus $\Gamma \models_{\mathcal{L}_1} \varphi$.

Now, suppose $\Gamma \models_{\mathcal{L}_1} \varphi$. Suppose $\mathcal{M}_0$ is an $\mathcal{L}_0$-structure with $\mathcal{M}_0 \models \Gamma$. Expand $\mathcal{M}_0$ to an $\mathcal{L}_1$-structure $\mathcal{M}_1$ in any way. $\mathcal{M}_1 \models \Gamma$, so $\mathcal{M}_1 \models \varphi$. Thus, $\mathcal{M}_0 \models \varphi$, so $\Gamma \models_{\mathcal{L}_0} \varphi$. ⊛

What is important about the proof system?

(1) Soundness and completeness, $\vdash \Leftrightarrow \models$.

(2) Proofs are finite, and use only finitely many hypotheses $\Rightarrow$ compactness.

Computational Properties

Consider the following.

Definition 2.6.3 (Computably enumerable). A set is *computably enumerable* or *computable listable* if there is a program that lists out its elements.

If $\mathcal{L}$ is finite, or computable (complete list of symbols and their arities), we have the following.

(a) We can compute with formulas.

(b) Given a formula, it's computable to check whether it's a logical axiom.

(c) It's computable to check whether a proof is valid.

(d) If Γ is a computably enumerable set of sentences, $\{\varphi \colon \Gamma \vdash \varphi\}$ is also computably enumerable.[3]

(e) There is no program that given φ can decide whether $\vdash \varphi$ at least for $\mathcal{L} = \{E\}$, E binary.

[3]We can list out all the valid proofs from Γ of any φ.

Chapter 3

The Beginning of Model Theory

We now discuss various properties of the models of some theories in our interest. In particular, we care about the size of the models, and how different models with different size relate to each other.

3.1 Complete Theories

Let's start with a proposition.

Proposition 3.1.1. Let T be an $\mathcal{L}$-theory with an infinite model, and let κ be an infinite cardinal with $\kappa \geq |\mathcal{L}|$. Then T has a model of cardinality κ.

Proof. Let $\mathcal{C}$ be a set of κ-many new constants, and let $\mathcal{L}^* = \mathcal{L} \cup \mathcal{C}$. Let

$$T^* = T \cup \{c \neq d \mid c, d \in \mathcal{C} \text{ distinct}\}.$$

If $\mathcal{M} \models T^*$, then $|M| \geq \kappa$; also, if T^* is satisfiable, it has a model of size at most $|\mathcal{L}^*| = \kappa$ since

$$\kappa = |\mathcal{C}| \leq |\mathcal{L}^*| \leq |\mathcal{C}| + |\mathcal{L}| \leq \kappa + \kappa = \kappa$$

from Theorem 2.6.5. Hence, if T^* is satisfiable, it has a model $\mathcal{M}$ with $|M| = \kappa$.

Claim. T^* is satisfiable.

Proof. It's enough to show that every finite $\Gamma \subseteq T^*$ is satisfiable from the compactness theorem. Let $\mathcal{M}$ be infinite, and $\Gamma \subseteq T^*$ finite, then we can write

$$\Gamma \subseteq T \cup \{c_i \neq c_j \mid i, j = 1, \ldots, n, i \neq j\}$$

for $c_1, \ldots, c_n \in \mathcal{C}$ since only finitely many c_i are involved. Without loss of generality, $\Gamma = T \cup \{c_i \neq c_j \mid i, j = 1, \ldots, n, i \neq j\}$. Pick $a_1, \ldots, a_n \in M$, distinct, we then turn $\mathcal{M}$ into an $\mathcal{L}^*$-structure $\mathcal{M}^*$ with $c_i^{\mathcal{M}^*} = a_i$,[a] resulting in $\mathcal{M}^* \models \Gamma$. ⊛

[a] And each other $d \in \mathcal{C}$ with $d^{\mathcal{M}^*} = a_1$.

■

Lecture 11: Algebraically Closed Fields

3.2 A Detour to Algebraically Closed Fields

9 Feb. 14:30

Algebraically closed fields are a great example of a *tame* theory (as opposed to e.g., $\mathbb{N}$, which are not *tame*). We detour to discuss some important and related definitions for the future discussion.

3.2.1 Rings

All rings R we refer to will be commutative.

Definition 3.2.1 (Ideal). Let R be a ring. An *ideal* I of R is a set $I \subseteq R$ such that

(a) $0 \in I$;

(b) if $a, b \in I$, then $a + b \in I$;

(c) if $a \in I$ and $r \in R$, $ra \in I$.

Intuition. An ideal is trying to act as a set of "zeros" (in order to be further mod out).

Definition 3.2.2 (Proper). An ideal is *proper* if $1 \notin I$, equivalently, $I \notin R$.

Definition. Let I be a proper ideal.

> **Definition 3.2.3** (Radical). I is *radical* if $a^n \in I$, then $a \in I$.
>
> **Definition 3.2.4** (Prime). I is *prime* if $ab \in I$, then $a \in I$ or $b \in I$.
>
> **Definition 3.2.5** (Maximal). I is *maximal* if there is no proper ideal $J \supsetneq I$.

Remark. Maximal $\supseteq$ Prime $\supseteq$ Radical.

Definition 3.2.6 (Polynomial ring). Let R be a ring. Then $R[x_1, \ldots, x_n]$ is the *polynomial ring* with coefficients in R on indeterminates $x_1, \ldots, x_n$.

Example. Let K be a field, $S \subseteq K^n$, and $I \subseteq K[x_1, \ldots, x_n]$ defined as

$$I = \{ f(\overline{x}) \mid f(\overline{s}) = 0 \text{ for all } \overline{s} \in S \}.$$

Then I is a radical ideal.

Definition 3.2.7 (Ideal generation). Let R be a ring. The ideal I *generated* by the set $\{x_1, \ldots, x_n \in R\}$, denoted as $I = (x_1, \ldots, x_n)$, is given by

$$I = \{ r_1 x_1 + \cdots + r_n x_n \mid r_i \in R \}.$$

Intuition. The ideal generated by $\{x_i\}$ is the "smallest" ideal containing all x_i's.

Definition 3.2.8 (Principal ideal). An ideal is a *principal ideal* if it's generated by a single element.

Definition 3.2.9 (Principal ideal ring). A ring R is a *principal ideal ring* if all its ideals are principal.

As previously seen (Zero divisor). If $a, b \neq 0$, but $ab = 0$, then a and b are *zero divisors* of the ring R.

Definition 3.2.10 (Integral domain). A nontrivial ring with no zero divisors is called an *integral domain.*[a]

[a]Some authors will just call *domain*.

Definition 3.2.11 (Principal ideal domain). An integral domain where all ideals are principal is called a *principal ideal domain* or *PID*.

Theorem 3.2.1. $K[x]$ is a PID, i.e., every ideal is generated by one element as $I = (f(x)) = \{g(x)f(x) \mid g(x) \in K[x]\}$.[a]

[a]It's clear that $K[x]$ is an integral domain.

Proof. We can let g be the polynomial of the least degree in I. Then for any other $h \in I$, by long division, $h = gs + r$, with $\deg(r) < \deg(g)$. But then $r = h - gs \in I$, so if r has lower degree than g, $r = 0$, hence $h = gs \in (g)$. ∎

If it's too much to ask for an ideal generated by a single element, then we might as well consider the finite case.

Definition 3.2.12 (Noetherian ring). A ring R is *Noetherian* if every ideal I of R is finitely generated.

Remark. Equivalently, there is no infinite proper ascending chain of ideals.

Theorem 3.2.2 (Hilbert basis theorem). If R is a Noetherian ring, then $R[x]$ is also Noetherian. In particular, $K[x_1, \ldots, x_n]$ is Noetherian and so every ideal in $K[x_1, \ldots, x_n]$ is finitely generated.

As previously seen (Ring homomorphism). Let R, S be rings. A *ring homomorphism* $\varphi \colon R \to S$ is a map satisfies

(a) $\varphi(x +_R y) = \varphi(x) +_S \varphi(y)$ for $x, y \in R$;

(b) $\varphi(x \times_R y) = \varphi(x) \times_S \varphi(y)$ for $x, y \in R$;

(c) $\varphi(1_R) = 1_S$.

Theorem 3.2.3. If $\alpha \colon R \to S$ is a ring homomorphism, then $\ker \alpha$ is an ideal of R, and the induced map $\overline{\alpha} \colon R \,/\, \ker \alpha \to S$ is injective.

Theorem 3.2.4. Let R be a ring, and I an ideal of R.[a]

(a) $R \,/\, I$ is an integral domain if and only if I is a prime.

(b) $R \,/\, I$ is a field if and only if I is maximal.

[a]Then $\pi \colon R \to R \,/\, I$ is a ring homomorphism with kernel I.

3.2.2 Field Extensions

Now, we can talk about field extension.

Definition 3.2.13 (Field extension). If $K \subseteq L$ is a subfield of L, we call $L \,/\, K$ a *field extension.*

Given a field extension $L \,/\, K$, then we have that L is a K-vector space, which suggests the following natural notion.

Definition 3.2.14 (Degree). The *degree* $[L\colon K]$ of L / K is the dimension of the K-vector space L.

Notation (Finite extension). If $[L\colon K]$ is finite, then we say L / K is a *finite extension*.

Example. $\mathbb{C}$ is a field extension over $\mathbb{R}$ with $[\mathbb{C}\colon \mathbb{R}] = 2$.

Proof. Since $\mathbb{C}$ is an $\mathbb{R}$-vector space with basis $\{1, i\}$. ⊛

Example. $\mathbb{Q}(\sqrt{2})$ is a field extension over $\mathbb{Q}$ with $[\mathbb{Q}(\sqrt{2})\colon \mathbb{Q}] = 2$.

Proof. Since $\mathbb{Q}(\sqrt{2})$ is a $\mathbb{Q}$-vector space with basis $\{1, \sqrt{2}\}$. ⊛

The following is the powerful way to calculate the degree of a field extension if it can be constructed by a "tower" of field extensions.

Theorem 3.2.5. If M / L and L / K are field extensions, then $[M\colon K] = [M\colon L][L\colon K]$.

3.2.3 Algebraically Closed Fields

We care about field extensions L / K that are algebraic. This start from defining what does it mean by a single element $a \in L$ is algebraic over K.

Definition. Let L / K be a field extension, and $a \in L$.

> **Definition 3.2.15** (Algebraic). If there is a non-zero $f(x) \in K[x]$ such that $f(a) = 0$, then a is *algebraic* over K.
>
> **Definition 3.2.16** (Transcendental). If a is not algebraic, then it is *transcendental* over K.

Definition 3.2.17 (Minimal polynomial). If a is algebraic over K, there is a non-zero, monic[a] $f(x) \in K[x]$ of least degree such that $f(a) = 0$ which we call the *minimal polynomial* of a over K.

[a] This is a common practice.

Intuition. An algebraic number a is the root of some polynomials f in this polynomial ring, and we can find the minimal such f.

As previously seen (Irreducible). A non-zero non-unit of an integral domain R is *irreducible* if it cannot be written as the product of two non-units.

Note. A minimal polynomial is irreducible.

Remark. If $f(x)$ is a minimal polynomial, then $(f(x)) = \{g(x) \in K[x] \mid g(a) = 0\}$.

Example. Consider a field extension $\mathbb{R} / \mathbb{Q}$ with $a = \sqrt{2} \in \mathbb{R}$. Then the minimal polynomial is $f(x) = x^2 - 2$.

Theorem 3.2.6. Let L / K be a field extension and $a \in L$, then a is algebraic over K if and only if $n = [K(a)\colon K] < \infty$. Furthermore, if a is algebraic over K, then n is the degree of the minimal polynomial of a, and $1, a, \ldots, a^{n-1}$ is a basis for $K(a)$ as a K-vector space.

Proof idea. Think about $f(a) = a^n + r_{n-1}a^{n-1} + \cdots + r_1 a + r_0 1 = 0$. ∎

The following example illustrates how can we combine Theorem 3.2.5 and Theorem 3.2.6,

Example. Let $f(x) = x^2 - 2$, $\mathbb{Q}(\sqrt{2}) = \left\{a1 + b\sqrt{2} \mid a, b \in \mathbb{Q}\right\}$.

$$\begin{tikzcd}
& \mathbb{Q}(\sqrt[6]{2}) & \\
\mathbb{Q}(\sqrt{2}) \arrow[ur, "3"] & & \mathbb{Q}(\sqrt[3]{2}) \arrow[ul, "2"'] \\
& \mathbb{Q} \arrow[ul, "2"] \arrow[ur, "3"'] \arrow[uu, "6"] &
\end{tikzcd}$$

Theorem 3.2.7. Let $L \,/\, K$ be a field extension, $a \in L$, and $f(x) \in K[x]$ be the minimal polynomial of a over K.

(a) $K[x] \,/\, (f(x)) \cong K(a)$.[a]

(b) If $b \in L$ has the same minimal polynomial as a, then $K(a) \cong K[x] \,/\, (f(x)) \cong K(b)$.

[a] Let $x \in K[x]$, then $\overline{x} = x + (f(x)) \in K[x] \,/\, (f(x))$, i.e., $\overline{x}$ is a root of f, hence the isomorphism is given by $\overline{x} \mapsto a$.

Example. Let $a = \sqrt{2}, b = -\sqrt{2}$, and $f(x) = x^2 - 2$ with $K = \mathbb{Q}$. Then

$$\begin{aligned}
\mathbb{Q}(\sqrt{2}) \cong \quad & \mathbb{Q}[x] \big/ (x^2 - 2) & \cong \quad & \mathbb{Q}(-\sqrt{2}); \\
a + b\sqrt{2} \mapsto \quad & [a + bx] & \mapsto \quad & a - b\sqrt{2}.
\end{aligned}$$

Then, it's now natural to talk about a algebraic extension.

Definition 3.2.18 (Algebraic extension). Let $L \,/\, K$ be a field extension. Then L is an *algebraic extension* of K if all $a \in L$ are algebraic over K.

If a is algebraic over K, then $K(a) \,/\, K$ is algebraic: If $b \in K(a)$, then $K(b) \subseteq K(a)$, so $[K(b)\colon K] \leq [K(a)\colon K] < \infty$, so b is algebraic over K.

Theorem 3.2.8. If $M \,/\, L$ and $L \,/\, K$ are algebraic extensions, then $M \,/\, K$ is an algebraic extension.

Proof. Let $a \in M$, and let $b_1, \ldots, b_n \in L$ be the coefficients of the minimal polynomial of a over L. Then $b_1, \ldots, b_n$ are algebraic over K. Since

$$\begin{aligned}
[K(a)\colon K] &\leq [K(a, b_1, \ldots, b_n)\colon K] \\
&= [K(a, b_1, \ldots, b_n)\colon K(b_1, \ldots, b_n)] \cdot [K(b_1, \ldots, b_n)\colon K(b_2, \ldots, b_n)] \cdots [K(b_n)\colon K].
\end{aligned}$$

Since each of these is a finite extension, so $[K(a)\colon K] < \infty$. ∎

Definition 3.2.19 (Algebraically closed). A field L is *algebraically closed* if any non-constant $f(x) \in L[x]$ has a root in L.

Definition 3.2.20 (Algebraic closure). If $L \,/\, K$, then L is an *algebraic closure* of K if L is algebraically closed and an algebraic extension of K.

Remark. Over an algebraically closed field K, any polynomial $f(x) \in K[x]$ factors completely into $f(x) = (x - a_1) \cdots (x - a_n)$ for $n = \deg f$.

Example. $\mathbb{C}$ is algebraically closed, while $\mathbb{R}$ is not.

Example. $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$, and $[\mathbb{C} \colon \mathbb{R}] = 2$.

Example. $\mathbb{Q}^{\mathrm{alg}} = \{a \in \mathbb{C} \mid a \text{ is algebraic over } \mathbb{Q}\}$ is the algebraic closure of $\mathbb{Q}$.

If L is algebraically closed, any $f(x) \in L[x]$ factors completely as $f(x) = (x - a_1) \cdots (x - a_n)$ and $a_1, \dots, a_n$ are the only roots of f.

Theorem 3.2.9. Every field K has an algebraic closure. If $L \,/\, K$ and $M \,/\, K$ are algebraic closures over K, then $L \cong_K M$.[a]

[a] There exists $\alpha \colon L \to M$ such that $\alpha(a) = a$ for $a \in K$.

Proof. First, we show the existence. Let $f_1, f_2, \dots$ be (non-constant) polynomials over K. Start with $K = K_0$, let $g_1(x)$ be an irreducible factor of $f_1(x)$ and consider

$$K_1 \coloneqq K_0[x] \big/ (g_1(x)).$$

Since g_1 is irreducible, $(g_1(x))$ is maximal, so K_1 is a field with a root of f_1. Now, we build

$$K = K_0 \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K^* = \bigcup_i K_i$$

in the same way such that K_i contains a root of $f_i(x)$. Since any $f(x) \in K$ has a root in K^*, so $K^* \,/\, K$ is algebraic. Now, we do the same construction for K^* to get

$$K \subseteq K^* \subseteq K^{**} \subseteq K^{***} \subseteq \cdots \subseteq L = \bigcup K^{*\cdots},$$

then L is algebraically closed since any non-constant polynomial with coefficients in L actually has coefficients in one of the $K^{*\cdots*}$, so it has a root in the next field. Now we prove the uniqueness.

Lemma 3.2.1. An algebraically closed field L has no proper algebraic extensions M.

Proof. If $a \in M$ is algebraic over L for some M, the minimal polynomial $f(x)$ of a factors completely (irreducible), so $f(x) = x - r$ for $r \in L$ with $f(a) = 0$, i.e., $a = r$, so $M = L$. ∎

Lemma 3.2.2. Let $L \,/\, K$ algebraic, $M \,/\, K$ algebraically closed. Then there is an embedding $\alpha \colon L \to M$ fixing K.

Proof. Consider the case that $L = K(a)$[a] with a algebraic over K, and let $f(x)$ be the minimal polynomial of a over K. Then there is a root $b \in M$ of f with $K(a) \cong K[x] \,/\, (f) \cong K(b) \subseteq L$ from Theorem 3.2.7. Let this isomorphism be our α. ∎

[a] Once this is done, repeat iteratively and get the general case by using Zorn's lemma or transfinite induction.

Hence, if $L \,/\, K$ and $M \,/\, K$ are algebraic closures over K, there is an embedding $\alpha \colon L \to M$ over K.

Finally, since $M \,/\, \alpha(L)$ is an algebraic extension, and $\alpha(L) \cong L$ is algebraically closed, by Lemma 3.2.1, $M = \alpha(L)$, so α is an isomorphism $L \to M$ over K. ∎

Lecture 12: The ACF Theory and Categorical

14 Feb. 14:30

Definition 3.2.21 (Characteristic)**.** A field F has finite *characteristic* $p > 0$ if $\underbrace{1 + \cdots + 1}_{p \text{ times}} = 0$.

Remark. p is always prime, otherwise, F has characteristic $p = 0$, i.e., $1 + \cdots + 1 \neq 0$, always.

The following notion comes up naturally.

Definition 3.2.22 (Prime field)**.** The *prime field* $\mathbb{F}_p$ in characteristic p such that $\mathbb{F}_p = \mathbb{Q}$ if $p = 0$, $\mathbb{F}_p = \mathbb{Z} / p\mathbb{Z}$ if $p > 0$.

Definition 3.2.23 (Transcendence basis)**.** Let L / K be a field extension. A set $S \subseteq L$ is called a *transcendence basis* of L / K if S is algebraically independent[a] and L is an algebraic extension of $K(S)$, i.e., S is maximal.

[a]No $a_1, \ldots, a_n \in S$ have non-zero polynomial $f(x_1, \ldots, x_n) \in K[\overline{x}]$ with $f(a_1, \ldots, a_n) = 0$.

Remark. Every field extension has a transcendence basis(and all transcendence basis have the same size).

Proof. On a combinatorial level, this is exactly the same as the proof that any two bases for a vector space have the same cardinality. ⊛

Example. Let $K(t_1, \ldots, t_n)$ be the fraction field of $K[x_1, \ldots, x_n]$, then $\{t_1, \ldots, t_n\}$ is a transcendence basis for $K(t_1, \ldots, t_n)$ over K.

Definition 3.2.24 (Transcendence degree)**.** The *transcendence degree* of L over K is the cardinality of any transcendence basis.

If we do not specify K, then K is the prime field $K = \mathbb{F}_p$.

Theorem 3.2.10. Any two algebraically closed fields of the same characteristic p and transcendence degree are isomorphic.

Proof. Let L, K be those fields, with transcendence basis S, T over $\mathbb{F}_p$ with $|S| = |T|$. L is the algebraic closure of $\mathbb{F}_p(S)$ and K is the algebraic closure of $\mathbb{F}_p(T)$. There is a bijection $f \colon S \to T$, and then f extends to $\overline{f} \colon \mathbb{F}_p(S) \to \mathbb{F}_p(T)$ such that

$$\overline{f}\left(\frac{\sum_\alpha r_\alpha \overline{x}^\alpha}{\sum_\alpha s_\alpha \overline{x}^\alpha}\right) = \frac{\sum_\alpha r_\alpha f(\overline{x})^\alpha}{\sum_\alpha s_\alpha f(\overline{x})^\alpha},$$

where $r_\alpha, s_\alpha \in \mathbb{F}_p$ and $\overline{x}^\alpha$ is some monomial from S, e.g., $x_1^2 x_2$ for $x_1, x_2 \in S$.[a]

$\mathbb{F}_p(S)$ and $\mathbb{F}_p(T)$ are the same (up to isomorphism), but the algebraic closures are unique from Theorem 3.2.9, so $K \cong L$ via an isomorphism extending $\overline{f}$. ■

[a]α can be thought as a tuple, in the case of $x_1^2 x_2$, $\alpha = (2, 1)$.

The above proof actually shows more.

Corollary 3.2.1. If L / K and M / K are field extensions with transcendence bases S and T, and $\alpha \colon S \to T$ is a bijection, then α extends to an isomorphism $L \cong_K M$.

If we apply this inside a single algebraically closed field, we have the following.

Theorem 3.2.11. Let K be the algebraic closure of k, and L, M be subfields of K which extend k. Suppose that $\alpha \colon M \to L$ is an isomorphism fixing k, then α extends to an automorphism of K.

3.3 The ACF Theory

Finally, we are ready to introduce the theory we're going to study, which is called ACF. It turns out that the models of which are exactly the algebraically closed fields with nice properties we're going to discuss.

Definition 3.3.1 (ACF)**.** ACF is the theory of algebraically closed fields consists of field axioms and formulas that for every $n \geq 1$,

$$\forall a_0 \ldots \forall a_n \left(a_n \neq 0 \to \exists b \; a_n b^n + a_{n-1} b^{n-1} + \cdots + a_0 = 0\right).$$

Remark. The models of ACF are exactly the algebraically closed fields with the language $\mathcal{L} = \mathcal{L}_{\text{ring}} = \{0, 1, +, -, \cdot\}$.

Notation (ACF_p)**.** For a prime $p > 0$, let $\mathrm{ACF}_p \coloneqq \mathrm{ACF} \cup \{\underbrace{1 + \cdots + 1}_{p} = 0\}$.

Notation (ACF_0)**.** Let $\mathrm{ACF}_0 \coloneqq \mathrm{ACF} \cup \{\underbrace{1 + \cdots + 1}_{n} \neq 0 \mid n \in \mathbb{N}\}$.

Definition 3.3.2 (Categorical)**.** Let κ be an infinite cardinal and T be an $\mathcal{L}$-theory. We say T is *κ-categorical* if for any $\mathcal{M}, \mathcal{N} \models T$ of size κ, we have $\mathcal{M} \cong \mathcal{N}$.

Definition 3.3.3 (Countably categorical)**.** If κ is countable, then T is *countably categorical*.

Definition 3.3.4 (Uncountably categorical)**.** If κ is uncountable, then T is *uncountably categorical*.

We see that for being uncountably categorical, we only need one uncountable κ.

Example. $(\mathbb{Q}, \leq)$ is countably categorical.

Lemma 3.3.1. If K has transcendence degree λ, then $|K| = \lambda + \aleph_0$.

Proof. Let K be algebraic over $\mathbb{F}_p(S)$, where S is a transcendence basis of size λ. By counting, $|\mathbb{F}_p(S)| = \lambda + \aleph_0$, so $|\mathbb{F}_p(S)[x]| = \lambda + \aleph_0$. But since each element of K satisfies some polynomials, and each polynomial has finitely many roots in K, so $|K| = \lambda + \aleph_0$. ■

Theorem 3.3.1. For each p, ACF_p is κ-categorical for every uncountable κ.

Proof. Let L, K be a model of ACF_p for size κ. From Theorem 3.2.10, if L, K have transcendence degree κ, then they are isomorphic. With the application of Lemma 3.3.1, we're done. ■

Example. $\mathbb{Q}^{\text{alg}}$, the algebraic closure of $\mathbb{Q}$, has size $\aleph_0$ with transcendence degree 0.

Example. $\mathbb{Q}(t)^{\text{alg}}$, the algebraic closure of $\mathbb{Q}(t) \cong \mathbb{Q}(\pi)$, has size $\aleph_0$ with transcendence degree 1.

The above implies $\mathbb{Q}^{\text{alg}} \ncong \mathbb{Q}(t)^{\text{alg}}$, which lead to the following.

Remark. ACF_0 is not countably categorical. The same with ACF_p for $p > 0$.

Proof. Notice that
$$\mathbb{Q}(t)^{\text{alg}} = \{z \in \mathbb{C} \mid z \text{ is algebraic over } \mathbb{Q}(\pi)\},$$
which is countable. With $\mathbb{Q}^{\text{alg}} \ncong \mathbb{Q}(t)^{\text{alg}}$, we have the result. ⊛

Note. ACF is not uncountably categorical.

Theorem 3.3.2 (Vaught's test). Let T be a satisfiable $\mathcal{L}$-theory with no finite models. If T is κ-categorical for some infinite $\kappa \geq |\mathcal{L}|$, then T is complete.

Proof. Suppose T is not complete, then we can pick φ with $T \not\models \varphi$ and $T \not\models \neg\varphi$, i.e., $T \cup \{\varphi\}$ and $T \cup \{\neg\varphi\}$ are satisfiable. By a consequence of the proof of completeness theorem (with a compactness argument),

- $T \cup \{\varphi\}$ has a model $\mathcal{M}$ of size κ, and
- $T \cup \{\neg\varphi\}$ has a model $\mathcal{N}$ of size κ.

But T is κ-categorical, so $\mathcal{M} \cong \mathcal{N}$, which is a contradiction $\lightning$ ■

Corollary 3.3.1. ACF_p is complete for each p.

Proof. Follows immediately by Theorem 3.3.1 and Vaught's test. ■

The axioms for ACF_p completely determines all first-order facts about algebraically closed fields of characteristic p.

Remark. $\{\varphi \mid \text{ACF} \models \varphi\}$ and $\{\varphi \mid \text{ACF}_p \models \varphi\}$ can be listed computably.

Proof. Since the axioms for ACF or ACF_p can be listed computably. ⊛

Definition 3.3.5 (Decidable). A theory T is *decidable* if there is a program that given φ, it determines whether $T \models \varphi$ or $T \not\models \varphi$.

Corollary 3.3.2. ACF_p is decidable for each p.

Proof. Given φ, either $\text{ACF}_p \models \varphi$ or $\text{ACF}_p \models \neg\varphi$ since ACF_p is complete. By looking for a proof of φ and a proof of $\neg\varphi$, eventually we will find one, telling us whether $\text{ACF}_p \models \varphi$. ■

Corollary 3.3.3. ACF is decidable.

Proof. Given φ, the algorithm simultaneously

- looks for a proof of $\text{ACF} \vdash \varphi$, and
- looks for p such that $\text{ACF}_p \vdash \neg\varphi$ (hence $\text{ACF} \not\models \varphi$).[a]

If $\text{ACF} \models \varphi$, then we will halt at the first case. Now suppose $\text{ACF} \not\models \varphi$, i.e., there is $\mathcal{M} \models \text{ACF}$ such that $\mathcal{M} \models \neg\varphi$, and also, there is p such that $\mathcal{M} \models \text{ACF}_p$. Since ACF_p is complete, $\text{ACF}_p \models \neg\varphi$, so the search of the second case will half, hence the whole search will eventually halt. ■

[a] It might not be true that $\text{ACF} \models \neg\varphi$, we don't know.

Lecture 13: Upward Löwenheim-Skolem Theorem

Another consequence of completeness is that since $\mathbb{C} \models \text{ACF}_0$, if K is any algebraically closed field of characteristic 0, $\mathbb{C} \equiv K$, i.e., we have the following. 16 Feb. 14:30

Remark. The sentences true of $\mathbb{C}$ are exactly the same as the sentences true of any algebraically closed field.

Essentially, the idea is that if one proves an algebraic statement about the complex numbers by analytic techniques of complex analysis, then there will be a proof of the same algebraic statement using purely algebraic tools, which works in any algebraically closed field of characteristic 0. The compactness theorem also gives connections to fields of finite characteristic.

Theorem 3.3.3 (Leftschetz principle)**.** Let $\mathcal{L}$ be the language of rings. For an $\mathcal{L}$-sentence φ, the following are equivalent:

(i) φ is true in $\mathbb{C}$;

(ii) φ is true in every algebraically closed field of characteristic 0;

(iii) φ is true in some algebraically closed fields of characteristic 0;

(iv) there is a number n such that φ is true in all algebraically closed fields of characteristic $p > n$;

(v) for each number n, φ is true in all algebraically closed fields of characteristic $p > n$.

Proof. Let $K \models \mathrm{ACF}_0$, then since it's complete, we know that $K \models \varphi \Leftrightarrow \mathrm{ACF}_0 \models \varphi$, which proves the first three. Others are left as homework. ■

We can use the Leftschetz principle to prove the following.

Theorem 3.3.4 (Ax-Grothendieck theorem)**.** Let $f\colon \mathbb{C}^n \to \mathbb{C}^n$ be a polynomial map.[a] If f is injective, then it's surjective. More generally, this is true for any $K \models \mathrm{ACF}_p$ for any p.

[a] I.e., $f(\overline{x}) = (f_1(\overline{x}), \ldots, f_n(\overline{x}))$ where $f_1, \ldots, f_n$ are polynomials.

Proof. The claim can be expressed by the sentences, so by Leftschetz principle, it's enough to prove that if for $K = \overline{\mathbb{F}_p}$, for each $p > 0$.

Let $f\colon \overline{\mathbb{F}}_p^n \to \overline{\mathbb{F}}_p^n$ be an injective polynomial map and $\overline{y} \in \overline{\mathbb{F}}_p^n$. Then there is a finite subfield $L \subseteq \overline{\mathbb{F}}_p$ which contains $\overline{y}$ and the coefficients of f. Then, f restricts to an injective function $L^n \to L^n$, which is surjective because L^n is finite, so $\exists \overline{x} \in L^n$ such that $f(\overline{x}) = \overline{y}$. ■

3.4 Up and Down

After introducing categorical, and seeing that ACF as an example, we now study something else: when does a model exists w.r.t. some original model, and what size is it?

3.4.1 Diagrams

One way to capture the structure in terms of theory is using the so-called "diagrams".

Definition. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Let $\mathcal{L}_M \supseteq \mathcal{L}$ be the expanded language with a new constant symbol $\underline{a}$ for each $a \in M$.

Definition 3.4.1 (Atomic diagram)**.** The *atomic diagram* of $\mathcal{M}$ is the $\mathcal{L}_M$-theory

$$\mathrm{Diag}(\mathcal{M}) \coloneqq \{\varphi(\underline{a}_1, \ldots, \underline{a}_n) \mid \mathcal{M} \models \varphi(m_1, \ldots, m_n) \text{ and } \varphi \text{ is atomic or negated of atomic}\}.$$

Definition 3.4.2 (Elementary diagram)**.** The *elementary diagram* of $\mathcal{M}$ is the $\mathcal{L}_M$-theory

$$\mathrm{Diag}_{\mathrm{el}}(\mathcal{M}) \coloneqq \{\varphi(\underline{a}_1, \ldots, \underline{a}_n) \mid \mathcal{M} \models \varphi(m_1, \ldots, m_n) \text{ and } \varphi \text{ an } \mathcal{L}\text{-formula}\}.$$

Intuition. Basically both $\mathrm{Diag}(\mathcal{M})$ and $\mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$ contain the information about the structure but in the form of a theory.

Notation. There's a canonical way of expanding $\mathcal{M}$ to an $\mathcal{L}_M$-structure with $\underline{a}^{\mathcal{M}} := a$, i.e., we write a for both the symbol and the element.

Lemma 3.4.1. Let $\mathcal{N}$ be an $\mathcal{L}_M$-structure.

(a) If $\mathcal{N} \models \mathrm{Diag}(\mathcal{M})$ then, viewing $\mathcal{N}$ as an $\mathcal{L}$-structure, there is an embedding $f\colon \mathcal{M} \to \mathcal{N}$.

(b) If $\mathcal{N} \models \mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$, then there is an elementary $\mathcal{L}$-embedding of $\mathcal{M}$ into $\mathcal{N}$.

Proof. Take $f(a) = \underline{a}^{\mathcal{N}}$, then $\mathcal{N} \models \mathrm{Diag}(\mathcal{M})$ means exactly that f is an embedding, and $\mathcal{N} \models \mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$ means that f is an elementary embedding. ■

3.4.2 Upward Löwenheim-Skolem theorem

Theorem 3.4.1 (Upward Löwenheim-Skolem theorem)**.** Let $\mathcal{M}$ be an infinite $\mathcal{L}$-structure and let κ be an infinite cardinal $\kappa \geq |\mathcal{M}| + |\mathcal{L}|$. Then there is an $\mathcal{L}$-structure $\mathcal{N}$ of cardinality κ such that $j\colon \mathcal{M} \to \mathcal{N}$ is elementary.

Proof. $\mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$ is satisfiable since $\mathcal{M} \models \mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$, by Proposition 3.1.1 it has a model $\mathcal{N}$ of cardinality $\kappa \geq |\mathcal{L}_M| = |\mathcal{M}| + |\mathcal{L}|$, so an elementary embedding exists $\mathcal{M} \to \mathcal{N}$ by Lemma 3.4.1. ■

Intuition. The upward Löwenheim-Skolem theorem says that every structure is an elementary substructure of many much bigger structures.

As previously seen. Our very first application of the compactness theorem: in the construction of the non-standard model of arithmetic, we built $\mathcal{N} \models \mathrm{Th}(\mathbb{N})$ not isomorphic to $\mathbb{N}$, which is exactly like this. Every element of $\mathbb{N}$ can already be expressed as a term of the form $1 + \cdots + 1$ without adding any new constants.

3.4.3 Downward Löwenheim-Skolem theorem

A "downward" version of the upward Löwenheim-Skolem theorem also exists, which says that big models contain smaller elementary substructures. This will take some more work to prove, so we first need a test for this.

Proposition 3.4.1 (Tarski-Vaught test)**.** Let $\mathcal{M}$ be a substructure of $\mathcal{N}$. Then $\mathcal{M}$ is an elementary substructure of $\mathcal{N}$ if and only if for any formula $\varphi(x, \overline{y})$ and $\overline{a} \in M^n$, if there is $b \in N$ such that $\mathcal{N} \models \varphi(b, \overline{a})$, then there is $c \in M$ such that $\mathcal{N} \models \varphi(c, \overline{a})$.

Proof. The forward direction follows from the fact that $\mathcal{M}$ is an elementary substructure, so the truth of $\exists x\ \varphi(x, \overline{y})$ is proved since $\mathcal{M} \models \exists x\ \varphi(x, \overline{a})$ if and only if $\mathcal{N} \models \exists x\ \varphi(x, \overline{a})$.

For the backward direction, suppose the condition holds. We show that $\mathcal{M} \models \varphi(\overline{a}) \Leftrightarrow \mathcal{N} \models \varphi(\overline{a})$ by induction on φ. Since $\mathcal{M}$ is a substructure of $\mathcal{N}$, this is true for all quantifier-free formulas, and in particular for the atomic formulas.

Suppose that the claim is true for ψ, then

$$\mathcal{M} \models \neg\psi(\overline{a}) \Leftrightarrow \mathcal{M} \not\models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \not\models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \models \neg\psi(\overline{a}),$$

so the claim is also true for $\neg\psi$. Similarly, suppose the claim holds for φ, ψ. Then,

$$\mathcal{M} \models (\varphi \wedge \psi)(\overline{a}) \Leftrightarrow \mathcal{M} \models \varphi(\overline{a}) \text{ and } \mathcal{M} \models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \models \varphi(\overline{a}) \text{ and } \mathcal{N} \models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \models (\varphi \wedge \psi)(\overline{a}).$$

Finally, suppose the claim holds for $\varphi(x, \overline{y})$, then

$$\mathcal{M} \models \exists x\ \varphi(x, \overline{a}) \Rightarrow \exists b \in M\ \mathcal{M} \models \varphi(b, \overline{a}) \Rightarrow \exists b \in M\ \mathcal{N} \models \varphi(b, \overline{a}) \Rightarrow \mathcal{N} \models \exists x\ \varphi(x, \overline{a}),$$

by induction hypotheses. Conversely, $\mathcal{N} \models \exists x\ \varphi(x, \overline{a})$, then $\exists b \in N$ such that $\mathcal{N} \models \varphi(b, \overline{a})$ by the condition from the statement, so $\exists c \in M$ such that $\mathcal{N} \models \varphi(c, \overline{a})$. By the induction hypotheses, we further have $\mathcal{M} \models \varphi(c, \overline{a})$, hence $\mathcal{M} \models \exists x\ \varphi(x, \overline{a})$. ∎

Example. The ring $\mathbb{Z}$ is a substructure of $\mathbb{Q}$, but $\mathbb{Q} \models \exists x\ (x + x = 1)$ while $\mathbb{Z} \not\models \exists x\ (x + x = 1)$.

Lecture 14: Downward Löwenheim-Skolem theorem

We will also need to introduce the Skolemizations of theories. 21 Feb. 14:30

Definition 3.4.3 (Built-in Skolem function). We say an $\mathcal{L}$-theory T has *built-in Skolem functions* if for all $\mathcal{L}$-formulas $\varphi(x, y_1, \ldots, y_n)$, there is a function symbol f such that

$$T \models \forall \overline{y}\ (\exists x\ \varphi(x, \overline{y}) \to \varphi(f(\overline{y}), \overline{y})).$$

Intuition. This is like a parametrized version of Henkin constants.

Lemma 3.4.2. Let T be an $\mathcal{L}$-theory, then there are $\mathcal{L}^* \supseteq \mathcal{L}$ and $T^* \supseteq T$ an $\mathcal{L}^*$-theory such that T^* has built-in Skolem functions. Moreover, if $\mathcal{M} \models T$, then we can expand $\mathcal{M}$ to $\mathcal{M}^* \models T^*$. Finally, we can choose $\mathcal{L}^*$ such that $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$.

Proof. Start with $\mathcal{L}_0 = \mathcal{L}$ and $T_0 = T$, we build $\mathcal{L}_0 \subseteq \mathcal{L}_1 \subseteq \ldots$ and $T_0 \subseteq T_1 \subseteq \ldots$ and let $\mathcal{L}^* = \bigcup_i \mathcal{L}_i$ and $T^* = \bigcup_i T_i$. Given $\mathcal{L}_i$ and T_i, define

$$\mathcal{L}_{i+1} = \mathcal{L}_i \cup \{f_\varphi \mid \varphi(x, \overline{y}) \text{ is an } \mathcal{L}_i\text{-formulas}\}$$

where the arity of f_φ is the same as $\overline{y}$, and

$$T_{i+1} = T_i \cup \{\forall \overline{y}\ (\exists x\ \varphi(x, \overline{y}) \to \varphi(f_\varphi(\overline{y}), \overline{y}))\}.$$

We now argue that if $\mathcal{M}_i \models T_i$, we can expand it to a model $\mathcal{M}_{i+1}$ of T_{i+1}. Pick $c \in M_i$ a "default value." Given φ and $\overline{a}$, define $f_\varphi^{\mathcal{M}_{i+1}}(\overline{a})$ to be some b with $\mathcal{M}_i \models \varphi(b, \overline{a})$ if such a b exists, or c otherwise.[a] Then, $\mathcal{M}_{i+1} \models T_{i+1}$. From this construction, we see that T^* has built-in Skolem functions since any $\mathcal{L}^*$-formula φ is in some $\mathcal{L}_i$ and has a Skolem function in $\mathcal{L}_{i+1}$.

Now, suppose $\mathcal{M} \models T$, i.e., $\mathcal{M} = \mathcal{M}_0 \models T_0$. From above, $\mathcal{M}_0$ has an expansion $\mathcal{M}_1 \models T_1$, which has an expansion $\mathcal{M}_2 \models T_2$, etc. By expanding $\mathcal{M}$ iteratively, we get a model $\mathcal{M}^*$ of T^*.

Finally, at each step, we add one symbol for each $\mathcal{L}_i$-formula to $\mathcal{L}_i$, hence by counting, $|\mathcal{L}_{i+1}| = |\mathcal{L}_i| + \aleph_0$, hence $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$. ∎

[a] If such b doesn't exist, then the left-hand side is false, so we don't really care about the right-hand side.

Note. We see that this is a similar argument to when we added Henkin constants, though it's simpler now because we can work semantically.

Notation (Skolemization). We call T^* in Lemma 3.4.2 a *Skolemization* of T.

Theorem 3.4.2 (Downward Löwenheim-Skolem theorem). Let $\mathcal{M}$ be an $\mathcal{L}$-structure and $X \subseteq M$. Then there is an elementary substructure $\mathcal{N}$ of $\mathcal{M}$ with $X \subseteq \mathcal{N}$ and $|N| \leq |X| + |\mathcal{L}| + \aleph_0$.

Proof. By expanding the language, we get $\mathcal{M}^*$ and $\mathcal{L}^*$-structure with $\mathrm{Th}(\mathcal{M}^*)$ has built-in Skolem functions (where $T = \mathrm{Th}(\mathcal{M})$ in Lemma 3.4.2). Hence, by replacing $\mathcal{M}$ by $\mathcal{M}^*$, etc., we may assume that we already had built-in Skolem functions.

Start with $X_0 = X \cup \left\{ c^{\mathcal{M}} \mid c \text{ a constant symbol} \right\}$. Given X_i, define X_{i+1} as

$$X_{i+1} = X_i \cup \left\{ f^{\mathcal{M}}(\overline{a}) \mid f \in \mathcal{L} \text{ a function symbol}, \overline{a} \in X_i^{n_f} \right\}.$$

Let $N = \bigcup_i X_i$, and let $\mathcal{N}$ be the substructure of $\mathcal{M}$ with domain N. This can be done by:

- for each function symbol f, let $f^{\mathcal{N}}$ be the restriction of $f^{\mathcal{M}}$ to N;[a]
- for each relation symbol R, let $R^{\mathcal{N}}$ be the restriction of $R^{\mathcal{M}}$ to N;
- for each constant symbol c, there is a Skolem function f such that $f(x) = c^{\mathcal{M}} \in N$ for all $x \in M$, so let $c^{\mathcal{N}} = c^{\mathcal{M}}$.

Now, to show that $\mathcal{N}$ is an elementary substructure of $\mathcal{M}$, we use the Tarski-Vaught test. Suppose that we have an $\mathcal{L}$-formula $\varphi(x, \overline{y})$, $\overline{a} \in N$, $b \in M$ such that $\mathcal{M} \models \varphi(b, \overline{a})$, and we must replace b by $c \in N$. Since $\mathcal{M} \models \exists x\ \varphi(x, \overline{a})$, so $\mathcal{M} \models \varphi(f_\varphi(\overline{a}), \overline{a})$. But since $\overline{a} \in N$, so $f_\varphi(\overline{a}) \in N$, so the Tarski-Vaught test says $\mathcal{N}$ is an elementary substructure of $\mathcal{M}$.

Finally, since $|X_0| \leq |X| + |\mathcal{L}| + \aleph_0$, with N being a countable union, $|N| \leq |X| + |\mathcal{L}| + \aleph_0$. ■

[a] From the definition of N, it is closed under the applications of functions f, so the restriction takes values in N.

Notation (Generated substructure)**.** $\mathcal{N}$ in the downward Löwenheim-Skolem theorem is called the *substructure generated* by X.

Example (Countable real closed filed)**.** Consider $\mathbb{R} = (\mathbb{R}, 0, 1, +, -, \cdot, \leq)$. Let $X \subseteq \mathbb{R}$ be countable, e.g., $X = \varnothing$ or $X = \{\pi, e\}$. Then there is $X \subseteq \mathcal{R} \preceq \mathbb{R}$ such that $\mathcal{R}$ is countable. In particular, $\mathrm{Th}(\mathcal{R}) = \mathrm{Th}(\mathbb{R})$ and R is a *countable real closed filed*, i.e.,

(a) -1 is not a sum of squares;

(b) for all a, there is b such that $a = b^2$ or $a = -b^2$;

(c) every odd degree polynomial has a root.

There is a whole theory of real closed fields just like for algebraically closed fields.

Intuition. Countable real closed fields are as algebraically closed as they can be while still being orderable.

Example (Skolem's paradox)**.** Let $\mathcal{L} = \{\in\}$ be the language of set theory, where $\in$ a binary relation symbol. Let $T = \mathrm{ZFC}$, and suppose ZFC is satisfiable,[a] i.e., there is a model $\mathcal{M}$ such that $\mathcal{M} \models T$.

Then, there is a countable $\mathcal{N} \preceq \mathcal{M}$, in particular, there is a countable model of ZFC. We can then write down the sentence

$$\varphi := \text{"there is no bijection between } \mathbb{R}^{\mathcal{N}} \text{ and } \mathbb{N}^{\mathcal{N}}\text{"}$$

such that $\mathcal{N} \models \varphi$, in $\mathcal{L}$. Observe that $\mathcal{N}$ thinks that it contains an uncountable set $\mathbb{R}^{\mathcal{N}}$, but $\left\{ a \in N \mid \mathcal{N} \models a \in \mathbb{R}^{\mathcal{N}} \right\} \subseteq N$ is countable! This is called *Skolem's paradox.*

[a] From Gödel's incompleteness theorem, in ZFC, one can't prove that ZFC is consistent.

We finish this section with two useful facts.

Definition 3.4.4 (Universally axiomatizable)**.** Let T be an $\mathcal{L}$-theory, then T is *universally axiomati-*

zable if there is a set Γ of universal sentences such that $T \models \Gamma$ and $\Gamma \models T$.[a]

[a] I.e., $\mathcal{M} \models T$ if and only if $\mathcal{M} \models \Gamma$.

Theorem 3.4.3. Let T be an $\mathcal{L}$-theory. T is universally axiomatized if and only if whenever $\mathcal{N} \models T$ and $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{M} \models T$.

Proof. The forward direction is easy: suppose that T is universally axiomatized by Γ. If $\mathcal{M} \subseteq \mathcal{N}$ and $\mathcal{N} \models T$, then $\mathcal{N} \models \Gamma$, and since Γ consists only of universal formulas, $\mathcal{M} \models \Gamma$, so $\mathcal{M} \models T$.

Now, to prove the backward direction, suppose that if $\mathcal{N} \models T$, $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{M} \models T$. Define

$$\Gamma = \{\varphi \text{ universal} \mid T \models \varphi\},$$

then $T \models \Gamma$. Now, we need to show that $\Gamma \models T$. We may assume that T is satisfiable[a] and let $\mathcal{M} \models \Gamma$, so we now want to prove that $\mathcal{M} \models T$. This can be done by finding $\mathcal{N} \supseteq \mathcal{M}$ and $\mathcal{N} \models T$, then from our assumption we have $\mathcal{M} \models T$. We build such an $\mathcal{N}$ by showing that $\operatorname{Diag}(\mathcal{M}) \cup T$ is satisfiable, and take the corresponding model to be $\mathcal{N}$. Then, trivially we have $\mathcal{N} \models T$.

This can be done by compactness theorem. Let $\Delta \subseteq \operatorname{Diag}(\mathcal{M}) \cup T$ be finite, then there is a finite set of atomic or negated atomic formulas $\varphi_1, \ldots, \varphi_\ell$ and $m_1, \ldots, m_k \in M$ such that

$$\Delta \subseteq \{\varphi_1(\overline{m}), \ldots, \varphi_\ell(\overline{m})\} \cup T,$$

assume that they are actually equal. To show that Δ is satisfiable, it is enough to show[b] that

$$\{\exists x_1 \ldots \exists x_k \ (\varphi_1(\overline{x}) \wedge \cdots \wedge \varphi_\ell(\overline{x}))\} \cup T$$

is satisfiable. If not, then since T is satisfiable,

$$T \models \forall x_1 \ldots \forall x_k \ \neg(\varphi_1(\overline{x}) \wedge \cdots \wedge \varphi_\ell(\overline{x})).$$

Since this is universal, hence in Γ, so it is true in $\mathcal{M}$. But it's also not true in $\mathcal{M}$ since $\mathcal{M} \models \varphi_1(\overline{m}) \wedge \cdots \wedge \varphi_\ell(\overline{m})$, a contradiction $\lightning$ Hence, Δ is satisfiable, so any finite subset is satisfiable, by compactness theorem, we're done. ∎

[a] Since otherwise $\Gamma \ni \forall x \ x \neq x$, and hence $\Gamma \models T$ trivially.
[b] Since the constant symbols $m_1, \ldots, m_k$ do not appear in T, we can interpret them as the witness to the $\exists$'s.

Lecture 15: The Random Graph Theory

Proposition 3.4.2. Suppose $\mathcal{M}_1 \preceq \mathcal{M}_2 \preceq \ldots$, and let $\mathcal{M} = \bigcup_i \mathcal{M}_i$. Then $\mathcal{M}_i \preceq \mathcal{M}$ for all i.[a] 23 Feb. 14:30

[a] Countability is not necessary.

Proof. By induction on formulas, for all i and $\overline{a} \in M_i$, we show that $\mathcal{M}_i \models \varphi(\overline{a})$ if and only if $\mathcal{M} \models \varphi(\overline{a})$.

(a) For φ is atomic, this is true since φ is quantifier-free and $\mathcal{M}_i \subseteq \mathcal{M}$.

(b) For $\neg, \vee, \wedge$, exactly the same as the Tarski-Vaught test.

(c) If φ is $\exists y \ \psi(\overline{x}, y)$:

- If $\mathcal{M}_i \models \exists y \ \psi(\overline{a}, y)$, there is $b \in M_i$ such that $\mathcal{M}_i \models \psi(\overline{a}, b)$. Then by the induction hypothesis, $\mathcal{M} \models \psi(\overline{a}, b)$, so $\mathcal{M} \models \exists y \ \psi(\overline{a}, y)$.
- If $\mathcal{M} \models \exists y \ \psi(\overline{a}, y)$, then there is $b \in M$ such that $\mathcal{M} \models \psi(\overline{a}, b)$. Since $M = \bigcup_j M_j$, there is $j \geq i$ such that $b \in M_j$. By the induction hypothesis, $\mathcal{M}_j \models \psi(\overline{a}, b)$, so $\mathcal{M}_j \models \exists y \ \psi(\overline{a}, y)$. Finally, since $\mathcal{M}_i \preceq \mathcal{M}_j$, so $\mathcal{M}_i \models \exists y \ \psi(\overline{a}, y)$.

∎

3.5 Back and Forth

We have examples of uncountably categorical theories, but no examples of countably categorical theories.

3.5.1 Dense Linear Order Theory

The simplest example of a countably categorical theory is the theory of "linear orders (without endpoints)," denoted as DLO.

Definition 3.5.1 (DLO)**.** Let $\mathcal{L} = \{\leq\}$. The theory of *dense linear orders (without endpoints)*, denoted as DLO, has the axioms:

(a) $\leq$ is a linear order;

(b) $\forall x \forall y\ (x < y \to \exists z\ x < z < y)$ (the density axiom);

(c) $\forall x \exists y \exists z\ (y < x < z)$ (the no-endpoints axiom).

Example. $(\mathbb{Q}, \leq)$ and $(\mathbb{R}, \leq)$ are both DLO's.

To create a new dense linear orders, given $\mathcal{M}_1, \mathcal{M}_2$ two DLO's, define $\mathcal{M}_1 + \mathcal{M}_2$ with domain $M \sqcup N$ and has each element of M less than each element of N, and within M and N, the orderings are the same as in $\mathcal{M}$ and $\mathcal{N}$. This is also a DLO.

Example. $\mathbb{Q} + \mathbb{Q}$ and $\mathbb{R} + \mathbb{R}$ are both DLO's.

Example. $\mathbb{R} + \mathbb{R} \ncong \mathbb{R}$.

Proof. Since $\mathbb{R}$ has the least upper bound property while $\mathbb{R} + \mathbb{R}$ does not (there is no least upper bound for the first copy). ⊛

Example. $\mathbb{Q} + \mathbb{Q} \cong \mathbb{Q}$.

Proof. For example, take some irrational, e.g., π. Then $\mathbb{Q} = \{x \mid x < \pi\} \cup \{x \mid x > \pi\}$, and we observe that we have

$$\{x \mid x < \pi\} \cong \mathbb{Q} \cong \{x \mid x > \pi\},$$

and hence piecing them together we have $\mathbb{Q} + \mathbb{Q} \cong \mathbb{Q}$. ⊛

Example. $\mathbb{Q} + \mathbb{R} \ncong \mathbb{R}$, so DLO is not $|\mathbb{R}| = 2^{\aleph_0}$-categorical. In fact, not κ-categorical for any $\kappa \geq 2^{\aleph_0}$.

We now show that DLO is actually countably categorical.

Theorem 3.5.1. The theory DLO is countably categorical and hence complete.

Proof. Let $(A, \leq)$ and $(B, \leq)$ be two countable DLO's, and let $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ be a listing of A and B, respectively. We build an isomorphism $f \colon A \to B$ stage-by-stage: at stage i, we have

- finite sets $A_i \subseteq A$ and $B_i \subseteq B$, and
- a bijection $f_i \colon A_i \to B_i$ called a *partial embedding*:[a] if $a < a' \in A_i$, then $f_i(a) < f_i(a')$.

In this way, $f_i \subseteq f_{i+1}$, $A_i \subseteq A_{i+1}$, and $B_i \subseteq B_{i+1}$, and we need to make sure that

- $\bigcup_i A_i = A$, i.e., each element of A is in the domain of f (ensured by odd stages);
- $\bigcup_i B_i = B$, i.e., each element of B is in the range of f (ensured by even stages),

so $f = \bigcup_i f_i$ is a bijection from $A \to B$. Then sine each f_i is a partial embedding, so f is an $\mathcal{L}$-embedding, hence an isomorphism. This will prove that DLO is countably categorical.

The construction works as follows.

- Stage 0: $A_0 = \varnothing$, $B_0 = \varnothing$, $f_0 = \varnothing$.
- Stage $i+1 = 2k+1$: the goal is to make sure $a_k \in A_{i+1} = \mathrm{dom}(f_{i+1})$:
  - if $a_k \in A_i$ already, then do nothing, i.e., $A_{i+1} = A_i, B_{i+1} = B_i, f_{i+1} = f_i$;
  - otherwise, $a_k \notin A_i$, define $f_{i+1} \supseteq f_i$ by adding a_k to $A_{i+1} = A_i \cup \{a_k\}$, and for elements $a \in A_i$, $f_{i+1}(a) = f_i(a)$. Now, we have three possibilities:
    * a_k is less than all of A_i: choose $b \in B$ less than all of B_i;[b]
    * a_k is greater than all of A_i: similar to above;
    * there are a and a' in A_i such that $a < a_k < a'$ with no other elements of A_i between a and a' since A_i is finite: pick b with $f_i(a) < b < f_i(a')$.[c]

    In all cases, we can choose b and let $B_{i+1} = B_i \cup \{b\}$ with $f_{i+1}(a_k) = b$.
- Stage $i+1 = 2k+2$: the goal is to make sure $b_k \in B_{i+1} = \mathrm{Im}(f_{i+1})$: this is exactly the same, but in the other direction (e.g., working with f_i^{-1} rather than f_i).

Now everything is checked, so DLO is countably categorical (hence complete by Vaught test). ■

[a] If f_i maps $\overline{a}$ to $\overline{b}$, then $\overline{a}$ satisfies the same atomic and negated atomic formulas in $(A, \leq)$ that $\overline{b}$ does in $(B, \leq)$.
[b] Such b exists since B_i is finite, and $(B, \leq)$ has no left endpoint.
[c] Such b exists since B_i is finite and $(B, \leq)$ is dense.

Note (Back-and-forth). We see that the above is the so-called *back-and-forth* argument.

Corollary 3.5.1. $\mathbb{Q} + \mathbb{R} \equiv \mathbb{R}$.

Proof. Since $\mathrm{Th}(\mathbb{Q} + \mathbb{R}) = \mathrm{Th}(\mathbb{R}) = \{\varphi \mid \mathrm{DLO} \models \varphi\}$. ■

Definition 3.5.2 (Complete). A linear order is *complete* if every subset bounded above has a least upper bound.

Corollary 3.5.2. There is no first order sentence φ such that $\mathcal{M} \models \varphi$ if and only if $\mathcal{M}$ is a complete linear order.

3.5.2 Random Graph Theory

Another example of a countably categorical theory is the theory of random graph.

Definition 3.5.3 (Random graph). A *random graph* we will consider is constructed as follows. Firstly, fix countably infinitely many vertices $v_1, v_2, \ldots$, and fix p such that $0 < p < 1$. For each pair of vertices, "flip a coin": with probability p, put an edge; with $1-p$, no edge.

Now, the question is, what graph do we get? It turns out to be interesting enough, so we will look into it:

Remark. With probability 1, we get the same graph up to isomorphism, not matter what p is.

Then, consider the following theory.

Definition 3.5.4 (Random graph theory). Let $\mathcal{L} = \{E\}$, where E is a binary relation. The *random graph theory* T has axioms:

(a) $\forall x\ \neg xEx$ and $\forall x \forall y\ (xEy \to yEx)$ (irreflexive, undirected);

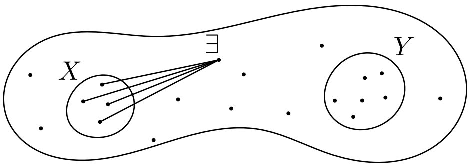
(b) $\exists x \exists y\ x \neq y$;

(c) for each n, define ψ_n as

$$\psi_n \coloneqq \forall x_1 \ldots \forall x_n \forall y_1 \ldots \forall y_n \left[\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{n} x_i \neq y_i \to \exists z \left(\bigwedge_{i=1}^{n} x_i E z \wedge \neg y_i E z \wedge z \neq x_i \wedge z \neq y_i \right) \right]$$

Intuition (Extension axiom). Think of ψ_n as an *extension axiom*: the property that for any finite disjoint sets X and Y, there is a vertex with an edge to each $x \in X$ and no edge to each $y \in Y$. This axiom happens with probability $p^{|X|} \cdot (1-p)^{|Y|}$ for a given z for some X, Y.

We see that a model of T is a graph with at least two elements with the extension property.



Note. In ψ_n's, we allow there to be repetitions among the x's and the y's. In particular, if $m \leq n$, then $\psi_n \models \psi_m$.

Lecture 16: Quantifier Elimination

Remark. T is actually countably categorical. 7 Mar. 14:30

Proof. We will show this on homework. But this is basically Theorem 3.5.1. ⊛

Now we show that it is also satisfiable. Fix N, $p \in (0,1)$, and vertices $\{1, \ldots, N\}$. Generate a random graph G_N on vertices $\{1, \ldots, N\}$ by, for each $i, j \leq N$, $i \neq j$, putting an edge between i and j with probability p and no edge with probability $1-p$. Let $\mathcal{G}_N$ be the class of all such graphs, and let $p_N(\varphi)$ be the probability that a randomly generated graph on N vertices satisfies φ.

Example. If $p = 1/2$, all graphs in $\mathcal{G}_N$ are equally likely.

Observe the following.

Proposition 3.5.1. For every n, then $\lim_{N\to\infty} p_N(\psi_n) = 1$.

Proof. Fix n, and let G be a random graph in $\mathcal{G}_N$, $N > 2n$. Fix $x_1, \ldots, x_n$, and $y_1, \ldots, y_n$, and z in G, all distinct. Consider the probability q that

$$\neg \left(\bigwedge_{i=1}^{n} E(x_i, z) \wedge \neg E(y_i, z) \right),$$

which is just $1 - p^n(1-p)^n \coloneqq q$ as we already know. The probability that

$$G \models \underbrace{\neg \exists z}_{\forall z\ \neg} \left(\bigwedge_{i=1}^{n} E(x_i, z) \wedge \neg E(y_i, z) \right)$$

is q^{N-2n}, where $N - 2n$ is the number of possible z's.[a] Let M be the number of different choices of x_i's and y_i's, then we have

$$p_N(\neg\psi_n) \leq M \cdot q^{N-2n} \leq N^{2n} q^{N-2n}$$

by the union bound.[b] Since $0 < q < 1$, so $p_N(\neg\psi_n) \to 0$, i.e., $p_N(\psi_n) \to 1$ as $N \to \infty$. ■

[a]Notice that we already assume that $x_i \neq z \wedge y_i \neq z$.
[b]This bound is not allowing the x's and y's to have repetitions. But this doesn't change anything.

Then, we have the following.

Theorem 3.5.2. T is satisfiable.

Proof. From Proposition 3.5.1, $\lim_{N\to\infty} p_N(\psi_n) = 1$. In particular, for each n, there is N such that $p_N(\psi_n) > 0$, i.e., there is at least one $G \in \mathcal{G}_N$ such that $G \models \psi_n$, hence $G \models \psi_m$ for $m \leq n$.[a] This means that for every finite $T^* \subseteq T$ is satisfiable, so T is satisfiable by the compactness theorem. ■

[a]Remember that for $n > m$, $\models \psi_n \to \psi_m$.

Moreover, we have the following.

Theorem 3.5.3 (Zero-one law for graphs)**.** For any $\mathcal{L}$-sentence φ, we either have $\lim\limits_{N\to\infty} p_N(\varphi) = 0$ or $\lim\limits_{N\to\infty} p_N(\varphi) = 1$. Moreover, T axiomatizes $\{\varphi \colon \lim_{N\to\infty} p_N(\varphi) = 1\}$, the "almost sure theory for graphs", which is decidable and complete.

Proof. Since T is countably categorical, and has only infinite models, so it is complete by the Vaught's test. If $T \models \varphi$, there is some n such that $\{\psi_n, (a), (b)\} \models \varphi$. Since $p_N(\psi_n) \leq p_N(\varphi)$, so $\lim_{N\to\infty} p_N(\varphi) = 1$ and the left-hand side goes to 1 as well. On the other hand, if $T \not\models \varphi$, $T \models \neg\varphi$, so $\lim_{N\to\infty} p_N(\neg\varphi) = 1$, i.e., $\lim_{N\to\infty} p_N(\varphi) = 0$ from the same argument. ■

Before ending this section, we note that there are other things we may explore.

Remark. There are also Fraïssé constructions, and Ryll-Nardzewski theorem, etc.

Chapter 4

Quantifier Elimination and Algebraic Applications

4.1 Quantifier Elimination

Let's start with a definition.

Definition 4.1.1 (Quantifier elimination)**.** A theory T has (admits) *quantifier elimination* if for every formula $\varphi(\overline{x})$ (with $\overline{x}$ containing at least one variable), there is a quantifier-free $\psi(\overline{x})$ such that

$$T \models \forall \overline{x}\ (\varphi(\overline{x}) \leftrightarrow \psi(\overline{x})).$$

The easiest example will be that DLO admits quantifier elimination. To prove this, we need some preliminaries.

Note. DLO has no constant symbols, so it has no quantifier-free sentences. For example,

$$\forall x \exists y\ y < x \leftrightarrow x = x,$$

where the right-hand side has a free variable. Another solution would be to allow $\top$ and $\bot$ for the true and false sentences.

Lemma 4.1.1. Let $(A, \leq)$ and $(B, \leq)$ be countable DLOs. Let $a_1, \ldots, a_n \in A$ have $a_1 < a_2 < \cdots < a_n$ and $b_1, \ldots, b_n \in B$ have $b_1 < \cdots < b_n$. Then there is an isomorphism $f\colon A \to B$ which maps $a_i \mapsto b_i$. Hence,

$$A \models \varphi(\overline{a}) \Leftrightarrow B \models \varphi(\overline{b}).$$

Proof. This is the same as the back-and-forth argument, but now we start with $\overline{a} \mapsto \overline{b}$. ■

Theorem 4.1.1. DLO admits quantifier elimination.

Proof. Fix $\varphi(\overline{x})$. If φ is actually a sentence then we're done since either

- DLO $\models \varphi$, so DLO $\models \forall x\ (\varphi \leftrightarrow x = x)$, or
- DLO $\models \neg\varphi$, so DLO $\models \forall x\ (\varphi \leftrightarrow x \neq x)$.

Now suppose $\varphi(\overline{x})$ has at least one free variable, $\overline{x} = (x_1, \ldots, x_n)$. Since DLO is complete, it's enough to find $\psi(\overline{x})$ with $\mathbb{Q} \models \forall \overline{x}\ (\varphi(\overline{x}) \leftrightarrow \psi(\overline{x}))$. Let σ be a map from pairs i, j to $\{1, 2, 3\}$. Define[a]

$$\theta_\sigma(x_1, \ldots, x_n) \coloneqq \bigwedge_{\sigma(i,j)=1} x_i < x_j \wedge \bigwedge_{\sigma(i,j)=2} x_i = x_j \wedge \bigwedge_{\sigma(i,j)=3} x_j < x_i.$$

If $\mathbb{Q} \models \theta_\sigma(\overline{a}) \wedge \theta_\sigma(\overline{b})$, then $\overline{a}$ and $\overline{b}$ satisfy the same formulas by Lemma 4.1.1, so[b]

$$\Sigma := \{\sigma \mid \mathbb{Q} \models \exists \overline{x}\ (\theta_\sigma(\overline{x}) \wedge \varphi(\overline{x}))\} = \{\sigma \mid \mathbb{Q} \models \forall \overline{x}\ (\theta_\sigma(\overline{x}) \to \varphi(\overline{x}))\}.$$

If $\Sigma = \varnothing$, then $\varphi(x) \leftrightarrow x \neq x$; if $\Sigma \neq \varnothing$, let $\psi(\overline{x}) = \bigvee_{\sigma \in \Sigma} \theta_\sigma(\overline{x})$, then $\mathbb{Q} \models \forall \overline{x}\ (\varphi(\overline{x}) \leftrightarrow \psi(\overline{x}))$. ■

[a] It's clear that some θ_σ might be inconsistent.
[b] The second equality follows from the fact that if one such $\overline{x}$ exists, all such $\overline{x}$ work.

Lecture 17: Quantifier Elimination

9 Mar. 14:30

Example. $\varphi(x, y) := \exists u\ (u > x \wedge u < y)$. This is equivalent, in DLO, to $\psi(x, y) := x < y$.

Example. In Problem Set 2, we have looked at $\mathcal{L} = \varnothing$. There, we showed that if $A \subseteq B$ and both infinite, then $A \preceq B$. The same idea show that $T = \left\{ \exists x_1 \ldots \exists x_n \bigwedge_{i \neq j} x_i \neq x_j \mid n \in \mathbb{N} \right\}$ admits quantifier elimination.

Proposition 4.1.1. Let T be a theory that admits quantifier elimination. If $\mathcal{A}, \mathcal{B} \models T$ and $\mathcal{A} \subseteq \mathcal{B}$, then $\mathcal{A} \preceq \mathcal{B}$.

Proof. Given $\phi(\overline{x})$ and $\overline{a} \in A$, then $\phi(x)$ is equivalent to a quantifier-free $\psi(\overline{x})$ (modulo T). Then,

$$\mathcal{A} \models \phi(\overline{a}) \Leftrightarrow \mathcal{A} \models \psi(\overline{a}) \Leftrightarrow \mathcal{B} \models \psi(\overline{a}) \Leftrightarrow \mathcal{B} \models \phi(\overline{a}).$$

■

Next, we want to show that ACF admits quantifier elimination. First, we need a test for quantifier elimination.

Theorem 4.1.2. Let $\mathcal{L}$ include at least one constant symbol c. Let T be an $\mathcal{L}$-theory, and $\phi(\overline{x})$ an $\mathcal{L}$-formula. Then the following are equivalent.

(a) There is a quantifier-free $\psi(\overline{x})$ such that $T \models \forall \overline{x}\ (\phi(\overline{x}) \leftrightarrow \psi(\overline{x}))$.

(b) If $\mathcal{M}, \mathcal{N} \models T$, and $\mathcal{A}$ is a common substructure, then for all $\overline{a} \in A$, $\mathcal{M} \models \phi(\overline{a})$ if and only if $\mathcal{N} \models \phi(\overline{a})$.

$$\begin{tikzcd} \mathcal{M} \models T & & \mathcal{N} \models T \\ & \mathcal{A} \ni \overline{a} \arrow[ul, hook] \arrow[ur, hook] & \end{tikzcd}$$

Proof. (a) implies (b) is easy, since we have

$$\mathcal{M} \models \phi(\overline{a}) \Leftrightarrow \mathcal{M} \models \psi(\overline{a}) \Leftrightarrow \mathcal{A} \models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \models \phi(\overline{a}).$$

To show (b) implies (a), we see that first, there are two easy cases:

- $T \models \forall \overline{x}\ \phi(\overline{x})$: take ψ to be $c = c$;
- $T \models \forall \overline{x}\ \neg\phi(\overline{x})$: take ψ to be $c \neq c$.

Now, suppose we are not in the above two cases. Let $\Gamma(\overline{x})$ be the set of all quantifier-free formulas $\psi(\overline{x})$ such that $T \models \forall \overline{x}\ (\phi(\overline{x}) \to \psi(\overline{x}))$, and let $\overline{d} = (d_1, \ldots, d_n)$ be new constant symbols.

Claim. It's enough to show $T \cup \Gamma(\overline{d}) \models \phi(\overline{d})$.

Proof. If we can do this, then by compactness, there are $\psi_1(\overline{x}), \ldots, \psi_m(\overline{x})$ such that

$$T \models \left[\psi_1(\overline{d}) \land \cdots \land \psi_m(\overline{d})\right] \to \phi(\overline{d})$$

By the choice of Γ, the $\leftarrow$ direction also holds, hence

$$T \models \left[\psi_1(\overline{d}) \land \cdots \land \psi_m(\overline{d})\right] \leftrightarrow \phi(\overline{d}),$$

i.e., $T \models \forall \overline{x}\ (\psi(\overline{x}) \leftrightarrow \phi(\overline{d}))$ where $\psi(x) \coloneqq \psi_1(\overline{x}) \land \cdots \land \psi_m(\overline{x})$. ⊛

Claim. $T \cup \Gamma(\overline{d}) \models \phi(\overline{d})$.

Proof. Suppose not. Then $T \cup \Gamma(\overline{d}) \cup \left\{\neg\phi(\overline{d})\right\}$ is satisfiable.[a] Let $\mathcal{M}$ be a model of $T \cup \Gamma(\overline{d}) \cup \left\{\neg\phi(\overline{d})\right\}$, and let $\mathcal{A} \subseteq \mathcal{M}$ be the "submodel"[b] generated by $\overline{d}$. Notice that since $c^{\mathcal{A}} \in \mathcal{A}$, so $\mathcal{A} \neq \varnothing$. Every element of $\mathcal{A}$ is $t^{\mathcal{A}}(\overline{d}) = t^{\mathcal{M}}(\overline{d})$ for some term t. Also, $\mathcal{A} \models \Gamma(\overline{d})$ because $\mathcal{M} \models \Gamma(\overline{d})$ and $\Gamma(\overline{d})$ is quantifier-free. We now show that $T \cup \mathrm{Diag}(\mathcal{A}) \cup \left\{\phi(\overline{d})\right\}$ is satisfiable, and take $\mathcal{N}$ to be a model. If not, some finite subset is not satisfiable. There are quantifier-free formulas $\gamma_1(\overline{d}), \ldots, \gamma_\ell(\overline{d})$ in $\mathrm{Diag}(\mathcal{A})$ such that[c]

$$T \models \left[\gamma_1(\overline{d}) \land \cdots \land \gamma_\ell(\overline{d})\right] \to \neg\phi(\overline{d}).$$

Equivalently, $T \models \phi(\overline{d}) \to \neg\left[\gamma_1(\overline{d}) \land \cdots \land \gamma_\ell(\overline{d})\right]$, so $T \models \forall \overline{x}\ (\phi(\overline{x}) \to [\neg\gamma_1(\overline{x}) \lor \cdots \lor \neg\gamma_\ell(\overline{x})])$. This implies that $\neg\gamma_1(\overline{x}) \lor \cdots \lor \neg\gamma_\ell(\overline{x})$ is in $\Gamma(\overline{x})$, hence $\mathcal{A} \models \neg\gamma_1(\overline{d}) \lor \cdots \lor \neg\gamma_\ell(\overline{d})$. But $\mathcal{A} \models \gamma_1(\overline{d}) \land \cdots \land \gamma_\ell(\overline{d})$, contradicts to the fact that $\gamma_1(\overline{d}), \ldots, \gamma_\ell(\overline{d})$ is in $\mathrm{Diag}(\mathcal{A})$.

Hence, there is $\mathcal{N} \models T \cup \mathrm{Diag}(\mathcal{A}) \cup \left\{\phi(\overline{d})\right\}$. This contradicts to (b):

$$\begin{tikzcd}
\mathcal{M} \models \neg\phi(\overline{d}) \text{ and } \mathcal{M} \models T & & \mathcal{N} \models \neg\phi(\overline{d}) \text{ and } \mathcal{N} \models T \\
& \mathcal{A} \ni \overline{d} \arrow[ul, "\subseteq"'] \arrow[ur, hook] &
\end{tikzcd}$$

Hence, $T \cup \Gamma(\overline{d}) \models \phi(\overline{d})$. ⊛

[a] Notice that $T \cup \Gamma(\overline{d})$ needs to be satisfiable, which is true from our assumption.
[b] $\mathcal{A}$ is the smallest model containing $\overline{d}$, and $\overline{a}$ might not be a model of T.
[c] We can take $\gamma_i(\overline{d})$ to be just about $\overline{d}$ but not $\gamma_i(\overline{a})$ for $\overline{a} \in A$ is because each $\overline{a} \in A$ is $t^{\mathcal{A}}(\overline{d})$, so we can replace $\overline{a}$ by this term.

■

We now show that we can eliminate one $\exists$ at a time.

Lemma 4.1.2. Let T be an $\mathcal{L}$-theory. Suppose that for every quantifier-free formula $\gamma(\overline{x}, y)$, there is a quantifier-free formula $\psi(\overline{x})$ such that

$$T \models \forall \overline{x}\ (\exists y\ \gamma(\overline{x}, y) \leftrightarrow \psi(\overline{x})),$$

then T has quantifier elimination.

Proof. By induction on formulas with the hypotheses of Theorem 4.1.2 for the $\exists$ quantifier case. ■

By putting together the last two results, i.e., Lemma 4.1.2 and Theorem 4.1.2, we get the following test for quantifier elimination. Compare this, for example, to Vaught's test for elementary-substructure.

Corollary 4.1.1. Let T be an $\mathcal{L}$-theory with at least one constant. Suppose that for all quantifier-free $\psi(\overline{x}, y)$, if $\mathcal{M}, \mathcal{N} \models T$, and $\mathcal{A} \subseteq \mathcal{M}$ and $\mathcal{A} \subseteq \mathcal{N}$, $\overline{a} \in A$ and $b \in M$[a] is such that $\mathcal{M} \models \psi(\overline{a}, b)$, then there is $c \in N$ such that $\mathcal{N} \vDash \psi(\overline{a}, c)$. Then T admits quantifier elimination.

$$\begin{tikzcd}
\mathcal{M} \ni b \arrow[rr, dashed, bend left] & & \mathcal{N} \ni c \\
& \mathcal{A} \ni \overline{a} \arrow[ul, hook] \arrow[ur, hook] &
\end{tikzcd}$$

[a]This is (b) for $\exists y\ \psi(\overline{x}, y)$.

4.1.1 Quantifier Elimination for ACF

Now, we start with our favorite ACF, and show that it admits quantifier elimination.

Theorem 4.1.3. ACF admits quantifier elimination.

Proof. We use Corollary 4.1.1. Suppose that $\mathcal{M}, \mathcal{N} \vDash$ ACF, and $\mathcal{A} \subseteq \mathcal{M}, \mathcal{N}$ an integral domain. Let $\psi(\overline{x}, y)$ be a quantifier-free formula and $\overline{a} \in \mathcal{A}$. Suppose there is $b \in \mathcal{M}$ such that $\mathcal{M} \vDash \psi(\overline{a}, b)$.

Firstly, we replace $\mathcal{N}$ by the algebraic closure of $\mathcal{A}$, then we can also assume that $\mathcal{N} \subseteq \mathcal{M}$ since any algebraic closure of $\mathcal{A}$ embeds in any algebraically closed field containing $\mathcal{A}$.

$$\begin{tikzcd}
\mathcal{M} \ni b & & \mathcal{N} \arrow[ll, dashed, hook'] \\
& \mathcal{A} \ni \overline{a} \arrow[ul, hook] \arrow[ur, hook] &
\end{tikzcd}$$

Problem. What does ψ look like?

Answer. Since it's quantifier-free, we may assume that it is $\theta_1(\overline{x}, y) \vee \theta_2(\overline{x}, y) \vee \ldots$, where each θ_i is a conjunction of atomic and negated atomic formula. We can replace ψ by whichever θ_i is satisfied by b. ⊛

Now, since in the language of rings, an atomic formula $\gamma(\overline{x}, y)$ is equivalent to a formula of the form $p(\overline{x}, y) = 0$, where $p \in \mathbb{Z}[\overline{X}, Y]$, so $\psi(\overline{x}, y)$ is equal to

$$p_1(\overline{x}, y) = 0 \wedge \cdots \wedge p_k(\overline{x}, y) = 0 \wedge q_1(\overline{x}, y) \neq 0 \wedge q_\ell(\overline{x}, y) \neq 0$$

for $p_i, q_i \in \mathbb{Z}[\overline{X}, Y]$. Then $\psi(\overline{a}, y)$ says $p_1(\overline{a}, y) = 0 \wedge \ldots$ where $p_i(\overline{a}, y)$ are now in $\mathcal{A}[y]$, i.e., polynomial in y with coefficient in $\mathcal{A}$. If any $p_i(\overline{a}, y)$ is non-trivial, then b is a solution, so $b \in \mathcal{N}$ since $\mathcal{N}$ is algebraically closed (and we just take $c = b$ for applying Corollary 4.1.1). Otherwise, assume $p_i(\overline{a}, y)$ is trivial, so $\psi(\overline{a}, y)$ is just

$$q_i(\overline{a}, y) \neq 0 \wedge \cdots \wedge q_\ell(\overline{a}, y) \neq 0.$$

Since b satisfies this, each $q_i(\overline{a}, y)$ is non-trivial, so $q_i(\overline{a}, y) = 0$ has only finitely many solutions. But $\mathcal{N}$ is infinite, so there is a $c \in \mathcal{N}$ that is not a solution to any $q_i(\overline{a}, y) = 0$, so $\mathcal{N} \vDash \psi(\overline{a}, c)$. ■

Lecture 18: Quantifier Elimination for Algebraically Closed Fields

Problem 4.1.1. What do we get from quantifier elimination? 14 Mar. 14:30

Answer. Understand the definable sets, they are defined by quantifier-free formulas. ⊛

Remark. We see that

$$\varphi(\overline{x}) \coloneqq \exists y\ p_n(\overline{x}) y^n + \cdots + p_1(\overline{x}) y + p_0(\overline{x}) = 0$$

is equivalent to

$$\psi(\overline{x}) \coloneqq p_n(\overline{x}) \neq 0 \vee \cdots \vee p_1(\overline{x}) \neq 0 \vee p_0(\overline{x}) = 0.$$

Definition 4.1.2 (Cofinite). A *cofinite* subset of a set X is a subset A such that $|A^c| < \infty$.

Proposition 4.1.2. The definable[a] subsets of an algebraically closed field K are exactly the finite and cofinite sets.

[a]Using parameters.

Proof. Let $\{a_1, \ldots, a_n\}$ be a finite set. This is definable by $x = a_1 \vee \cdots \vee x = a_n$ using $a_1, \ldots, a_n$ as parameters; also, the complement of $\{a_1, \ldots, a_n\}$ is definable by $x \neq a_1 \wedge \cdots \wedge x \neq a_n$.[a]

On the other hand, let $X = \{x \in K \mid K \models \varphi(x, \overline{a})\}$ be a definable subset of K. By quantifier elimination, we may assume that φ is quantifier-free, so φ is a boolean combination of atomic and negated atomic formulas. Notice that an atomic formula is of the form

$$p_n(\overline{a})x^n + \cdots + p_1(\overline{a})x + p_0(\overline{a}) = 0,$$

hence this atomic formula defines either a finite set or all of K. A negated atomic formula defines a cofinite set or $\varnothing$. Boolean combinations of finite and cofinite sets are finite or cofinite, so X is finite or cofinite. ■

[a]We did not use anything about fields here.

Remark. Proposition 4.1.2 is not true for $X \subseteq K^2$.

Proof. $X = \{(x, y) \mid x^2 + y^2 + 1\}$. ⊛

4.2 Definable and Constructible Sets

Proposition 4.1.2 shows some desirable properties, so we come up with the following definition.

Definition 4.2.1 (Strongly minimal). A theory T is *strongly minimal* if for any $\mathcal{M} \models T$, and $X \subseteq M$ definable, X is either finite or cofinite.

Now let us consider definable sets of higher arities.

Definition 4.2.2 (Algebraic). Let K be a field, and $X \subseteq K^n$. We say that X is *algebraic* if there is a set S of polynomials over K such that X is the zero set of S.

As previously seen. Recall Definition 3.2.15 and compared it to the above.

Example. $X = \{(x, y) \mid x^2 + y^2 = 1\}$ is algebraic since $S = \{x^2 + y^2 - 1\}$.

The complement of an algebraic set is usually not algebraic.

Definition 4.2.3 (Constructible). The *constructible* sets are the boolean combinations of algebraic sets.

Remark. The constructible sets are exactly the definable sets in $K \models \mathrm{ACF}$.

Proof. The definable sets, by quantifier elimination, boolean combinations of sets defined by atomic formulas, which are algebraic. So definable implies constructible.

On the other hand, it is enough to see that algebraic set are definable. The issue is that $S \subseteq K[\overline{X}]$ might be infinite. Let I be the ideal generated by S. Then the set $X \subseteq K^n$ of common zeros of S is also the set of common zeros of I.[a] By the Hilbert's basis theorem, $I = (f_1, \ldots, f_m)$ is finitely generated. Hence,

$$X = \{\overline{a} \in K^n \mid f_1(\overline{a}) = 0 \wedge \cdots \wedge f_m(\overline{a}) = 0\},$$

hence it's definable. ⊛

[a]Each $f \in I$ is $f = r_1 g_1 + \cdots + r_n g_n$, where $g_1, \ldots, g_n \in S$.

Theorem 4.2.1 (Chevalley's theorem)**.** Let K be an algebraically closed field and $X \subseteq K^n$ be constructible. Let $p \colon K^n \to K^m$ be a polynomial map, i.e., $p(\overline{x}) = (q_1(\overline{x}), \ldots, q_n(\overline{x}))$ for polynomials q_i. Then, $p(X)$, the image of X under p, is also constructible.

Proof. Since we know that constructible is the same as definable, so consider

$$p(X) = \{\overline{y} \mid \exists \overline{x}\ (\overline{x} \in X \wedge p(\overline{x}) = \overline{y})\}$$

where for $\overline{x} \in X$, there is a formula expressing this, and for $p(\overline{x}) = \overline{y}$, $y_1 = q_1(\overline{x}) \wedge \cdots \wedge y_m = q_m(\overline{x})$. Hence, $p(X)$ is definable (hence constructible), since X was. ■

Example. $p(x_1, x_2, x_3) = (x_1, x_3)$.

Theorem 4.2.2 (Weak Hilbert's Nullstellensatz)**.** Let K be algebraically closed, and $f_1, \ldots, f_n \in K[\overline{x}]$. Then there is $\overline{a} \in K^m$ sch that $f_1(\overline{x}) = \cdots = f_n(\overline{x}) = 0$ if and only if $1 \notin (f_1, \ldots, f_n)$, i.e., there are no $r_1, \ldots, r_n \in K[\overline{x}]$ such that $1 = r_1 f_1 + \cdots + r_n f_n$.

Proof. If $1 \in (f_1, \ldots, f_n)$, there are $r_1, \ldots, r_n \in K[\overline{x}]$ such that $1 = r_1 f_1 + \cdots + r_n f_n$. If $\overline{a}$ was a common zero of $f_1, \ldots, f_n$, then

$$1 = r_1(\overline{a})\cancel{f_1(\overline{a})} + \cdots + r_n(\overline{a})\cancel{f_n(\overline{a})} = 0,$$

so no such $\overline{a}$ exists.

Now, suppose $1 \notin (f_1, \ldots, f_n)$, so $(f_1, \ldots, f_n) \neq K[\overline{x}]$. Let I be a maximal ideal containing $f_1, \ldots, f_n$, and let $L = K[\overline{x}] \,/\, I$, which is a field extension of K, $K \hookrightarrow L$. There is $\overline{a} \in L^m$ which is a common root of $f_1, \ldots, f_n$, namely $a_i = x_i + I$. Let M be the algebraic closure of L, with $\overline{a} \in M^n$. By quantifier elimination, $K \succeq M$. $M \models \exists \overline{y}\ f_1(\overline{y}) = 0 \wedge \cdots \wedge f_n(\overline{y}) = 0$, which is a formula about elements of K (the coefficients). Because $K \succeq M$, $K \models \exists \overline{y}\ f_1(\overline{y}) = 0 \wedge \cdots \wedge f_n(\overline{y}) = 0$, which says that $f_1, \ldots, f_n$ have a common zero of K. ■

Note. We leave the full version in the note, which relates to algebraic geometry (if you care).

Remark. The weak Hilbert's Nullstellensatz says that whether $1 \in (f_1, \ldots, f_n)$ is the only barrier for $f_1, \ldots, f_n$ having a common zero.

4.3 Algebraic Closure

As previously seen (Definable closure)**.** The *definable closure* $\operatorname{dcl}(A)$ of A is the set of all $b \in M$ which are definable over A.

On the homework, we looked at dcl, which is not really useful. In strongly minimal theories, we look at acl instead, which is very important.

Definition 4.3.1 ((Model theoretical) algebraic closure)**.** Let $\mathcal{M}$ be a structure and $A \subseteq M$. Then the *(model-theoretic) algebraic closure* $\operatorname{acl}(A)$ of A is the set of all $a \in M$ such that there are $\overline{b} \in A$ and a formula $\varphi(x, \overline{b})$ such that $\mathcal{M} \models \varphi(a, \overline{b})$ and there are only finitely many other a' with $\mathcal{M} \models \varphi(a', \overline{b})$.

Note. There are some properties that acl always satisfies.

- $\mathrm{dcl}(A) \subseteq \mathrm{acl}(A)$.
- $A \subseteq \mathrm{acl}(A)$.
- If $A \subseteq B$, then $\mathrm{acl}(A) \subseteq \mathrm{acl}(B)$.
- $\mathrm{acl}(A) = \mathrm{acl}(\mathrm{acl}(A))$.[a]
- If $a \in \mathrm{acl}(A)$, then there is a finite $F \subseteq A$ such that $a \in \mathrm{acl}(F)$.

[a]This is a good exercise!

On homework 5, we will show that in models of a strongly minimal theory, acl also satisfies the "exchange property":

Remark (Exchange property). If $a \in \mathrm{acl}(X \cup \{b\})$ and $a \notin \mathrm{acl}(X)$, then $b \in \mathrm{acl}(X \cup \{a\})$. This makes acl a *pregeometry* or *matroid*.

In algebraically closed fields, $a \in \mathrm{acl}(X)$ just means that a is algebraic over (the field generated by) X , i.e., there is $p(y) \in F_X[\overline{x}]$ such that $p(a) = 0$. In vector spaces, $a \in \mathrm{acl}(X)$ if and only if $a \in \mathrm{span}(X)$.

Intuition. This makes model has notions like dimensions, independence, etc.

Lecture 19: acl in Algebraically Closed Fields

16 Mar. 14:30

Theorem 4.3.1. Let $K \models \mathrm{ACF}$, and $A \subseteq K$. The $\mathrm{acl}(A)$ is the set of all elements $b \in K$ which are algebraic over A.[a]

[a]I.e., that satisfy a non-zero polynomial with coefficients in the ring generated by A.

Proof. If $p(x)$ is a non-zero polynomial over A, with $p(b) = 0$, then $\{c \mid p(c) = 0\}$, so $p(x) = 0$ witnesses that $b \in \mathrm{acl}(A)$.

Suppose that $b \in \mathrm{acl}(A)$, as witnessed by $\varphi(x, \overline{a})$ for $\overline{a} \in A$. We may assume that φ is quantifier-free. Write

$$\varphi(x, \overline{a}) = \psi_1(x, \overline{a}) \vee \cdots \vee \psi_n(x, \overline{a}),$$

where each ψ_i is a conjunction of atomic or negated atomic formulas. We may replace φ with some ψ_i, choosing one that b satisfies. Then,

$$\varphi(x, \overline{a}) = p_1(x, \overline{a}) \wedge \cdots \wedge p_k(x, \overline{a}) = 0 \wedge q_1(x, \overline{a}) \neq 0 \wedge \cdots \wedge q_\ell(x, \overline{a}) \neq 0.$$

In order for this to have finitely many solutions, we must have $k \geq 1$ and some $p_i(x, \overline{a})$ non-zero polynomial in x. But then b satisfies a polynomial over A. ∎

4.4 Types

Now, unless we specify, we will now work in a model of a strongly minimal theory.

Definition 4.4.1 (Independent). A set X is *independent* if for all $a \in X$, $a \notin \mathrm{acl}(X - \{a\})$.

Intuition. Think about the case of vector spaces.

Definition 4.4.2 (Basis). A set X is a *basis* of Y if it's the maximal independent set in Y.

Note. In homework, we will show that each set Y always contains a basis.

Definition 4.4.3 (Dimension)**.** The *dimension* $\dim Y$ of Y is this cardinality of its basis.

Remark. If X_1, X_2 are two bases for Y, then $|X_1| = |X_2|$, i.e., Definition 4.4.3 is well-defined.

Definition 4.4.4 (Type)**.** Let $\mathcal{M}$ be an $\mathcal{L}$-structure, $\bar{c} \in M$, $A \subseteq M$. The *type* of $\bar{c}$ over A (in $\mathcal{M}$) is

$$\operatorname{tp}^{\mathcal{M}}(\bar{c}/A) = \{\varphi(\bar{x}, \bar{a}) \mid \mathcal{M} \models \varphi(\bar{c}, \bar{a}) \text{ and } \bar{a} \in A\}.$$

Notation. Where we omit $/A$, we just mean $/\varnothing$.

Lemma 4.4.1. Let T be a complete strongly minimal theory, and $\mathcal{M}, \mathcal{N} \models T$. Let $\bar{a} \in M$ and $\bar{b} \in N$ be independent tuples of the same size, then $\operatorname{tp}^{\mathcal{M}}(\bar{a}) = \operatorname{tp}^{\mathcal{N}}(\bar{b})$, i.e., $\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{N} \models \varphi(\bar{b})$.

Proof. We do an induction on the lengths of $\bar{a}, \bar{b}$. Start with $n = 1$: let a, b be independent, then

$$\operatorname{tp}(a) = \{\varphi(x) \mid \varphi(x) \text{ has cofinitely many solutions in } \mathcal{M}\}$$

since a is independent, $a \notin \operatorname{acl}(\varnothing)$, i.e., $\varphi(\mathcal{M}) \coloneqq \{c \in M \mid \mathcal{M} \models \varphi(c)\}$ is infinite for all φ. And since $\mathcal{M}$ is strongly minimal, it's cofinite.

Example. In ACF, $\operatorname{tp}(a) = \{p(a) \neq 0 \mid p \in \mathbb{Q}[x], p \neq 0\}$ (or "generated by" these). Also, $(1+1)x$ being 0 depends on the characteristic.

Similarly, we have

$$\operatorname{tp}(b) = \{\varphi(x) \mid \varphi(x) \text{ has cofinitely many solutions in } \mathcal{N}\}.$$

Now, suppose that φ has k non-solutions in $\mathcal{M}$, $\mathcal{M} \models \exists^{=k} x\ \neg\varphi(x)$, with T being complete, $T \models \exists^{=k} x\ \neg\varphi(x)$, hence $\mathcal{N} \models \exists^{=k} x\ \neg\varphi(x)$. So if φ has cofinitely many (all but k) solutions in $\mathcal{M}$, the same is true in $\mathcal{N}$. Hence, $\operatorname{tp}(a) = \operatorname{tp}(b)$.

Example. The completeness is important: ACF is a non-example.

For $n+1$, let $\bar{a}a'$ and $\bar{b}b'$ be independent $(n+1)$-tuples with $\operatorname{tp}(\bar{a}) = \operatorname{tp}(\bar{b})$. Suppose $\mathcal{M} \models \varphi(\bar{a}, a')$. Since $\mathcal{M} \models T$ is strongly minimal and $a' \notin \operatorname{acl}(\bar{a})$, $\varphi(\bar{a}, \mathcal{M}) = \{c \in M \mid \mathcal{M} \models \varphi(\bar{a}, c)\}$ is cofinite with complement of size k. Then $\mathcal{M} \models \exists^{=k} x\ \neg\varphi(\bar{a}, x)$, so $\exists^{=k} x\ \neg\varphi(\bar{y}, x) \in \operatorname{tp}(\bar{a}) = \operatorname{tp}(\bar{b})$, i.e., $\mathcal{N} \models \exists^{=k} x\ \neg\varphi(\bar{b}, x)$. Then, since $b' \notin \operatorname{acl}(\bar{b})$, so $b' \in \varphi(\bar{b}, \mathcal{N})$,[a] hence $\mathcal{N} \models \varphi(\bar{b}, b')$. ■

[a] Since $\varphi(\bar{b}, \mathcal{N})$ is cofinite, and if b' is in the complement of $\varphi(\bar{b}, \mathcal{N})$, which is finite, $b' \in \operatorname{acl}(\bar{b})$ by $\neg\varphi(\bar{b}, x)$ ↯

Lemma 4.4.2. Let T be a strongly minimal theory. If $\mathcal{M} \models T$ and $c \in \operatorname{acl}(A)$, then there is a formula $\varphi(x, \bar{a}) \in \operatorname{tp}(c/A)$ such that for any other $c' \in M$ with $\mathcal{M} \models \varphi(c', \bar{a})$, $\operatorname{tp}(c/A) = \operatorname{tp}(c'/A)$.

Proof. Let $\varphi(x, \bar{a}) \in \operatorname{tp}(c/A)$ be such that

$$|\{x \in M \mid \mathcal{M} \models \varphi(x, \bar{a})\}| = k$$

is minimal. We claim that $\varphi(x, \bar{a})$ witnesses the statement. If not, there are some c' and a formula $\psi(x, \bar{a}') \in \operatorname{tp}(c/A)$ with $\mathcal{M} \models \varphi(c', \bar{a}) \wedge \neg\psi(c', \bar{a}')$. So

$$\{x \in M \mid \mathcal{M} \models \varphi(x, \bar{a}) \wedge \psi(x, \bar{a}')\} \subsetneq \{x \in M \mid \mathcal{M} \models \varphi(d, \bar{a})\}$$

since c' is in the right-hand side but not the left-hand side. This implies that the left-hand side has cardinality $< k$, a contradiction, so φ does imply ψ. ■

Theorem 4.4.1. Let T be a complete strongly minimal theory. If $\mathcal{M}, \mathcal{N} \models T$, then $\mathcal{M} \cong \mathcal{N}$ if and only if $\dim \mathcal{M} = \dim \mathcal{N}$.

Proof. Suppose $\dim \mathcal{M} = \dim \mathcal{N}$ with A, B being the bases for $\mathcal{M}, \mathcal{N}$. Then, we have $|A| = |B|$, so there exists a bijection $f\colon A \to B$, which is a partial elementary map since for f, $\overline{a} \in A$, $\operatorname{tp}(\overline{a}) = \operatorname{tp}(f(\overline{a}))$ from Lemma 4.4.1.

Definition 4.4.5 (Paritla elementary map)**.** A map $g\colon U \subseteq M \to N$ is a *partial elementary map* if $\mathcal{M} \models \varphi(\overline{a}) \Leftrightarrow \mathcal{N} \models \varphi(g(\overline{a}))$ for $\overline{a} \in \operatorname{dom} g$.[a]

[a]This implies injectivity since if $\overline{a} \neq \overline{a}'$ with $g(\overline{a}) = g(\overline{a}')$, there exists some φ differentiate them.

By Zorn's lemma, there exists a maximal partial elementary map g from $\mathcal{M} \to \mathcal{N}$ extending f.

Claim. g is an isomorphism, i.e., $\operatorname{dom} g = M$ ($\operatorname{Im} g = N$ is automatic).

Proof. Assume that $c \in M - \operatorname{dom} g$. We know that $c \in \operatorname{acl}(A) \subseteq \operatorname{acl}(\operatorname{dom} g)$ since A is a basis, so let $\varphi(x, \overline{d})$ be the formula from Lemma 4.4.2 that isolates $\operatorname{tp}(c / \operatorname{dom} g)$, i.e., whenever $\mathcal{M} \models \varphi(c', \overline{d})$, $\operatorname{tp}(c / \operatorname{dom} g) = \operatorname{tp}(c' / \operatorname{dom} g)$. Then, $\mathcal{M} \models \exists x\ \varphi(x, \overline{d})$, implying $\mathcal{N} \models \exists x\ \varphi(x, g(\overline{d}))$. Let $c' \in N$ witness this. Then $\operatorname{tp}^{\mathcal{M}}(c / \operatorname{dom} g) = \operatorname{tp}^{\mathcal{N}}(c' / \operatorname{Im} g)$ after identifying $\operatorname{dom} g$ and $\operatorname{Im} g$.[a] Then we can define $g(c) = c'$ to extend g but remain partial elementary, contradicting to the maximality of g, so $\operatorname{dom} g = M$.

Following the same argument, we can show that $\operatorname{Im} g = N$, hence g is an isomorphism. ⊛

[a]Suppose $\mathcal{M} \models \psi(c, \overline{b})$ for $\overline{b} \in \operatorname{dom} g$, then in $\mathcal{M}$, $\mathcal{M} \models \forall x\ \varphi(x, \overline{a}) \to \psi(x, \overline{b})$. Since g is elementary, $\mathcal{N} \models \forall x\ \varphi(x, g(\overline{a})) \to \psi(x, g(\overline{b}))$, so $\mathcal{N} \models \psi(c', \overline{b})$.

The other direction is trivial, since if $\mathcal{M} \cong \mathcal{N}$, then clearly $\dim \mathcal{M} = \dim \mathcal{N}$. ■

Lecture 20: Fraïssé Theorem

4.5 Other Examples of Quantifier Elimination

21 Mar. 14:30

As previously seen. ACF, DLO, and also the random graph theory.[a]

[a]This is very similar to the argument we gave for DLO.

We discuss some other examples of theories which admit quantifier elimination.

4.5.1 Vector Spaces

On the problem set, we will show that the theory of infinite (dimensional) vector spaces admits quantifier elimination and is strongly minimal.

4.5.2 Torsion-Free Abelian Groups

As previously seen (Torsion-free)**.** An Abelian group G is said to be *torsion-free* if no element other than e has finite order.

Definition 4.5.1 (Divisible)**.** A torsion-free group G is *divisible* if for every $n \in \mathbb{N}$ and $g \in G$, there is $h \in G$ such that

$$nh = \underbrace{h + \cdots + h}_{n \text{ times}} = g.$$

Since G is torsion-free, for each g, the corresponding h is unique.[1] The theory DAG of torsion-free divisible Abelian groups (i.e., $\mathbb{Q}$-vector spaces in language $\{+, 0, -\}$) admits quantifier elimination.

[1]As if $nh = g = nh'$, then $n(h - h') = 0$ so $h = h'$.

4.5.3 Ordered Torsion-Free Divisible Abelian Groups

On top of DAG, if we also add an ordering, we get ODAG, the theory of ordered torsion-free divisible Abelian groups, which also admits quantifier elimination.

4.5.4 Presburger Arithmetic

Presburger arithmetic is $\mathrm{Th}(\mathbb{N}, 0, 1, +)$.

4.5.5 Real Closed Fields

As previously seen (Real-closed field)**.** A *real-closed field* is a field F that has the same first-order properties as the field of real numbers.

The theory of real-closed fields, denoted as RCF, does not admit quantifier elimination in the language $\mathcal{L} = \{0, 1, +, -, \cdot\}$. This is because the formula

$$\varphi(x, y) = \exists x \ x - y = z^2$$

is not equivalent to a quantifier-free formula.

Intuition. This formula is defining the ordering of the field.

Claim. $\varphi(x, y) = \exists z \ x - y = z^2$ is not equivalent to a quantifier-free formula in $\mathcal{L} = \{0, 1, +, -, \cdot\}$.

Proof. The first way one might try to prove this is to show that there are real-closed fields $\mathcal{M} \subseteq \mathcal{N}$ and $a, b \in \mathcal{N}$ such that $\mathcal{N} \models \varphi(a, b)$ but $\mathcal{M} \models \neg\varphi(a, b)$. But this strategy will not work, since $\varphi(x, y)$ is also equivalent, in any real-closed field, to the universal formula $\forall z \ y - x \neq z^2$. This is because only one of $x - y$ and $y - z$ can have a square root in a real-closed field. Instead, one should think about the test for quantifier elimination.

Proposition 4.5.1. There is no quantifier-free formula $\psi(x, y)$ such that RCF proves that $\psi(x, y)$ and $\varphi(x, y)$ are equivalent.

⊛

But if we add in the ordering as a symbol in the language, then in the language $\mathcal{L}_<$, the theory $\mathrm{Th}(\mathbb{R}, 0, 1, +, -, \cdot, <)$ of real-closed ordered fields does admit quantifier elimination. This was shown by Tarski in the 1940s as part of showing that the theory of $\mathbb{R}$ is decidable.

Example. An example of eliminating quantifiers from a formula that we already know is

$$\mathrm{RCF} \models \exists x \ (a \neq 0 \land ax^2 + bx + c = 0) \leftrightarrow a \neq 0 \land b^2 - 4ac \geq 0.$$

After proving quantifier elimination, we can analyze the definable sets like we did for algebraically closed fields.

Remark (o-minimal)**.** The theory of real-closed fields is not strongly minimal, but instead what is called *o-minimal.*[a]

[a] There is a well-developed theory of o-minimality which would be an entire course in itself.

Proof. The definable sets in one variable are finite unions of intervals and points. ⊛

Remark. Tarski-Seidenberg algorithm can solve all Euclidean geometry problems.

Chapter 5

Fraïssé Limits

Our two examples of countably categorical theories, DLO and the random graph, are both characterized by an extension axiom.[1] This influences a general way of construction countably categorical structures.

5.1 Substructures' Properties

Definition 5.1.1 (Generated substructure)**.** Given an $\mathcal{L}$-structure $\mathcal{M}$, and a set $A \subseteq M$, we write $\langle A \rangle$ for the substructure of M *generated* by A, defined as the smallest substructure of $\mathcal{M}$ whose domain contains A.

Intuition. Equivalently, it's the substructure of $\mathcal{M}$ containing A and all the constants, and is closed under the application of functions.

Note. Compare Definition 5.1.1 to the substructure generated in the downward Löwenheim-Skolem theorem.

And it's natural to talk about finiteness.

Definition 5.1.2 (Finitely generated)**.** A substructure $\mathcal{N}$ of $\mathcal{M}$ is *finitely generated* if it is $\mathcal{N} = \langle A \rangle$ for some finite $A \subseteq M$.

Now, the question is how should we build structures? Idea is that to build "universal" countable structures with "all possible finitary behaviors". Assume (for now) $\mathcal{L}$ is relational.

Definition 5.1.3 (Age)**.** For an $\mathcal{L}$-structure $\mathcal{M}$, the *age* of $\mathcal{M}$, $\mathrm{Age}(\mathcal{M})$, is the class of all finite $\mathcal{L}$-substructure which extend into $\mathcal{M}$.

Definition 5.1.4 (Hereditary property)**.** A class $\mathbb{K}$ of finite $\mathcal{L}$-structure has the *hereditary property* if for all $\mathcal{B} \in \mathbb{K}$ and $\mathcal{A}$ which extends into $\mathcal{B}$, $\mathcal{A} \in \mathbb{K}$.

Intuition. If it's downward-closed under embedding.

Definition 5.1.5 (Joint embedding property)**.** A class $\mathbb{K}$ of finite $\mathcal{L}$-structure has the *joint embedding property* if for all $\mathcal{A}, \mathcal{B} \in \mathbb{K}$, there exists $\mathcal{C} \in \mathbb{K}$ with embeddings $\mathcal{A} \hookrightarrow \mathcal{C} \hookleftarrow \mathcal{B}$:

$$\begin{tikzcd} & \mathcal{C} & \\ \mathcal{A} \arrow[ur, dashed] & & \mathcal{B} \arrow[ul, dashed] \end{tikzcd}$$

[1] Though not explicit for DLO, but essentially given $a < b$, we can "extend" this order by finding c such that $a < c < b$.

5.2 “Baby” Fraïssé Theorem

Consider the following.

Theorem 5.2.1 (Baby Fraïssé theorem). A class $\mathbb{K}$ of finite $\mathcal{L}$-structure is $\mathrm{Age}(\mathcal{M})$ for some countable $\mathcal{M}$ if and only if

- $\mathbb{K}$ is countable up to isomorphism;
- $\mathbb{K} \neq \varnothing$;
- $\mathbb{K}$ has the hereditary property;
- $\mathbb{K}$ has the joint embedding property.

Proof. The forward direction is clear. For the backward direction, let $\mathbb{K} = \{\mathcal{A}_0, \mathcal{A}_1, \ldots\}$, we construct $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \mathcal{M}_2 \subseteq \ldots$ inductively as follows:

- $\mathcal{M}_0 \coloneqq \mathcal{A}_0$.
- $\mathcal{M}_{n+1}$ is chosen from $\mathbb{K}$ using joint embedding property such that $\mathcal{M}_n \hookrightarrow \mathcal{M}_{n+1} \hookleftarrow \mathcal{A}_{n+1}$.

$$\begin{tikzcd}
\mathbb{K} = & \{\mathcal{A}_0, \arrow[dr, hook, dashed] & \mathcal{A}_1, \arrow[dr, hook, dashed] & \mathcal{A}_2, \arrow[dr, hook, dashed] & \ldots\} & & \\
& & \mathcal{M}_0 \arrow[r, hook, dashed] & \mathcal{M}_1 \arrow[r, hook, dashed] & \mathcal{M}_2 \arrow[r, hook, dashed] & \ldots \arrow[r, hook, dashed] & \mathcal{M} \coloneqq \bigcup_n \mathcal{M}_n
\end{tikzcd}$$

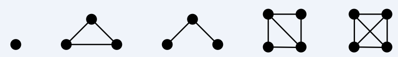
Let $\mathcal{M} \coloneqq \bigcup_n \mathcal{M}_n$. Then,

- $\mathbb{K} \subseteq \mathrm{Age}(\mathcal{M})$: because each $\mathcal{A}_n \in \mathcal{K}$ is embedded into $\mathcal{M}_n \subseteq \mathcal{M}$, so $\mathcal{A}_n$ is embedded into $\mathcal{M}$;
- $\mathrm{Age}(\mathcal{M}) \subseteq \mathbb{K}$: for a finite subset $\mathcal{N} \subseteq \mathcal{M}$, we have $\mathcal{N} \subseteq \mathcal{M}_n$ for some n, so by hereditary property, $\mathcal{M}_n \in \mathbb{K}$, hence $\mathcal{N} \in \mathbb{K}$.

■

Example. $\mathrm{Age}(\mathbb{Q}, <) = \mathrm{Age}(\mathbb{N}, <) = \mathrm{Age}(\mathbb{Z}, <) = \{\text{all finite linear orders}\}$, and in fact, it's also the age of any infinite linear order.

Example. $\mathrm{Age}(\text{random graph theory}) = \{\text{all finite graphs}\} = \mathrm{Age}(\coprod \{\text{finite graphs}\})$.



5.3 Fraïssé Theorem

Fraïssé asked when the class $\mathbb{K}$ determines a single structure. It turns out that in addition to the hereditary property and the joint embedding property, we need a third property of a class $\mathbb{K}$.

Definition 5.3.1 (Ultrahomogeneous). A countable structure $\mathcal{M}$ is *ultrahomogeneous* if for any finite subsets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{M}$ and isomorphism $g\colon \mathcal{A} \cong \mathcal{B}$, there is an automorphism $\widetilde{g}\colon \mathcal{M} \cong \mathcal{M}$ extending g.

$$\begin{tikzcd}
\mathcal{A} \arrow[r, hook, two heads, "g"] \arrow[d, hook] & \mathcal{B} \arrow[d, hook] \\
\mathcal{M} \arrow[r, hook, two heads, dashed, "\widetilde{g}"] & \mathcal{M}
\end{tikzcd}$$

Definition 5.3.2 (Amalgamation property). A class $\mathbb{K}$ of finite $\mathcal{L}$-structure has the *amalgamation property* if for all $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{K}$ such that
$$\begin{tikzcd} & \mathcal{B} \\ \mathcal{A} \arrow[ur, hook, "f"] \arrow[dr, hook, "g"'] & \\ & \mathcal{C} \end{tikzcd}$$
, there exists $\mathcal{D} \in \mathbb{K}$ such that
$$\begin{tikzcd} \mathcal{B} \arrow[dr, dashed, hook, "\widetilde{f}"] & \\ & \mathcal{D} \\ \mathcal{C} \arrow[ur, dashed, hook, "\widetilde{g}"'] & \end{tikzcd}$$
such that $\widetilde{f} \circ f = \widetilde{g} \circ g$, i.e., the following diagram commutes:

$$\begin{tikzcd} & \mathcal{B} \arrow[dr, dashed, hook, "\widetilde{f}"] & \\ \mathcal{A} \arrow[ur, hook, "f"] \arrow[dr, hook, "g"'] & & \mathcal{D} \\ & \mathcal{C} \arrow[ur, dashed, hook, "\widetilde{g}"'] & \end{tikzcd}$$

Intuition. We can "glue" $\mathcal{B}$ and $\mathcal{C}$ along their "common part" $\mathcal{A}$ to get $\mathcal{D}$.

Definition 5.3.3 (Extension property). A countable structure $\mathcal{M}$ has the *extension property* w.r.t. a class $\mathbb{K}$ of (finite) structure if for all $\mathcal{A}, \mathcal{B} \in \mathbb{K}$ and $f : \mathcal{A} \to \mathcal{M}$ and $g : \mathcal{A} \hookrightarrow \mathcal{B}$, there exists $h : \mathcal{B} \hookrightarrow \mathcal{M}$ such that $h \circ g = f$.

$$\begin{tikzcd} \mathcal{A} \arrow[rr, hook, "g"] \arrow[dr, hook, "f"'] & & \mathcal{B} \arrow[dl, dashed, hook, "h"] \\ & \mathcal{M} & \end{tikzcd}$$

Intuition. This is a direct generalization of extension axiom (as the name suggests).

Note. $\mathcal{M}$ ultrahomogeneous implies $\mathcal{M}$ has extension property w.r.t. $\text{Age}(\mathcal{M})$.[a]

[a]In homework, we will show the converse. More generally, $\mathcal{M}$ has extension property w.r.t. $\text{Age}(\mathcal{N})$ implies $\mathcal{N}$ embeds into $\mathcal{M}$, so $\mathcal{M}$ is a Fraisse limit, hence $\text{Age}(\mathcal{N}) \subseteq \text{Age}(\mathcal{M}) \Rightarrow \mathcal{N} \hookleftarrow \mathcal{M}$.

Proof. Without loss of generality, let $\mathcal{A}, \mathcal{B} \subseteq \mathcal{M}$ and f is the inclusion, we have

$$\begin{tikzcd} \mathcal{A} \arrow[r, hook, two heads, "g"] \arrow[d, hook] & \text{Im}(g) \arrow[r, hook] & \mathcal{B} \arrow[d, hook] \arrow[dll, dashed, "h"'] \\ \mathcal{M} \arrow[rr, dashed, "\exists \widetilde{g}"'] & & \mathcal{M} \end{tikzcd}$$

where $h := \widetilde{g}^{-1}|_{\mathcal{B}}$. ⊛

We now see the generalized version of baby Fraïssé theorem.

Theorem 5.3.1 (Fraïssé theorem). A class $\mathbb{K}$ of finite $\mathcal{L}$-structure is $\text{Age}(\mathcal{M})$ for an ultrahomogeneous countable $\mathcal{M}$ if and only if

- $\mathbb{K}$ is countable up to isomorphism;
- $\mathbb{K} \neq \varnothing$;
- $\mathbb{K}$ has the hereditary property;
- $\mathbb{K}$ has the joint embedding property;

- $\mathbb{K}$ has the amalgamation property.

Moreover, in that case, there exists a unique-up-to-isomorphism countable ultrahomogeneous $\mathcal{M}$ such that $\mathrm{Age}(\mathcal{M}) = \mathbb{K}$.

Definition 5.3.4 (Fraïssé class)**.** A class $\mathbb{K}$ with properties described in Fraïssé theorem is called a *Fraïssé class*.

Definition 5.3.5 (Fraïssé limit)**.** The countable ultrahomogeneous $\mathcal{M}$ such that $\mathrm{Age}(\mathcal{M}) = \mathbb{K}$ is the *Fraïssé limit* of $\mathbb{K}$, denoted as $\mathrm{Flm}(\mathbb{K})$.

Lecture 21: Proof of Fraïssé Theorem

23 Mar. 14:30

As previously seen. We want to build an infinite structure from finite pieces.

To prove the Fraïssé theorem, we need the following lemma.

Lemma 5.3.1. Let $\mathbb{K}$ have the amalgamation property. Then for $\mathcal{B}, \mathcal{A}_1, \ldots, \mathcal{A}_n, \mathcal{C}_1, \ldots, \mathcal{C}_n \in \mathbb{K}$ with embeddings

$$\begin{tikzcd} \mathcal{B} & & \mathcal{C}_i \\ & \mathcal{A}_i \arrow[ul, "f_i"] \arrow[ur, "g_i"'] & \end{tikzcd}$$

, there exists $\mathcal{D} \in \mathbb{K}$ with embedding

$$\begin{tikzcd} & \mathcal{D} & \\ \mathcal{B} \arrow[ur, "\widetilde{f}"] & & \mathcal{C}_i \arrow[ul, "\widetilde{g}_i"'] \end{tikzcd}$$

such that $\widetilde{f} \circ f_i = \widetilde{g}_i \circ g_i$, i.e.,

$$\begin{tikzcd} & \mathcal{D} & & & \\ \mathcal{B} \arrow[ur, "\widetilde{f}"] & & \mathcal{C}_1 \arrow[ul, "\widetilde{g}_1"] & \mathcal{C}_2 \arrow[ull, "\widetilde{g}_2"] & \mathcal{C}_3 \arrow[ulll, "\widetilde{g}_3"] & \cdots \\ & \mathcal{A}_1 \arrow[ul, "f_1"] \arrow[ur, "g_1"'] & \mathcal{A}_2 \arrow[ull, "f_2"] \arrow[ur, "g_2"'] & \mathcal{A}_3 \arrow[ulll, "f_3"] \arrow[ur, "g_3"'] & \cdots \end{tikzcd}$$

Proof. By induction on n with amalgamation property, we have

$$\begin{tikzcd} & & \mathcal{D}_2 & \\ & \mathcal{D}_1 \arrow[ur, dashed] & & \\ \mathcal{B} \arrow[ur, dashed] & & \mathcal{C}_1 \arrow[ul, dashed] & \mathcal{C}_2 \arrow[uul, dashed] \\ & \mathcal{A}_1 \arrow[ul] \arrow[ur] & \mathcal{A}_2 \arrow[ull] \arrow[ur] \arrow[uul, dotted, bend left] & \end{tikzcd}$$

∎

Now, we try to prove the Fraïssé theorem.

Proof of Theorem 5.3.1. To prove the forward direction, we need only to check AP for $\mathrm{Age}(\mathcal{M})$. We use EP to extend $\mathcal{A} \hookrightarrow \mathcal{B} \hookrightarrow \mathcal{M}$ along $\mathcal{A} \hookrightarrow \mathcal{C}$. Let $\mathcal{D}$ be the union of images of $\mathcal{C} \hookrightarrow \mathcal{M}$ and $\mathcal{B} \hookrightarrow \mathcal{M}$.

$$\begin{tikzcd} \mathcal{A} \arrow[r, hook, "f"] \arrow[d, hook] \arrow[rr, bend left] & & \mathcal{B} \arrow[d, hook] \\ \mathrm{Age}(\mathcal{M}) \ni \mathcal{C} \arrow[rr, dashed, "\exists \text{ by EP}"'] & & \mathcal{D} \subseteq \mathcal{M} \end{tikzcd}$$

For the backward direction, we'll build $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \ldots$ as in baby Fraïssé theorem to ensure

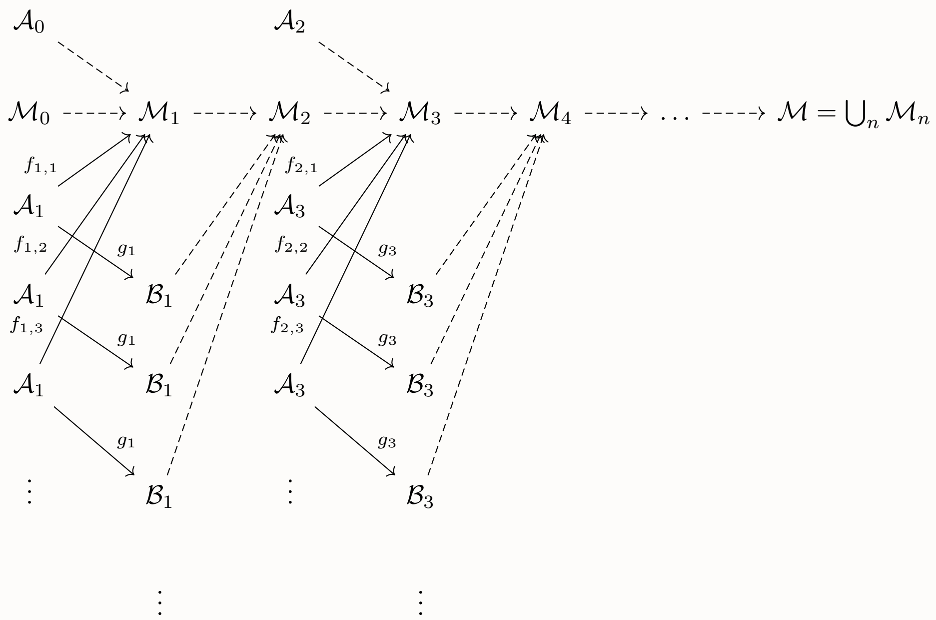
$\mathcal{M}$ obeys countably many "conditions". A "condition" means:

- a structure $\mathcal{A} \in \mathbb{K}$ (means we need to embed $\mathcal{A}$ into $\mathcal{M}_n$ at some stage),
- a tuple $(\mathcal{A}, \mathcal{B}, g)$ where $\mathcal{A}, \mathcal{B} \in \mathbb{K}$ and g is an embedding $\mathcal{A} \hookrightarrow \mathcal{B}$ (means we need to amalgamate g into $\mathcal{M}_n$).

  Then, we list al conditions

$$(\mathcal{A}_0, (\mathcal{A}_1, \mathcal{B}, g_1), \mathcal{A}_2, (\mathcal{A}_3, \mathcal{B}_3, g_3), \dots)$$

such that each condition of second type occurs infintely often. Consider

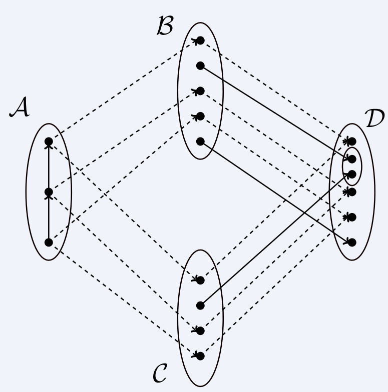


where $f_{1,i}$ are all embeddings $\mathcal{A}_1 \hookrightarrow \mathcal{M}_1$, for example. Then as before, $\mathrm{Age}(\mathcal{M}) = \mathbb{K}$. To check extension property w.r.t. $\mathbb{K}$, we have

$$\begin{tikzcd} \mathbb{K} \ni \mathcal{A} \arrow[r, "f"] \arrow[d] & \mathcal{M}_n \subseteq \mathcal{M} \\ \mathbb{K} \ni \mathcal{B} & \end{tikzcd}$$

since each triple we listed are infinitely often, at some $m \geq b$, $(\mathcal{A}_m, \mathcal{B}_m, g_m) = (\mathcal{A}, \mathcal{B}, g)$, so we just amalgamate this into $\mathcal{M}_{m+1}$. ■

Example. Let $\mathbb{K}$ be finite linear orders which has amalgamation property:



Note. $\mathbb{K}$ has the strong amalgamation property (SAP) if it's always possible to amalgamate such that $\mathrm{Im}(\widetilde{f}) \cap \mathrm{Im}(\widetilde{g}) = \mathrm{Im}(\widetilde{f} \circ f)$.[a]

$$\begin{tikzcd}
 & \mathcal{D} & \\
\mathcal{B} \arrow[ur, dashed, "\widetilde{f}"] & & \mathcal{C} \arrow[ul, dashed, "\widetilde{g}"'] \\
 & \mathcal{A} \arrow[ul, "f"] \arrow[ur, "g"'] &
\end{tikzcd}$$

[a] Recall that $\widetilde{f} \circ f = \widetilde{g} \circ g$.

Theorem 5.3.2. A Fraïssé limit has $\mathrm{Age}(\mathcal{M})$ with SAP if and only if $\mathcal{M}$ has trivial acl, i.e., $\mathrm{acl}(A) = A$ for all $A \subseteq M$.

Appendix

Bibliography

[HH97] W. Hodges and S.M.S.W. Hodges. *A Shorter Model Theory*. Cambridge University Press, 1997. ISBN: 9780521587136. URL: https://books.google.com/books?id=S6QYeuo4p1EC.

[Hin05] P.G. Hinman. *Fundamentals of Mathematical Logic*. Taylor & Francis, 2005. ISBN: 9781568812625. URL: https://books.google.com/books?id=xA6D8o72qAgC.

[Mar02] D. Marker. *Model Theory : An Introduction*. Graduate Texts in Mathematics. Springer New York, 2002. ISBN: 9780387987606. URL: https://books.google.com/books?id=gkvogoiEnuYC.

[Rob49] Julia Robinson. "Definability and decision problems in arithmetic". In: *The Journal of Symbolic Logic* 14.2 (1949), pp. 98–114. DOI: 10.2307/2266510.