220A Class Notes

Daniel Rui - Fall 2022

Professor: Artem Chernikov. Recommends YouTube videos for recorded versions of lectures. Proposed releasing all homeworks, and allowing us to do them at our own pace. Final grade based on homework and final exam. Texts: *A First Journey Through Logic* by Hils-Loeser, and later *Model Theory, an Introduction* by Marker. This is my first time taking notes on computer live during class; consider this an experiment. I will use `\section` to number lectures based off of 2021 YouTube recordings, and `\subsection` to mark the 2022 lecture dates.

Lecture 0

0.1 9/23/22 Class 1

Historic remarks of subject (Mathematical Logic): in Ancient Greece, only one subject, philosophy (everything part of it somehow). Aristotle had idea of make precise implications and connections ("probably invented a million times, in Western canon we like to attribute everything to the Ancient Greeks"). Later, ideas developed by Leibniz (thinking can be reduced to calculations), Frége (formalize ideas, in particular restricted formal languages, formal deducibility), Boole (famous for boolean algebra — "highest honor in mathematics, name becomes lowercase adjective").

Early 20th century, major push to make everything rigorous due to "foundational crisis in mathematics": mathematics expanded in scope (people used to study concrete/calculational things like numbers, polynomials, and then people started studying more abstract math. Example: function used to be computation formula, later abstracted. Actually purely existence proofs were more and more common, e.g. Hilbert basis theorem, but controversial). Unfortunately, paradoxes arose, e.g. Russell paradox. "Not very deep", but pointed out we should be careful. The response to this was mathematical logic: through formal language, we can control very precisely what is "true", what can be "proved". Originally, "protective role", protect us from doing nonsense.

ZFC arose (formalization of set theory), everything in mathematics can be encoded in ZFC, so everything in principle can be boiled down to formal manipulations in ZFC. "Pretty much understood" — yes, many different theories one can work with, but they all essentially encode each other, so birds-eye-view they are all sort of the same. But interesting thing happened ("late" 20th century): originally "protective role", it grew into tool to prove things in classical areas of mathematics.

Since then, it is big subject now. Traditionally, split into 4 main branches:

- model theory A
- recursion/computability (or rather, for now, incomputability) theory — computer science (CS) branched out of here, e.g. Turing machines, diagonalizability, etc. B
- set theory C
- proof theory, "a bit diluted from other branches". A

UCLA 220A,B,C courses subjects notated above.

SYNTAX	SEMANTICS
formal languages	"meaningful interpretation to symbols"
"can put in computer"	
formal deduction	models, properties of models

Gödel's completeness theorem tells us syntax and semantics are connected in precise, tight way. Important consequence: the "0th theorem of model theory", compactness theorem.

Next part of course, foundations of model theory: Löwenheim-Skolem (make models smaller or larger), types of structures (saturated, everything that can happen does; or prime, "opposite" everything that doesn't "have to" happen doesn't happen). Once we have this, we prove some "crazy theorems": Vaught's "no-2" theorem (no theory in any language can have exactly 2 countable models — can have 1, 3, 4, so on, but not 2??? "Quite amusing").

Hope is by the end, prove (or get as close as possible to proving) Morley's categoricity theorem: let T be any first order (FO) theory in countable language $\mathscr{L}$ (mostly works with countable languages in this course, because "closer to reality"), and assume that it has unique (up to isomorphism) model of size κ for some uncountable cardinal κ, i.e. $\kappa > \aleph_0$; then it has a unique (up to isomorphism) model of size λ for ANY cardinal $\lambda > \aleph_0$ (cf. principle that FO logic doesn't understand cardinality well). May not seem that interesting, but we'll see by the end. "Important part is not statement, but techniques it brings to logic, including topological principles". Viewed as "birth of modern model theory". So beginning course study (first-order) logic, last part study theories.

Bonus: in combinatorics, somehow "large finite things more complicated than infinite things"; if one can take some infinite limit ("ultraproduct") and use infinitary logic/methods of model theory, one can understand asymptotics of some combinatorial things. Model theory o-minimality, stability have implications to number theory things like Mordell-Lang ("somehow first proofs of those things are model theory in a sense").

1 Lecture 1

1.1 9/26/22 Class 2

Let us begin by defining syntax: languages, formulas, etc.. Our choice of language is not completely canonical, but it is only clear if we have made a "good" choice *a postieri*. This *a postieri* knowledge comes from experience already doing math:

Example 1.1: Symbols and interpretations we already understand

Consider the sentence $\forall x(x > 0 \implies \exists y(y \cdot y = x))$. It is just a string written in marks we recognize as mathematical symbols. With the meaning these symbols hold (from experience using them), we know that this holds/is valid in an ordered ring $\iff$ every positive element is a square, i.e./e.g. in $\mathbb{R} = (\mathbb{R}, 0, 1, +, \cdot, <)$ but not in $(\mathbb{Q}, 0, 1, +, \cdot, <)$. That is to say, we have an intuitive understanding of giving this string of symbols meaning, and the meaning depends on what context we are considering.

At this point in the course, we will not deal with foundational issues (we will not do anything subtle with the concept of "set", yet), so we use the word "set" without any further fuss.

Definition 1.1: First-order (FO) language

A (FO) language is a set of formal symbols $\mathscr{L}$ (which for this definition I will write use a different typewriter-style font), consisting of the two disjoint subsets:

(a) Logical symbols (common to all languages, hence often left unwritten), all distinct:
- parentheses `(` and `)`.
- the set of (pairwise distinct) variables $\mathscr{V} := \{\mathtt{v}_n : n \in \mathbb{N}\}$ (not dependent on $\mathscr{L}$, so as opposed to like $\mathscr{C}^{\mathscr{L}}$, only write as $\mathscr{V}$)
- the equality symbol `=`
- logical connectives: negation $\neg$, conjunction $\wedge$
- the existential quantifier $\exists$.

(b) and the signature of $\mathscr{L}$, denoted $\sigma^{\mathscr{L}}$, or the "non-logical symbols" (specific to choice of language $\mathscr{L}$ — because logical symbols are common, we often abuse notation and say $\mathscr{L} = \sigma^{\mathscr{L}}$), which is the disjoint union of:
- a set of (pairwise distinct) constant symbols $\mathscr{C}^{\mathscr{L}}$
- a sequence of sets $\mathscr{F}_n^{\mathscr{L}}$ for $n \in \mathbb{N}^+$ where the elements of $\mathscr{F}_n^{\mathscr{L}}$ are called "n-ary function symbols". The "function symbols" are $\mathscr{F}^{\mathscr{L}} = \bigcup_{n=1}^{\infty} \mathscr{F}_n^{\mathscr{L}}$ (disjoint union!)
- a sequence of sets $\mathscr{R}_n^{\mathscr{L}}$ for $n \in \mathbb{N}^+$, called the "n-ary any relational symbols" or "n-ary predicates". The "relational/predicate symbols" are $\mathscr{R}^{\mathscr{L}} = \bigcup_{n=1}^{\infty} \mathscr{R}_n^{\mathscr{L}}$ (disjoint union!)

The sets of non-logical symbols can be infinite or empty, or whatever. Again, let me emphasize that these are just marks on a paper/screen, no meaning imbued/attached.

From now on, we leave out (FO) since everything in this course is first-order. What is first-order? Only means we quantify over elements of the structure, instead of say 2nd-order logic which can quantify over subsets of structure (or 3rd-order, which can quantify over subsets of subsets of structures...). Basically past 1st-order, set theory rears its head.

Definition 1.2: Cardinality of $\mathscr{L}$

Let $|\mathscr{L}|$ denote the cardinality of $\mathscr{L}$. It is infinite because infinitely many variables; more precisely, $|\mathscr{L}| = \aleph_0 + |\mathscr{F}^{\mathscr{L}}| + |\mathscr{C}^{\mathscr{L}}| + |\mathscr{R}^{\mathscr{L}}|$.

Example 1.2: Some languages

Some examples of languages:

- the empty language $\mathscr{L}_\varnothing :=$ [logical part]$\cup\varnothing$ where again signature of this language $\sigma^{\mathscr{L}_\varnothing} := \varnothing$.
- the language of orders $\mathscr{L}_{\text{order}} :=$ [logical part] $\cup \{\texttt{<}\}$ where $\texttt{<} \in \mathscr{R}_2^{\mathscr{L}}$, binary relation symbol.
- (I'm tired of typing `\code`, just imagine all symbols below are in typewriter font) the language of rings

$$\mathscr{L}_{\text{ring}} := \ldots \cup \{\underbrace{0, 1}_{\in\mathscr{C}}, \underbrace{-}_{\in\mathscr{F}_1}, \underbrace{+, \cdot}_{\in\mathscr{F}_2}\}$$

- the language of ordered rings $\mathscr{L}_{\text{ord.ring}} := \mathscr{L}_{\text{order}} \cup \mathscr{L}_{\text{ring}}$
- the language of sets $\mathscr{L}_{\text{set}} := \ldots \cup \{\underbrace{\in}_{\in\mathscr{R}_2}\}$
- the language of groups $\mathscr{L}_{\text{group}} := \ldots \cup \{\underbrace{1}_{\in\mathscr{C}}, \underbrace{\bullet^{-1}}_{\in\mathscr{F}_1}, \underbrace{\cdot}_{\mathscr{F}_2}\}$
- the language of graphs $\mathscr{L}_{\text{graph}} := \ldots \cup \{\underbrace{E}_{\in\mathscr{R}_2}\}$
- the language of arithmetic

$$\mathscr{L}_{\text{arith}} := \ldots \cup \{\underbrace{0}_{\in\mathscr{C}}, \underbrace{S}_{\in\mathscr{F}_1}, \underbrace{+, \cdot}_{\in\mathscr{F}_2}, \underbrace{\leq}_{\in\mathscr{R}_2}\}$$

We may now imbue abstract/arbitrary/formal sets of symbols with some meaning/interpretation:

Definition 1.3: $\mathscr{L}$-structure

An $\mathscr{L}$-structure $\mathcal{A}$ consists of a nonempty set A (called the base set/(underlying) universe of $\mathcal{A}$), together with

- an element $\texttt{c}^{\mathcal{A}} \in A$ for every $\texttt{c} \in \mathscr{C}^{\mathscr{L}}$,
- a function $\texttt{f}^{\mathcal{A}} : A^n \to A$ for every $\texttt{f} \in \mathscr{F}_n^{\mathscr{L}}$ (for every $n \in \mathbb{N}^+$),
- and a subset $\texttt{R}^{\mathcal{A}} \subseteq A^n$ for each $\texttt{R} \in \mathscr{R}_n^{\mathscr{L}}$.

We write $\mathcal{A} := \left(A; (\texttt{Z}^{\mathcal{A}})_{\texttt{Z}\in\sigma^{\mathscr{L}}}\right)$. Such $\texttt{Z}^{\mathcal{A}}$ are called interpretations of symbols $\texttt{Z} \in \sigma^{\mathscr{L}}$ in the $\mathscr{L}$-structure $\mathcal{A}$. Although I do like the typewriter font distinction, just imagine it in future lectures since again I'm tired of typing `\code` and will stop doing so after this lecture.

Example 1.3: Some structures

Some examples of structures:

- We can define an $\mathscr{L}_{\text{arith}}$-structure $\mathcal{N}$ with base set $\mathbb{N}$ and interpretations

$$\begin{array}{c||c||c|c|c||c|}
\mathtt{Z} & \mathtt{0} & \mathtt{S} & \cdot & + & < \\ \hline
\mathtt{Z}^{\mathcal{N}} & \underbrace{0}_{\in\mathbb{N}} & \underbrace{x\mapsto x+1}_{:\mathbb{N}\to\mathbb{N}} & \underbrace{(x,y)\mapsto x\cdot y}_{:\mathbb{N}^2\to\mathbb{N}} & \underbrace{(x,y)\mapsto x+y}_{:\mathbb{N}^2\to\mathbb{N}} & \underbrace{[x<y]}_{\subseteq\mathbb{N}^2}
\end{array}$$

- We can define an $\mathscr{L}_{\text{ring}}$-structure $\mathcal{C}$ by

$$\mathcal{C} := \left(\mathbb{C};\ \begin{array}{c||c|c||c|c|c|}
\mathtt{Z} & \mathtt{0} & \mathtt{1} & \mathtt{-} & + & \cdot \\ \hline
\mathtt{Z}^{\mathcal{C}} & \underbrace{0}_{\in\mathbb{C}} & \underbrace{1}_{\in\mathbb{C}} & \underbrace{x\mapsto -x}_{:\mathbb{C}\to\mathbb{C}} & \underbrace{(x,y)\mapsto x+y}_{:\mathbb{C}^2\to\mathbb{C}} & \underbrace{(x,y)\mapsto x\cdot y}_{:\mathbb{C}^2\to\mathbb{C}}
\end{array}\right)$$

- We can define the $\mathscr{L}_{\text{ord.ring}}$-structure

$$\mathcal{R} := \left(\mathbb{R};\ \begin{array}{c||c|c||c|c|c||c|}
\mathtt{Z} & \mathtt{0} & \mathtt{1} & \mathtt{-} & + & \cdot & < \\ \hline
\mathtt{Z}^{\mathcal{A}} & \underbrace{0}_{\in\mathbb{R}} & \underbrace{1}_{\in\mathbb{R}} & \underbrace{x\mapsto -x}_{:\mathbb{R}\to\mathbb{R}} & \underbrace{(x,y)\mapsto x+y}_{:\mathbb{R}^2\to\mathbb{R}} & \underbrace{(x,y)\mapsto x\cdot y}_{:\mathbb{R}^2\to\mathbb{R}} & \underbrace{[x<y]}_{\subseteq\mathbb{R}^2}
\end{array}\right)$$

From now on, defining structures I will not be so explicit; as you can see the interpretations of many symbols are obvious once I specify base set.

Definition 1.4: Structure isomorphism

Given two $\mathscr{L}$-structures $\mathcal{A},\mathcal{B}$, we say that $\mathcal{A},\mathcal{B}$ are isomorphic as $\mathscr{L}$-structures (generalizing notions of isomorphism in algebra and graph theory), denoted $\mathcal{A}\simeq\mathcal{B}$, if there exists an $\mathscr{L}$-structure isomorphism $F:\mathcal{A}\xrightarrow{\cong}\mathcal{B}$, i.e. bijection between the base sets A,B, which commutes with interpretations of symbols in $\sigma^{\mathscr{L}}$, i.e.

- $F(\mathtt{c}^{\mathcal{A}})=\mathtt{c}^{\mathcal{B}}$ for all constant symbols $c\in\mathscr{C}^{\mathscr{L}}$
- $F(\mathtt{f}^{\mathcal{A}}(a_1,\dots,a_n)) = \mathtt{f}^{\mathcal{B}}(F(a_1),\dots,F(a_n))$ for all $f\in\mathscr{F}_n^{\mathscr{L}}$ and $\bar{a}:=(a_1,\dots,a_n)\in A^n$ (ranging over $n\in\mathbb{N}^+$).
- and $(a_1,\dots,a_n)\in\mathtt{R}^{\mathcal{A}} \iff (F(a_1),\dots,F(a_n))\in\mathtt{R}^{\mathcal{B}}$ for all $\mathtt{R}\in\mathscr{R}_n^{\mathscr{L}}$ and $\bar{a}:=(a_1,\dots,a_n)\in A^n$ (ranging over $n\in\mathbb{N}^+$).

As with all structures/structure-preserving isos, can just consider structure preserving property without the bijection part, leading to notion of (homo)morphism $F:\mathcal{A}\to\mathcal{B}$. Injective morphism are as usual called embeddings.

We can now build up more complicated syntax based on this setup:

Definition 1.5: Words

A word w over a set/alphabet E is just a finite string $\mathtt{a}_0\mathtt{a}_1\dots\mathtt{a}_{k-1}$ with symbols $\mathtt{a}_i\in E$ (for all $i\in[k-1]$), where k is the word-length, denoted $\operatorname{len}(w)$, of the word w and E^* denotes the set of all words over E. Examples: words in $\mathscr{L}_{\text{ord.ring}}$ include $\mathtt{<0S\text{-}0abxyz}\neg\exists\mathtt{S+\text{-}}\wedge\wedge\mathtt{00<\text{-}}$ and $\mathtt{0<S0}\wedge\neg\exists\mathtt{x\ Sx=0}$ (one of these words is gibberish and one "makes sense"...).

Definition 1.6: Terms

Let $\mathscr{L}$ be a language. The set $\mathscr{T}^{\mathscr{L}}$ of $\mathscr{L}$-terms is the smallest (in the sense of inclusion) subset D of $\mathscr{L}^*$ (the set of all words/strings made from symbols in $\mathscr{L}$) containing the variables and constants of $\mathscr{L}$ s.t. D is closed under functions, i.e. if $\mathtt{f} \in \mathscr{F}_n^{\mathscr{L}}$ and $t_1, \ldots, t_n \in D$, then the string (made by concatenating $\mathtt{f}$ with $t_1, \ldots, t_n$, which I write in math italics not typewriter font because t_i is not a formal symbol in the language, as opposed to say $\mathtt{f} \in \mathscr{F}_n^{\mathscr{L}}$) $\mathtt{f}t_1 \ldots t_n \in D$. One can also define this inductively: $\mathscr{T}^{\mathscr{L}} := \bigcup_{k=0}^{\infty} \mathscr{T}_k^{\mathscr{L}}$ where

$$\mathscr{T}_0 := \mathscr{C}^{\mathscr{L}} \cup \mathscr{V}, \text{ and } \mathscr{T}_{k+1}^{\mathscr{L}} := \mathscr{T}_k^{\mathscr{L}} \cup \left\{ \mathtt{f}t_1 \ldots t_n : n \in \mathbb{N}^+, \mathtt{f} \in \mathscr{F}_k^{\mathscr{L}}, t_1, \ldots, t_n \in \mathscr{T}_k^{\mathscr{L}} \right\}$$

Remark: convention is concatenation $\mathtt{f}t_1 \ldots t_n$, but doing parentheses are OK too $\mathtt{f(}t_1 \ldots t_n\mathtt{)}$, and even for binary functions $\mathtt{(}t_1\mathtt{f}t_2\mathtt{)}$, e.g. instead of `·+xyz` we can write `((x+y)·z)`

Terms are more naturally understood in tree structure, e.g. https://en.wikipedia.org/wiki/Binary_expression_tree, but collapsing them down into a string format doesn't lose information, i.e. we can reconstruct the syntax tree exactly:

Proposition 1.1: Unique reading of terms

Any term $t \in \mathscr{T}^{\mathscr{L}}$ satisfies one and only one of the following possibilities:

(a) $t \in \mathscr{V}$,

(b) $t \in \mathscr{C}^{\mathscr{L}}$,

(c) or there is unique $n \in \mathbb{N}^+$, unique $\mathtt{f} \in \mathscr{F}_n^{\mathscr{L}}$ and unique sequence of terms $t_1, \ldots, t_n$ s.t. t is exactly $\mathtt{f}t_1 \ldots t_n$(string equality is character by character match).

Proof (sketch): first prove that no proper initial segment of any term can be a term, using induction on word-length of a given term. ■

Unique reading of terms does the "painful" part of dealing with lengths of terms, so that we may deal with easier notion of height of term, because we can induct on how many functions we compose (no need to mess around with checking arity of relations, functions, etc.).

Definition 1.7: Height of term

The <u>height</u> of a term t, denoted $\mathrm{ht}(t)$ is the least natural number k s.t. $t \in \mathscr{T}_k^{\mathscr{L}}$. From uniqueness of reading terms, $\mathrm{ht}(\mathtt{f}t_1 \ldots t_n) = 1 + \max_{i \in [n]^+} \mathrm{ht}(t_i)$.

Using terms, we can now construct formulas, which are formal words/strings to which we may begin attaching/interpreting notions of "truth", as opposed to terms defined above (in later lectures, we will see "Tarski's definition of truth", i.e. formula satisfaction in structures):

Definition 1.8: Atomic formulas

An <u>atomic $\mathscr{L}$-formula</u> is either

- a word of the form $t_1\mathtt{=}t_2$ for $\mathscr{L}$-terms t_1, t_2,
- or a word of the form $\mathtt{R}t_1 \ldots \cdot t_n$ for $\mathtt{R} \in \mathscr{R}_n^{\mathscr{L}}$ and $\mathscr{L}$-terms $t_1, \ldots, t_n$.

Definition 1.9: Formulas

The set of $\mathscr{L}$-formulas, $\mathcal{Fml}^{\mathscr{L}}$ is the smallest subset D of $\mathscr{L}^*$ containing all atomic $\mathscr{L}$-formulas, and closed under negation, conjunction, and existential quantifier, i.e. if $\mathtt{x} \in \mathscr{V}$ and $\varphi, \psi \in D$, then words $\neg\varphi$, $\mathtt{(}\varphi \wedge \psi\mathtt{)}$ and $\exists\mathtt{x}\varphi$ are all in D as well. Again, we can approach this more inductively, and define $\mathcal{Fml}^{\mathscr{L}} = \bigcup_{n=0}^{\infty} \mathcal{Fml}_n^{\mathscr{L}}$ where $\mathcal{Fml}_0^{\mathscr{L}}$ is the set of all atomic $\mathscr{L}$-formulas, and

$$\mathcal{Fml}_{n+1}^{\mathscr{L}} := \mathcal{Fml}_n^{\mathscr{L}} \cup \{\neg\varphi : \varphi \in \mathcal{Fml}_n^{\mathscr{L}}\} \cup \{\mathtt{(}\varphi \wedge \psi\mathtt{)} : \varphi, \psi \in \mathcal{Fml}_n^{\mathscr{L}}\} \cup \{\exists\mathtt{x}\varphi : \varphi \in \mathcal{Fml}_n^{\mathscr{L}}, \mathtt{x} \in \mathscr{V}\}$$

Remark: could add more parentheses anywhere basically, but parentheses are necessary around conjunctions since $\mathtt{(}\exists\mathtt{xx=1} \wedge \mathtt{x=0)}$ is different than $\exists\mathtt{x(x=1} \wedge \mathtt{x=0)}$, and without parentheses these two are indistinguishable.

Metanotation alert 10/4/22: originally I envisioned typewriter font to be used to express a specific symbol in the language, e.g. $\mathtt{0}, \mathtt{+}, \mathtt{-}$ in $\mathscr{L}_{\text{ring}}$. However, unwittingly I've been using it all along to denote a more arbitrary symbol in the language, e.g. $\mathtt{c} \in \mathscr{C}^{\mathscr{L}}$ or $\mathtt{x} \in \mathscr{V}$. Such $\mathtt{c}, \mathtt{x}$ are not literally symbols in the language $\mathscr{L}$, but I guess I used typewriter font to indicate that they represent just one formal symbol in $\mathscr{L}$, as opposed to t, φ (written in math italics) which stand for whole words. But I ran into a problem yesterday in typing 220A Homework 1 Question 1: in unique reading/no proper initial segments of formulas of form $\exists\mathtt{x}\psi$, I got to the point where I wanted to write "so the variable $\mathtt{y}$ must actually be exactly $\mathtt{x}$". The confusion/cognitive dissonance is then that in the original envisioning where typewriter font symbols are actual symbols in the language, equality is literal character by character equality (so $\mathtt{y} = \mathtt{x}$ would be impossible). I will try to phase out bad usage. Using $\mathtt{v}_i$ still OK, since arbitrary/undetermined index i still in math italics. Argh, but I still want to use $\mathtt{f}, \mathtt{R}$ for arbitrary functions/relations...maybe different font? Solution: $\mathfrak{f} t_1 \dots t_n$, and $\exists\mathsf{x}\varphi$ vs. $\exists\mathtt{v}_i\varphi$.

2 Lecture 2

2.1 9/28/22 Class 3

Exactly like we had with terms, we again have unique readability, or like I said above converting between string representation and expression tree representation (the rules we have for forming formulas, including parentheses when forming conjunctions, ensure this):

Proposition 2.1: Unique reading of formulas

Any $\mathscr{L}$-formula φ satisfies exactly one of the following possibilities:

(a) φ is atomic,

(b) φ is equal as words (literally character by character match) to $\neg\psi$ for some unique $\mathscr{L}$-formula ψ,

(c) φ is equal as words to $(\psi \wedge \chi)$ for some unique $\mathscr{L}$-formulas ψ, χ,

(d) or φ is equal as words to $\exists \mathtt{x}\psi$ for some unique $\mathtt{x} \in \mathscr{V}$ and some $\mathscr{L}$-formula ψ.

In words, the intuition (from the inductive definition) is that φ is either atomic (in $\mathscr{Fmt}_0^{\mathscr{L}}$) or generated by negation, conjunction, or existential quantification of previously generated formulas.

Proof (sketch): exactly like we did with unique reading of terms, we first prove (by induction on word-length of formula) that no proper initial segment of formula is formula.■

As with terms, we can use unique readability to define a nicer notion of "length":

Definition 2.1: Height of formula

The height of a formula φ, denoted $\mathrm{ht}(\varphi)$ is the least natural number n s.t. $t \in \mathscr{Fmt}_n^{\mathscr{L}}$. From uniqueness of reading terms,

- $\mathrm{ht}(\neg\psi) = 1 + \mathrm{ht}(\psi)$.
- $\mathrm{ht}((\psi \wedge \chi)) = 1 + \max\{\mathrm{ht}(\psi), \mathrm{ht}(\chi)\}$.
- $\mathrm{ht}(\exists \mathrm{x}\psi) = 1 + \mathrm{ht}(\psi)$.

This is VERY VERY useful; we will use induction on formula height to prove many many things in the future.

We now define (syntatically) the notion of a variable occurring "freely" in a formula.

Definition 2.2: Free and bound variables

Let $\mathtt{v}_k \in \mathscr{V}$. We define a free occurrence of $\mathtt{v}_k$ in a formula φ by induction on $\mathrm{ht}(\varphi)$. The cases of atomic formulas, negation, and conjunction are straightforward; only existential quantification presents small subtleties:

- if φ atomic (i.e. $t_1 \mathtt{=} t_2$ or $\mathtt{R}t_1 \ldots, t_n$) then all occurrences of $\mathtt{v}_k$ in φ (i.e. occurrences of the substring v_k in the string φ) are free.
- if φ is equal to $\neg\psi$ for some $\mathscr{L}$-formula ψ, then free occurrences of $\mathtt{v}_k$ in φ are exactly those

in ψ.

- if φ is equal to $(\psi \wedge \chi)$ for some $\mathscr{L}$-formulas ψ, χ, then free occurrences of $\mathtt{v}_k$ in φ are exactly those in ψ and those in χ.
- if φ equals $\exists \mathtt{v}_l \psi$ for a $\mathscr{L}$-formula ψ and variable $\mathtt{v}_l$ and $l \neq k$, then the free occurrences of $\mathtt{v}_k$ in φ are exactly those in ψ.
- if φ equals $\exists \mathtt{v}_k \psi$ for a $\mathscr{L}$-formula ψ, then there are NO free occurrences of $\mathtt{v}_k$ in φ.

On the other hand, occurrences of a variable $\mathtt{v}_k$ in a formula φ which are not free are called bound.

Finally we define the free variables of φ, denoted Free(φ), to be the variables having at least one free occurrence in φ.

Definition 2.3: Sentences

A sentence is a formula with no free variables.

We have defined the notion of sentence purely syntactically, based on an inductive rule. Trying to formalize notion that in a formula, some variables may be quantified, some not (perhaps with the same name too!), we quickly see that due to complicated inductive nature of formulas, it may be hard to tell whether or not subformulas/variables fall under the scope of a quantifier. Hence the need for a careful definition.

Example 2.1: Example of free and bound variables

Consider the formula $\varphi := (\exists \mathtt{v}_0 \mathtt{v}_0 < \mathtt{v}_1 \wedge \mathtt{v}_0 = \mathtt{v}_1)$. The first two occurrences of (the substring) $\mathtt{v}_0$ in φ are bound, while the 3rd occurrence is free. All appearances of $\mathtt{v}_1$ are free. Thus Free$(\varphi) = \{\mathtt{v}_0, \mathtt{v}_1\}$.

We now formalize some other common notations we use (as "abbreviations"/"syntactic sugar" for terms or formulas written in the formal FO language we've been defining)

Definition 2.4: Notations/abbreviations we use

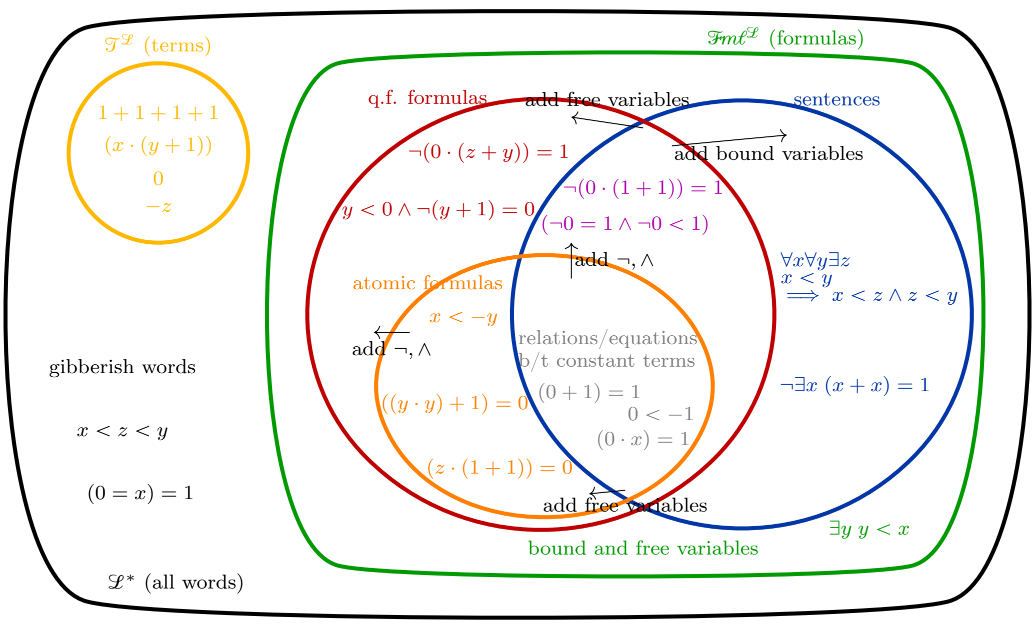
- $(\varphi \vee \psi)$ (read "disjunction") is shorthand for $\neg(\neg\varphi \wedge \neg\psi)$
- $\varphi \to \psi$ (read "implies/(forwards) implication") is shorthand for $\neg(\varphi \wedge \neg\psi)$ (or tautologically equivalent $\neg\varphi \vee (\varphi \wedge \psi)$, i.e. $\varphi \to \psi$ means that either φ is not true, or φ and ψ are both true; this is tautologically equivalent to $(\neg\varphi \vee \psi)$ of course).
- $\varphi \leftarrow \psi$ (read "implied by/backwards implication") is shorthand for $\psi \to \varphi$.
- $\varphi \leftrightarrow \psi$ (read "equivalence") shorthand for $(\varphi \to \psi \wedge \psi \to \varphi)$.
- $\forall \mathtt{x} \varphi$ (read "for all/universal quantification") is shorthand for $\neg \exists \mathtt{x} \neg \varphi$.
- As I wrote already in Def. 1.6, the string $\mathtt{f}(t_1, \ldots, f_n)$ ("written like a function applied to terms") is alternative notation for $\mathtt{f} t_1, \ldots, t_n$, and for binary functions, $(t_1 \mathtt{f} t_2)$ is another alternative notation for $\mathtt{f} t_1 t_2$, or without the parentheses if the reading is unambiguous.
- Similarly, the string $\mathtt{R}(t_1, \ldots, f_n)$ ("written like a function applied to terms") is alternative notation for $\mathtt{R} t_1, \ldots, t_n$, and for binary relations, $(t_1 \mathtt{R} t_2)$ is another alternative notation for

$\mathtt{R}t_1t_2$, or without the parentheses if the reading is unambiguous.

- $(\varphi_1 \wedge \ldots \wedge \varphi_n)$ or $\bigwedge_{i=1}^{n} \varphi_i$ is shorthand for $(\ldots((\varphi_0 \wedge \varphi_1) \wedge \varphi_2) \wedge \ldots)$. Similarly for disjunction $\vee$.

"We will not be so strict about parentheses... we don't want to count brackets in this class! The point of syntax in this class at least is just to do it once, make everything on solid logical/rigorous foundation, and never think about it again". However in our addition or omission of parentheses (in a way not following formal rules established earlier, but for ease of human understanding or writing), there are still conventions for "order of binding power", namely $\{\neg, \exists, \forall\} > \{\wedge\} > \{\vee\} > \{\rightarrow, \leftrightarrow\}$. For example $\forall \mathtt{x}\varphi \wedge \psi \rightarrow \chi$ is equivalent to $(((\forall \mathtt{x}\varphi) \wedge \psi) \rightarrow \chi)$ or in "standard" formal language $\neg((\neg(\exists \mathtt{x}(\neg\varphi)) \wedge \psi) \wedge \neg\chi)$.

Anyways, for $\mathscr{L} := \mathscr{L}_{\text{ord.ring}}$, the terminology we have defined (and which we will define later, in the case of q.f. formulas $\equiv$ quantifier-free formulas) can be pictured in the following diagram:



We can now define the interpretations of terms and formulas in a given $\mathscr{L}$-structure $\mathcal{A} := (A; \ldots)$.

Definition 2.5: Semantic interpretation

Let $\mathcal{A} := (A; \ldots)$ be an $\mathscr{L}$-structure (for FO language $\mathscr{L}$). An $\mathcal{A}$-assignment or assignment (with values in $\mathcal{A}$) is a function $\alpha : \mathscr{V} \to A$. Given an assignment α, we can define the interpretation (w.r.t. α) of a given term t, denoted $t^{\mathcal{A}}[\alpha]$, by induction on term-height:

- if $t \in \mathscr{T}_0^{\mathscr{L}}$, i.e. t is a variable $\mathtt{v}_i \in \mathscr{V}$ (resp. a constant $\mathtt{c} \in \mathscr{C}^{\mathscr{L}}$), then $\mathtt{v}_i^{\mathcal{A}}[\alpha] := \alpha(\mathtt{v}_i) \in A$ (resp. $\mathtt{c}^{\mathcal{A}}[\alpha] := \alpha(\mathtt{c}) \in A$).
- otherwise, t is equal to $\mathtt{f}t_1 \ldots t_n$ where t_i are terms of less height (hence already have interpretation defined by induction), we define $t^{\mathcal{A}}[\alpha] := (\mathtt{f}t_1 \ldots t_n)^{\mathcal{A}}[\alpha] := \mathtt{f}^{\mathcal{A}}(t_1^{\mathcal{A}}[\alpha], \ldots, t_n^{\mathcal{A}}[\alpha])$. Here, the "function notation/shorthand" $\mathtt{f}(t_1, \ldots, t_n)$ is now particularly evocative since

the interpretation (w.r.t. α) of the syntactic string $\mathtt{f}(t_1\mathtt{,}\ldots\mathtt{,}t_n\mathtt{)}$ is exactly what it looks like: (the interpretation in $\mathcal{A}$ of) $\mathtt{f}$ evaluated at (the interpretations in $\mathcal{A}$ w.r.t. α of) all the terms.

Thus, given a fixed $\mathscr{L}$-structure $\mathcal{A}$, a term t can be viewed as a function from $\mathrm{Maps}(\mathscr{V}, A) \simeq A^\infty$ to A. However, much of the information of the assignment is irrelevant when considering only interpretation of t:

Lemma 2.1: Only local properties of α matter (for fixed term)

If two $\mathcal{A}$-assignments α, β coincide on all variables occurring in a term t, then $t^{\mathcal{A}}[\alpha] = t^{\mathcal{A}}[\beta]$.

Proof: induction on term-height. ■

This tells us that a term t with variables $x_1, \ldots, x_n$ can be viewed as a function from $A^n \to A$. Indeed:

Definition 2.6: Terms as functions, notation

If t is a term, we may write $t(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ if the $\mathtt{x}_i$ are all all distinct variables, and all variables occurring in t belong in $\{\mathtt{x}_1, \ldots, \mathtt{x}_n\}$. Then given a term $t(\mathtt{x}_!, \ldots, \mathtt{x}_n)$ and $a_1, \ldots, a_n \in A$, we define/notate $t(a_1, \ldots, a_n) := t^{\mathcal{A}}(a_1, \ldots, a_n) := t^{\mathcal{A}}[a_1, \ldots, a_n] := t^{\mathcal{A}}[\alpha]$ for any $\alpha : \mathscr{V} \to A$ with $\alpha(\mathtt{x}_i) = a_i$ (above Lemma 2.1 tells us this is well-defined definition).

Phrased one last time, a given term t can be thought of as a function $A^n \to A$ as long as the number of variables occurring in t is $\leqslant n$ (and the notation we have chosen conflates standard function/evaluation notation with this term interpretation business).

Finally, we can define the satisfaction of formulas in structures ("Tarski's definition of Truth"):

Definition 2.7: Tarski's definition of truth/formula satisfaction

Let $\mathcal{A}$ be an $\mathscr{L}$-structure. By induction on $\mathrm{ht}(\varphi)$, we define (true/false-valued) $\mathcal{A} \models \varphi[\alpha]$ (read "φ is satisfied in $\mathcal{A}$ by/on α" or "φ on α is true in $\mathcal{A}$" or something similar):

- $\mathcal{A} \models t_1\mathtt{=}t_2[\alpha] \;:\Longleftrightarrow\; t_1^{\mathcal{A}}[\alpha] = t_2^{\mathcal{A}}[\alpha]$ (equal elements of A). Otherwise, we say $\mathcal{A} \not\models t_1\mathtt{=}t_2[\alpha]$.
- $\mathcal{A} \models \mathtt{R}t_1 \ldots t_n[\alpha] \;:\Longleftrightarrow\; (t_1^{\mathcal{A}}[\alpha], \ldots, t_n^{\mathcal{A}}[\alpha]) \in \mathtt{R}^{\mathcal{A}} \subseteq A$.
- $\mathcal{A} \models \neg\psi[\alpha] \;:\Longleftrightarrow\; \mathcal{A} \not\models \psi[\alpha]$.
- $\mathcal{A} \models (\psi \wedge \chi)[\alpha] \;:\Longleftrightarrow\; \mathcal{A} \models \psi[\alpha]$ and $\mathcal{A} \models \chi[\alpha]$.
- $\mathcal{A} \models \exists\mathtt{x}\psi[\alpha] \;:\Longleftrightarrow\;$ there exists $a \in A$ s.t. $\mathcal{A} \models \psi[\alpha_{a/\mathtt{x}}]$ where $\alpha_{a/\mathtt{x}}$ denotes the assignment defined by $\alpha_{a/\mathtt{x}}(\mathtt{x}) := a$ and $\alpha_{a/\mathtt{x}}(\mathtt{y}) := \alpha(\mathtt{y})$ for all variables $\mathtt{y} \neq \mathtt{x}$.

10/3/22: given formula $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ we write $\mathcal{A} \models \varphi$ if $\mathcal{A} \models \forall\mathtt{x}_1 \ldots \forall\mathtt{x}_n\varphi$.

"It's almost obvious, but still took Tarski to do it. That's the best outcome in mathematics: to be the first to define something that a postieri seems obvious, but took until you to figure it out/isolate it!"

My remarks: Given any formula φ and assignment α, there is a definitive yes/no answer to whether

$\mathcal{A}$ satisfies $\varphi[\alpha]$, because (as I wrote in 8/29/22 email draft) $\varphi[\alpha]$ interpreted in $\mathcal{A}$ is always of the form "some element in A equals another element in A" or "some element in A is in some subset of A" or "there exists some element in A satisfying some property" (or negation/conjunction of these), and these queries always have a definitive yes-or-no answer in set theory, via the "equality oracle" or "membership oracle" or "existence oracle" that I guess all sets implicitly come with? So we have merely used the truth oracles of set theory to define truth values of well-formed syntactic strings (in the context of a given $\mathscr{L}$-structure $\mathcal{A}$).

My further remarks (based on material not yet covered, very speculative/philosophical, could be completely off base): by Gödel completeness theorem, the only "knowledge" the truth oracles hold about base sets A of models $\mathcal{A}$ of a theory T that can be formally proven with axioms of T is "knowledge" shared by all models. And although models always have complete theory $\mathrm{Th}(\mathcal{A})$, by Gödel's incompleteness theorem, for models $\mathcal{A}$ "powerful enough to do basic arithmetic", the complete theory $\mathrm{Th}(\mathcal{A})$ can not be recursively enumerated. So for base sets of many models, the "knowledge" of the truth oracles of those base sets is forever beyond human reach (or at least reach of FO logic...)....?

2.2 9/30/22 Class 4

As we saw before with terms,

> *Proposition 2.2: Only local properties of α matter (for fixed formula φ)*
>
> If two assignments α, β coincide on $\mathrm{Free}(\varphi)$, then $\mathcal{A} \models \varphi[\alpha] \iff \mathcal{A} \models \varphi[\beta]$.

Proof: by induction on $\mathrm{ht}(\varphi)$: the atomic case is handled by Lemma 2.1 (assignments coincide $\rightsquigarrow$ same term-interpretation) and Def. 2.7 defining satisfaction of atomic formulas completely in terms of term-interpretation.

Inductive cases: let us consider φ equal to $\exists \mathtt{x} \psi$ where ψ has smaller height (the other cases φ equals $\neg\psi$ and $(\psi \wedge \chi)$ are easy — free variables in these cases are resp. $\mathrm{Free}(\varphi) = \mathrm{Free}(\psi)$ and $\mathrm{Free}(\varphi) = \mathrm{Free}(\psi) \cup \mathrm{Free}(\chi)$ because free variables don't change unless possibly when $\exists$ gets involved). If $\mathcal{A} \models \varphi[\alpha]$, by definition there is $a \in A$ s.t. $\mathcal{A} \models \psi[\alpha_{a/\mathtt{x}}]$. Note that any variable $\mathtt{y} \neq \mathtt{x}$ that is free in ψ is also free in φ (by def. of free). Thus α, β agreeing on $\mathrm{Free}(\varphi) \supseteq \mathrm{Free}(\psi)$ and $\alpha \equiv \alpha_{a/\mathtt{x}}$ on $\mathscr{V} \setminus \{x\}$ (same for β) implies that $\alpha_{a/\mathtt{x}}, \beta_{a/\mathtt{x}}$ agree on $\mathrm{Free}(\psi) \setminus \{\mathtt{x}\}$. Of course they agree on $\{x\}$ too, so $\alpha_{a/\mathtt{x}} \equiv \beta_{a/\mathtt{x}}$ on $\mathrm{Free}(\psi)$. Hence by the induction hypothesis, $\mathcal{A} \models \psi[\beta_{a/\mathtt{x}}]$, so by definition of satisfaction $\mathcal{A} \models \varphi[\beta]$. ■

Proofs by induction on height will always look like this; base case of atomic formulas, and inductive case splits into 3 cases, 2 of which (negation, conjunction) will usually be easy, with maybe the last case (existential quantification) being a bit trickier. Anyways, as with terms, we can introduce the notation

> *Definition 2.8: Formulas as functions, notation*
>
> If φ is a $\mathscr{L}$-formula, we may write $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ if the $\mathtt{x}_i$ are all all distinct variables, and all variables occurring (at least once) freely in φ belong in $\{\mathtt{x}_1, \ldots, \mathtt{x}_n\}$, i.e. $\mathrm{Free}(\varphi) \subseteq \{\mathtt{x}_1, \ldots, \mathtt{x}_n\}$ (yes, some of the $\mathtt{x}_i$ needn't occur freely or occur at all in φ). Then given a formula $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ and $a_1, \ldots, a_n \in A$, we define/notate $\mathcal{A} \models \varphi(a_1, \ldots, a_n) \;:\Longleftrightarrow\; \mathcal{A} \models \varphi[a_1, \ldots, a_n] \;:\Longleftrightarrow\; \mathcal{A} \models \varphi[\alpha]$

for any $\alpha : \mathscr{V} \to A$ with $\alpha(\mathtt{x}_i) = a_i$ (above Prop. 2.2 tells us this is well-defined definition).

In particular when φ is a sentence (i.e. a formula with no free variables $\mathrm{Free}(\varphi) = \varnothing$), then $\mathcal{A} \models \varphi \;:\Longleftrightarrow\; \mathcal{A} \models \varphi[\alpha]$ for any $\alpha : \mathscr{V} \to A \iff \mathcal{A} \models \varphi[\alpha]$ for some $\alpha : \mathscr{V} \to A$. This is read "$\varphi$ holds in $\mathcal{A}$" or "φ is satisfied/true in structure $\mathcal{A}$" or "$\mathcal{A}$ is a model of φ". Artem: "Never realized how many ways of saying this there are! In the literature there may be more ways/combinations of these ways of saying it,"

Thus, $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ defines a

Definition 2.9: Definable relation (in structure $\mathcal{A}$)

An n-ary relation on A given by $\{(a_1, \ldots, a_n) \in A^n : \mathcal{A} \models \varphi[a_1, \ldots, a_n]\}$, denoted $\varphi[\mathcal{A}]$.

3 Lecture 3

Last time we discussed satisfaction $\mathcal{A} \models \varphi[\alpha]$, and notation $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ to emphasize the free variables of $\varphi \subseteq \{\mathtt{x}_1, \ldots, \mathtt{x}_n\}$. Given $\mathscr{L}$-formula $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ we also defined definable (in $\mathcal{A}$) n-ary relations $\varphi[\mathcal{A}] := \{(a_1, \ldots, a_n) \in A^n : \mathcal{A} \models \varphi[a_1, \ldots, a_n]\}$. We can think of these relations purely as sets (subsets of A^n) too:

Definition 3.1: Definable sets

Let $\mathcal{A}$ be an $\mathscr{L}$-structure and $D \subseteq A^n$.

(a) the set D is called $\varnothing$-definable or 0-definable if $D = \varphi[\mathcal{A}]$ for some $\mathscr{L}$-formula $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$.

(b) let $B \subseteq A$. The set D is called B-definable or definable over B if there is $m \in \mathbb{N}^+, (b_1, \ldots, b_m) \in B^m$ s.t. $D = \varphi[\mathcal{A}, b_1, \ldots, b_m] := \{(a_1, \ldots, a_n) \in A^n : \mathcal{A} \models \varphi[a_1, \ldots, a_n, b_1, \ldots, b_m]\}$ for some $\mathscr{L}$-formula $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n, \mathtt{y}_1, \ldots, \mathtt{y}_m)$. So we allow some parameters to be used to define a set.

In many fields of math, want to fix category of sets (algebraic varieties, manifolds, etc.). The definable sets are the relevant category of sets for us; we will see that this category of sets has nice properties. "These are sets accessible through the logic." Only countably many formulas, so only really can work with sets defined by these formulas. This will come up mostly later in the course.

Example 3.1: Examples of definable sets

- Graphs (binary symmetric relation, irreflexive) are all structures in the language $\mathscr{L}_{\text{graph}} := \{\mathtt{E}\}$ where $E \in \mathscr{R}_2^{\mathscr{L}}$ satisfying the sentence $\varphi := \forall\mathtt{x}\forall\mathtt{y}\mathtt{Exy} \leftrightarrow \mathtt{Eyx} \wedge \forall\mathtt{x}\neg\mathtt{Exx}$. So graphs are axiomatized by this sentence (and in fact axiomatizability will be central question for us).
- In the $\mathscr{L}_{\text{ring}}$-structure $(\mathbb{N}, 0, 1, +, \cdot)$, the set of primes is $\varnothing$-definable subset by the formula $\varphi(\mathtt{x}) := \forall\mathtt{y}(\exists\mathtt{z}\ \mathtt{y}\cdot\mathtt{z}{=}\mathtt{x} \to (\mathtt{y}{=}1 \vee \mathtt{y}{=}\mathtt{x}))$.
- The integers/natural numbers $\mathbb{N} \subseteq \mathbb{R}$ are not definable in $\mathscr{L}_{\text{ring}}$ or $\mathscr{L}_{\text{ord.ring}}$ (not obvious). (Note language is important, since can e.g. add unary relation saying when real number is natural!) This is very important result, because if integers definable, Gödel incompleteness kicks in, and definable sets become too complicated to understand ("Gödelian phenomenon"). Later we will see Tarski-Seidenberg theorem using quantifier elimination to show that definable sets in real closed fields in $\mathscr{L}_{\text{ord.ring}}$ are nice.

Questions of definable sets is a rich question, and we will investigate more later in course.

We now discuss the substitution of terms in place of variables. I think the most natural perspective is the semantic viewpoint: given a term $t(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ or formula $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ and a $\mathcal{A}$-assignment α, we already know how to "plug in" the value $\alpha(\mathtt{x}_i) \in A$ to the variable $\mathtt{x}_i$ to produce (resp. in the case of term t or formula φ) the evaluation of (perhaps nested) function calls on the values $\{\alpha(\mathtt{x}_1), \ldots, \alpha(\mathtt{x}_n)\} \cup \{\mathtt{c}^{\mathcal{A}}\}_{\mathtt{c} \in \mathscr{C}^{\mathscr{L}}}$ or a true/false statement regarding the values $\{\alpha(\mathtt{x}_1), \ldots, \alpha(\mathtt{x}_n)\} \cup \{\mathtt{c}^{\mathcal{A}}\}_{\mathtt{c} \in \mathscr{C}^{\mathscr{L}}}$. What if we are concerned with/considering different values, e.g. the interpretation of some terms $s_1, \ldots, s_n$ we understand/want to understand $\equiv$ evaluation of some understood/particular functions on the "known/familiar" values $\{\alpha(\mathtt{x}_1), \ldots, \alpha(\mathtt{x}_n)\} \cup \{\mathtt{c}^{\mathcal{A}}\}_{\mathtt{c} \in \mathscr{C}^{\mathscr{L}}}$? Then we can "plug in" these values into t or φ in place of the $\mathtt{x}_i$ by using the assignment $\alpha_{s_1^{\mathcal{A}}[\alpha]/\mathtt{x}_1, \ldots, s_n^{\mathcal{A}}[\alpha]/\mathtt{x}_n}$. We can now ask whether there is a purely syntactic formula of this, i.e. some term $t_{s_1/\mathtt{x}_1,\ldots,s_n/\mathtt{x}_n}$ or formula $\varphi_{s_1/\mathtt{x}_1,\ldots,s_n/\mathtt{x}_n}$ "substituting s_i in place of $\mathtt{x}_i$" s.t. $t^{\mathcal{A}}_{\overline{s}/\overline{\mathtt{x}}}[\alpha] = t^{\mathcal{A}}[\alpha_{s_1^{\mathcal{A}}[\alpha]/\mathtt{x}_1,\ldots,s_n^{\mathcal{A}}[\alpha]/\mathtt{x}_n}]$ or $\varphi_{\overline{s}/\overline{\mathtt{x}}}[\alpha] = \varphi[\alpha_{s_1^{\mathcal{A}}[\alpha]/\mathtt{x}_1,\ldots,s_n^{\mathcal{A}}[\alpha]/\mathtt{x}_n}]$ (in the sense of mutual satisfiability by $\mathcal{A}$). In summary, we want to define substitution so that the following lemma holds:

Lemma 3.1: Substitution lemma

Let $\mathsf{x}_0, \ldots, \mathsf{x}_r$ be distinct variables, α an $\mathcal{A}$-assignment. Then,

- for every term t, $t^{\mathcal{A}}_{\overline{s}/\overline{\mathsf{x}}}[\alpha] = t^{\mathcal{A}}[\alpha_{s_0^{\mathcal{A}}[\alpha]/\mathsf{x}_0,\ldots,s_r^{\mathcal{A}}[\alpha]/\mathsf{x}_r}]$
- for every formula φ, $\mathcal{A} \models \varphi_{\overline{s}/\overline{\mathsf{x}}}[\alpha] \iff \mathcal{A} \models \varphi[\alpha_{s_0^{\mathcal{A}}[\alpha]/\mathsf{x}_0,\ldots,s_r^{\mathcal{A}}[\alpha]/\mathsf{x}_r}]$

Note if term t is of the form $t(x_0, \ldots, x_r)$ then $t^{\mathcal{A}}_{\overline{s}/\overline{\mathsf{x}}}[\alpha] = t^{\mathcal{A}}(s_0^{\mathcal{A}}[\alpha], \ldots, s_r^{\mathcal{A}}[\alpha])$, a "composition of functions" (as in notation of Def. 2.6 and Def. 2.8) — yes, could have $s_i = \mathsf{x}_i$ "unchanged", which would leave semantics on both sides the same since $\mathsf{x}_i^{\mathcal{A}}[\alpha] := \alpha(\mathsf{x}_i)$. Phrased another way, we want the syntactic composition to match the semantic composition of terms/formulas thought of as functions, i.e. commutativity of term/formula composition/substitution and semantic interpretation of syntax.

We focus on formulas. Thus the goal naively is to replace all free instances of $\mathtt{x}_i$ in φ by s_i. We must in fact replace for all $i \in [n]^+$ simultaneously because it's possible $\mathtt{x}_2$ appears in s_1, and replacing one i at a time would result in a s_2 term in an s_1 term, which is not what we want. There is fiddlier issue though: it's possible that some variable in term s becomes bound by a quantifier in the resulting formula. For example, given a formula $\varphi(\mathtt{v}_0) := \exists \mathtt{v}_1 \neg \mathtt{v}_1 = \mathtt{v}_0$ and term $s := \mathtt{v}_1$, the resulting substituted formula is $\exists \mathtt{v}_1 \neg \mathtt{v}_1 = \mathtt{v}_1$, which is satisfied in no structure while OTOH $|A| \geqslant 2 \rightsquigarrow \mathcal{A} \models \varphi[\alpha]$ for any $a \in A$ (i.e. this botched substitution changed semantic meaning drastically for no reason).

Artem: "Once we have behavior we want, we can work backwards and cook up the correct syntactic definition. However, I don't want to present it in class because it's quite tedious. Recommend read it in textbook. Can answer questions next time." Although the details are long to write out, the idea is very simple: the only issue crops up if quantified variable $\mathtt{x}_i$ appears in some s_j; the fix is to rename the quantified variable (choose the first variable in $\mathscr{V}$ not already used).

3.1 Limbo

I will just continue typing up my handwritten YouTube notes.

Definition 3.2: Simultaneous substitution of terms into variables

Let $\mathtt{x}_0, \ldots, \mathtt{x}_r$ be distinct variables, and $s_0, \ldots, s_r$ terms. We define the simultaneous substitution of the $\mathtt{x}_i$ by the s_i by:

(a) let t be a term (cf. unique reading of terms Prop. 1.1). Then $t_{s_0/\mathsf{x}_0,\ldots,s_r/\mathsf{x}_r}$, also denoted/abbreviated $t_{\overline{s}/\overline{\mathsf{x}}}$, is the word (defined inductive):

- for a variable $\mathsf{x} \in \mathscr{V}$, $\mathsf{x}_{\overline{s}/\overline{\mathsf{x}}} := \begin{cases} \mathsf{x} & \text{if } \mathsf{x} \neq \mathsf{x}_0 \text{ or } \ldots \text{ or } \mathsf{x} \neq \mathsf{x}_r \\ s_i & \text{if } \mathsf{x} = \mathsf{x}_i \end{cases}$
- $\mathsf{c}_{\overline{s}/\overline{\mathsf{x}}} := \mathsf{c}$ for all $\mathsf{c} \in \mathscr{C}^{\mathscr{L}}$.
- $[\mathsf{f}t^1 \ldots t^n]_{\overline{s}/\overline{\mathsf{x}}} := \mathsf{f}t^1_{\overline{s}/\overline{\mathsf{x}}} \ldots t^n_{\overline{s}/\overline{\mathsf{x}}}$

(b) for formulas,

- $[t = t']_{\overline{s}/\overline{\mathsf{x}}}$ is $t_{\overline{s}/\overline{\mathsf{x}}} = t'_{\overline{s}/\overline{\mathsf{x}}}$.

- $[\mathsf{R}t^1 \dots t^n]_{\overline{s}/\overline{\mathsf{x}}}$ is $\mathsf{R}t^1_{\overline{s}/\overline{\mathsf{x}}} \dots t^n_{\overline{s}/\overline{\mathsf{x}}}$.
- $[\neg\varphi]_{\overline{s}/\overline{\mathsf{x}}}$ is $\neg[\varphi]_{\overline{s}/\overline{\mathsf{x}}}$.
- $[\psi \wedge \chi]_{\overline{s}/\overline{\mathsf{x}}}$ is $\psi_{\overline{s}/\overline{\mathsf{x}}} \wedge \chi_{\overline{s}/\overline{\mathsf{x}}}$.
- the only remaining possibility is that the formula is of the form $\exists \mathsf{x}\psi$. Let $\mathsf{x}_{i_1}, \dots, \mathsf{x}_{i_k}$ ($i_1 < \dots < i_k$) be variables in $x_0, \dots, x_r$ that are free in $\exists\mathsf{x}\psi$ (in particular, $\mathsf{x} \neq \mathsf{x}_{i_j}$ for any $j \in [k]^+$). There are two further (mutually exclusive) possibilities:
  - if x doesn't occur in any of $s_{i_1}, \dots, s_{i_k}$, define $[\exists\mathsf{x}\psi]_{\overline{s}/\overline{\mathsf{x}}}$ to be $\exists\mathsf{x}[\psi]_{s_{i_1}/\mathsf{x}_{i_1},\dots,s_{i_k}/\mathsf{x}_{i_k}}$.
  - otherwise define to be define it to be $\exists\mathsf{u}[\psi]_{s_{i_1}/\mathsf{x}_{i_1},\dots,s_{i_k}/\mathsf{x}_{i_k},\mathsf{u}/\mathsf{x}}$ (simultaneous substitution so all the x's in original ψ replaced by u's, but not the x's in the s_i terms), where u is the first variable in the enumeration $\mathsf{v}_1, \mathsf{v}_2, \dots$ not occurring in any of the words $\exists\mathsf{x}\psi$, $s_{i_1}, \dots, s_{i_k}$.

It's a long definition, but basically everything is routine/sensible/obvious, except the very last part of the existential quantifier case.

Remark: must check (exercise): if t term, $\overline{s}$ tuple of terms, then $t_{\overline{s}/\overline{\mathsf{x}}}$ is term still; and if φ formula, $\overline{s}$ tuple of terms, then $\varphi_{\overline{s}/\overline{\mathsf{x}}}$ is formula still, and $\mathrm{ht}(\varphi_{\overline{s}/\overline{\mathsf{x}}}) = \mathrm{ht}(\varphi)$. Proof by induction on height, base case verify by hand.

We already made this definition before, but here for the record

Definition 3.3: Assignment substitution

For $\mathsf{x}_0, \dots, \mathsf{x}_r$ distinct variables, α is $\mathcal{A}$-assignment, $a_0, \dots, a_r \in A$, define $\alpha_{a_0/\mathsf{x}_0,\dots,a_r/\mathsf{x}_r} =: \alpha_{\overline{a}/\overline{\mathsf{x}}}$ in routine manner.

We now prove the promised substitution lemma, which I now restate:

Lemma 3.1: Substitution lemma

Let $\mathsf{x}_0, \dots, \mathsf{x}_r$ be distinct variables, α an $\mathcal{A}$-assignment. Then,

- for every term t, $t^{\mathcal{A}}_{\overline{s}/\overline{\mathsf{x}}}[\alpha] = t^{\mathcal{A}}[\alpha_{s_0^{\mathcal{A}}[\alpha]/\mathsf{x}_0,\dots,s_r^{\mathcal{A}}[\alpha]/\mathsf{x}_r}]$
- for every formula φ, $\mathcal{A} \models \varphi_{\overline{s}/\overline{\mathsf{x}}}[\alpha] \iff \mathcal{A} \models \varphi[\alpha_{s_0^{\mathcal{A}}[\alpha]/\mathsf{x}_0,\dots,s_r^{\mathcal{A}}[\alpha]/\mathsf{x}_r}]$

Proof: (2) only, leave (1) as exercise. By induction on $\mathrm{ht}(\varphi)$. The atomic, negation, conjunction cases left as exercise. The only difficult case is if φ is an existential statement, i.e. φ equals $\exists\mathsf{x}\psi$. As before, $\mathsf{x}_{i_1}, \dots, \mathsf{x}_{i_k}$ denote variables in $\{\mathsf{x}_0, \dots, \mathsf{x}_r\}$ that are free in φ; in particular $\mathsf{x} \neq \mathsf{x}_{i_j}$ for any $j \in [k]^+$.

- Assume first that x occurs in $\geqslant 1$ of the terms $s_{i_1}, \dots, s_{i_k}$. Let u be the first variable (w.r.t. enumeration $\mathsf{v}_0, \dots$ of $\mathscr{V}$) which doesn't occur in any of the words $\varphi, s_{i_1}, \dots, s_{i_k}$. Then,

$$\begin{aligned}
\mathcal{A} \models (\exists x\psi)_{\overline{s}/\overline{x}}[\alpha] &\underset{\substack{\text{def. of}\\ \text{substitution}}}{\iff} \exists\mathsf{u}(\psi)_{s_{i_1}/\mathsf{x}_{i_1},\dots,s_{i_k}/\mathsf{x}_{i_k},\mathsf{u}/\mathsf{x}}[\alpha] \\
&\underset{\text{def. of} \models}{\iff} \text{exists } a \in A \text{ s.t. } \mathcal{A} \models \underbrace{\psi_{s_{i_1}/\mathsf{x}_{i_1},\, \dots,\, s_{i_k}/\mathsf{x}_{i_k},\, \mathsf{u}/\mathsf{x}}}_{\text{has height } < \mathrm{ht}(\exists\mathsf{x}\psi)}[\alpha_{a/\mathsf{u}}] \\
&\underset{\substack{\text{induction}\\ \text{hypo.}}}{\iff} \psi\Big[\alpha_{s_{i_1}^{\mathcal{A}}[\alpha_{a/\mathsf{u}}]/\mathsf{x}_{i_1},\, \dots,\, s_{i_k}^{\mathcal{A}}[\alpha_{a/\mathsf{u}}]/\mathsf{x}_{i_k},\, \underbrace{u^{\mathcal{A}}[\alpha_{a/\mathsf{u}}]}_{=a}/\mathsf{x},\, a/\mathsf{u}}\Big] \text{ (for some } a \in A)
\end{aligned}$$

(we do use that $(\alpha_{\overline{s}/\overline{x}})_{\overline{t}/\overline{y}} = \alpha_{\overline{s}/\overline{x},\overline{t}/\overline{y}}$ if $\overline{x},\overline{y}$ are distinct variables)

$$\underset{\substack{\text{because}\\ s_{i_j}^{\mathcal{A}}[\alpha]=s_{i_j}^{\mathcal{A}}[\alpha_{a/u}]}}{\iff} \psi\big[\alpha_{s_{i_1}^{\mathcal{A}}[\alpha]/\mathsf{x}_{i_1},\,\ldots,\,s_{i_k}^{\mathcal{A}}[\alpha]/\mathsf{x}_{i_k},\,a/\mathsf{x},\,a/\mathsf{u}}\big] \text{ (for some } a \in A)$$

(more explicitly, $s_{i_j}^{\mathcal{A}}[\alpha] = s_{i_j}^{\mathcal{A}}[\alpha_{a/\mathsf{u}}]$ because u doesn't occur (freely) in s_{i_j} so α and $\alpha_{a/\mathsf{u}}$ agree on $\mathrm{Free}(s_{i_j})$, so true by Prop. 2.2 from last lecture)

$$\underset{\text{def. of }\models}{\iff} \mathcal{A} \models \exists\mathsf{x}\psi\big[\alpha_{s_{i_1}^{\mathcal{A}}[\alpha]/\mathsf{x}_{i_1},\,\ldots,\,s_{i_k}^{\mathcal{A}}[\alpha]/\mathsf{x}_{i_k},\,a/\mathsf{u}}\big]$$

$$\underset{\text{Prop. 2.2}}{\iff} \mathcal{A} \models \exists\mathsf{x}\psi\big[\alpha_{s_0^{\mathcal{A}}[\alpha]/\mathsf{x}_0,\,\ldots,\,s_r^{\mathcal{A}}[\alpha]/\mathsf{x}_r}\big]$$

where the last equivalence comes from Prop. 2.2 which tells us that that it doesn't matter what we substitute/assign for variables not in $\mathrm{Free}(\exists\mathsf{x}\psi)$ — similarly u doesn't appear in ψ so the dangling a/u disappears

- If otherwise x does not occur in any of $s_{i_1},\ldots,s_{i_k}$, then (following much the same logic as before)

$$\mathcal{A} \models (\exists x\psi)_{\overline{s}/\overline{x}}[\alpha] \underset{\substack{\text{def. of}\\ \text{substitution}}}{\iff} \exists\mathsf{u}(\psi)_{s_{i_1}/\mathsf{x}_{i_1},\ldots,s_{i_k}/\mathsf{x}_{i_k},\mathsf{u}/\mathsf{x}}[\alpha]$$

$$\underset{\text{def. of }\models}{\iff} \text{exists } a \in A \text{ s.t. } \mathcal{A} \models \underbrace{\psi_{s_{i_1}/\mathsf{x}_{i_1},\,\ldots,\,s_{i_k}/\mathsf{x}_{i_k},\,\mathsf{u}/\mathsf{x}}}_{\text{has height } <\mathrm{ht}(\exists\mathsf{x}\psi)}[\alpha_{a/\mathsf{x}}]$$

$$\underset{\substack{\text{induction}\\ \text{hypo.}}}{\iff} \psi\big[\alpha_{s_{i_1}^{\mathcal{A}}[\alpha]/\mathsf{x}_{i_1},\,\ldots,\,s_{i_k}^{\mathcal{A}}[\alpha_{a/\mathsf{u}}]/\mathsf{x}_{i_k},\,a/\mathsf{x}}\big] \text{ (for some } a \in A)$$

$$\underset{\text{def. of }\models}{\iff} \mathcal{A} \models \exists\mathsf{x}\psi\big[\alpha_{s_{i_1}^{\mathcal{A}}[\alpha]/\mathsf{x}_{i_1},\,\ldots,\,s_{i_k}^{\mathcal{A}}[\alpha]/\mathsf{x}_{i_k}}\big]$$

$$\underset{\text{Prop. 2.2}}{\iff} \mathcal{A} \models \exists\mathsf{x}\psi\big[\alpha_{s_0^{\mathcal{A}}[\alpha]/\mathsf{x}_0,\,\ldots,\,s_r^{\mathcal{A}}[\alpha]/\mathsf{x}_r}\big]$$

This concludes the proof of (2). ■

10/3/22 Class 4

Example 3.2: Some substituted formulas

- Returning to $\varphi(\mathsf{v}_0) := \exists\mathsf{v}_1\ \neg\mathsf{v}_1{=}\mathsf{v}_0$, let $s := \mathsf{v}_1$. The variable u as denoted in the above proof is then v_2. Then φ_{s/v_0} is equal to $\big(\exists\mathsf{v}_2\ \neg\mathsf{v}_1{=}\mathsf{v}_0\big)_{\mathsf{v}_1/\mathsf{v}_0,\mathsf{v}_2/\mathsf{v}_1}$, which is equal to $\exists\mathsf{v}_2\ \neg\mathsf{v}_2{=}\mathsf{v}_1$.
- **Claim:** let φ be $\mathscr{L}$-formula, and $\mathsf{x}_1,\ldots,\mathsf{x}_n$ be distinct variables, with $s_1,\ldots,s_n$ terms not containing any variables that occur in φ. Then $\varphi_{\overline{s}/\overline{\mathsf{x}}}$ is the word obtained by replacing every occurence of x_i by s_i. *Proof:* by induction on height.
  - Again let us only consider $\varphi := \exists\mathsf{x}\psi$. Again let us have indices $i_1,\ldots,i_k$ as above. By assumption, x doesn't appear in any $s_{i_1},\ldots,s_{i_k}$. So $(\exists\mathsf{x}\psi)_{\overline{s}/\overline{\mathsf{x}}} := \exists\mathsf{x}(\psi)_{s_{i_1}/\mathsf{x}_1,\ldots,s_{i_k}/\mathsf{x}_k}$. As the s_{i_j} do not contain any variables occuring in ψ, formula $\psi_{s_{i_1}/\mathsf{x}_1,\ldots,s_{i_k}/\mathsf{x}_k}$ by induction hypothesis is given by replacement (as the **Claim** claimed). ■
- If y is variable with no occurrence in φ, then $[\varphi_{\mathsf{y}/\mathsf{x}}]_{\mathsf{x}/\mathsf{y}}$. *Proof:* again by induction on height.

■

4 Lecture 4

Recall (Lec 3): for formula φ, variables $\mathtt{x}_1, \ldots, \mathtt{x}_n$ and terms $s_1, \ldots, s_n$, we defined $\varphi_{\overline{s}/\overline{\mathtt{x}}}$ ("φ with tuple of terms $\overline{s}$ substituted in place of tuple of variables $\overline{\mathtt{x}}$"), which satisfies the semantic property: for all assignments $\alpha : \mathscr{V} \to A$,

$$\mathcal{A} \models \varphi_{\overline{s}/\overline{\mathtt{x}}}[\alpha] \iff \mathcal{A} \models \varphi[\alpha_{s_1^{\mathcal{A}}[\alpha]/\mathtt{x}_1, \ldots, s_n^{\mathcal{A}}[\alpha]/\mathtt{x}_n}]$$

As mentioned before, this is composition of functions (arising from formulas/terms)/commutativity of substitution and semantic evaluation... justifies our formal syntactic rule. To emphasize this semantic (composition of functions) meaning of substitution (and to avoid subscript notation "which sucks and gets very complicated"), may often use the notation

Definition 4.1: Substitution as function composition, notation

If $t(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ is $\mathscr{L}$-term and $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ is a $\mathscr{L}$-formula (recall this notation from Def. 2.6 and Def. 2.8), and $s_1, \ldots, s_n$ are $\mathscr{L}$-terms, write $t(s_1, \ldots, s_n)$ for $t_{\overline{s}/\overline{\mathtt{x}}}$ and similarly $\varphi(s_1, \ldots, s_n)$ for $\varphi_{\overline{s}/\overline{\mathtt{x}}}$

In the rest of this lecture, we discuss:

Definition 4.2: Universally valid formulas

An $\mathscr{L}$-formula φ is <u>universally valid</u>, written $\models \varphi$, if it is satisfied in **every** $\mathscr{L}$-structure $\mathcal{A}$ and by **every** $\mathcal{A}$-assignment $\alpha : \mathscr{V} \to A^{\mathcal{A}}$ (i.e. $\mathcal{A} \models \varphi[\alpha]$ for all $\mathcal{A}$ and all α).

Remark: a formula $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n)$ is universally valid iff the sentence ("the <u>universal closure of φ</u>") $\forall \mathtt{x}_1 \ldots \forall \mathtt{x}_n \varphi$ is universally valid. *Proof:* sentence valid means for all $\mathscr{L}$-structures $\mathcal{A}$ and all $a_1, \ldots, a_n \in A$, $\mathcal{A} \models \varphi(a_1, \ldots, a_n)$; which is exactly the same as $\mathcal{A} \models \varphi[\alpha]$ for every $\mathscr{L}$-structure $\mathcal{A}$ and every $\mathcal{A}$-assignment α.

Example 4.1: Some universally valid formulas

- $\exists \mathtt{x}\ \mathtt{x}=\mathtt{x}$ is universally valid (remember: by definition of $\mathscr{L}$-structure, base set is nonempty; this convention isn't so important for most intents and purposes, and it looks pretty natural to allow empty structure, but leads to some more edge cases so we stick with this convention).
- for φ, ψ $\mathscr{L}$-formulas, then $\varphi \to (\psi \to \varphi)$ is universally valid. Notice it doesn't depend on what φ, ψ are, nor their boolean truth values. Can cook up more examples like this, e.g. $\neg\neg\varphi \leftrightarrow \varphi$. Purpose of this lecture is to understand formulas of this nature.

Note: the notation $\models \varphi$ does not mention language $\mathscr{L}$ anywhere (as opposed to $\mathcal{A} \models \varphi$, in which $\mathscr{L}$ is implicitly mentioned in $\mathscr{L}$-structure $\mathcal{A}$). This is important since *a priori* there may be more $\mathscr{L}$'-structures than $\mathscr{L}$-structures (take reduct, defined below). The notation is justified by the following lemma (first a definition):

Definition 4.3: Reduction and expansion

- Let $\mathscr{L} \subseteq \mathscr{L}'$ be two languages. Given $\mathscr{L}'$-structure $\mathcal{A}' := (A; (Z^{\mathcal{A}'})_{Z \in \sigma^{\mathscr{L}'}})$, the <u>reduct of $\mathcal{A}'$ to $\mathscr{L}$</u> or <u>$\mathscr{L}$-reduct of $\mathcal{A}$'</u> is the $\mathscr{L}$-structure $\mathcal{A} := \mathcal{A}' \upharpoonright_{\mathscr{L}} := (A; (Z^{\mathcal{A}'})_{Z \in \sigma^{\mathscr{L}}})$.

- Here $\mathcal{A}'$ is an expansion of $\mathcal{A}$ to $\mathscr{L}'$. More generally, one can expand $\mathscr{L}$-structure $\mathcal{A}$ to $\mathscr{L}'$-structure $\mathcal{A}'$ by defining $Z^{\mathcal{A}'} := Z^{\mathcal{A}}$ for all $Z \in \sigma^{\mathscr{L}}$, and the remaining non-logical symbols (constants, functions, relations) defined arbitrarily.

Lemma 4.1: Universal validity independent of language expansion

Let φ be $\mathscr{L}$-formula, and $\mathscr{L}' \supseteq \mathscr{L}$ be an arbitrary language extending $\mathscr{L}$. Then, φ is universally valid as an $\mathscr{L}$-formula $\iff$ it is universally valid as an $\mathscr{L}'$-formula

Proof: ($\implies$): let $\mathcal{A}'$ be arbitrary $\mathscr{L}'$-structure, and $\alpha : \mathscr{V} \to A^{\mathcal{A}'}$ be any $\mathcal{A}'$-assignment. Let $\mathcal{A} := \mathcal{A}'\restriction_{\mathscr{L}}$ be the $\mathscr{L}$-reduct of $\mathcal{A}'$. Variables $\mathscr{V}$ independent of language, and reduct means base sets of $\mathcal{A}$ and $\mathcal{A}'$ are the same, so α is also $\mathcal{A}$-assignment. Because φ is $\mathscr{L}$-formula and $\mathcal{A}$ is restriction of $\mathcal{A}'$, $Z^{\mathcal{A}}$ agrees with $Z^{\mathcal{A}'}$ for all $Z \in \sigma^{\mathscr{L}}$, so (can verify rigorously by induction) $\mathcal{A}' \models \varphi[\alpha] \iff \mathcal{A} \models \varphi[\alpha]$, so indeed all such $\mathcal{A}'$ and α satisfy φ.
($\impliedby$): essentially the same, but need to expand from $\mathscr{L}$-structure $\mathcal{A}$ to $\mathscr{L}'$-structure $\mathcal{A}'$, as defined above (i.e. $Z^{\mathcal{A}'} := Z^{\mathcal{A}}$ for all $Z \in \sigma^{\mathscr{L}}$, and the rest (constants/functions/relations) set arbitrarily.) ∎

So universal validity doesn't depend on language, as long as language contains all the symbols appearing in the formula. Artem: "may be trivial, but good to go through it once".

We now study/build a framework for a particular simple class of universally valid formulas: the formulas like $\varphi \to (\psi \to \varphi)$ or $\varphi \wedge \neg\varphi$ which are always satisfied independent of φ, ψ.

Definition 4.4: Propositional calculus

- Fix a set $\mathscr{P} := \{\mathtt{p}_i, i \in \mathbb{N}^+\}$ called the propositional variable ("propositional" to indicate only take values in $\{0,1\}$, as opposed to arbitrary base set).
- A propositional formula is a word over the alphabet $\mathscr{P} \cup \{\neg, \wedge, (,)\}$ formed by the following inductive rules:
  - all $\mathtt{p}_i \in \mathscr{P}$ are propositional formulas
  - if F, G are propositional formulas, then $\neg F$ and $(F \wedge G)$ are propositional formulas.
- As before, we define $\mathscr{Fml}_{\mathscr{P}}$ to be the set of all propositional formulas, and we use the same abbreviations ($\vee, \to, \leftrightarrow$, etc.). Truth values are $\{0,1\}$, where we identify "false" with 0, and "true" with 1 (could take vice versa, doesn't matter, but this is convention).
- An assignment is a function $\delta : \mathscr{P} \to \{0,1\}$. Any assignment induces a function $\delta^* : \mathscr{Fml}_{\mathscr{P}} \to \{0,1\}$, defined inductively by
  - $\delta^*(\mathtt{p}_i) := \delta(\mathtt{p}_i)$ for all $i \in \mathbb{N}$;
  - $\delta^*(\neg F) := 1 - \delta^*(F)$;
  - $\delta^*(F \wedge G) := \delta^*(F) \cdot \delta^*(G)$.

  If $\delta^*(F) = 1$, we write $\delta \models F$. Similar to before, we write F as $F(\mathtt{q}_1, \dots, \mathtt{q}_n)$ if $\mathtt{q}_i$ are all distinct propositional variables, and all variables occurring in F lie in $\{\mathtt{q}_1, \dots, \mathtt{q}_n\}$.

My remark: this definition very similar to how we define satisfaction but we don't have to deal with non-logical symbols (i.e. those in $\sigma^{\mathscr{L}}$); so purely logical structure. Artem: "a simpler version of logic. Instead of arbitrary structures, only take values in $\{0,1\}$. Simplify satisfaction to calculations in a boolean ring".

Remark: the boolean operations we allowed ($\wedge, \neg$) give complete set of all boolean functions. More precisely, for all functions $g : \{0,1\}^n \to \{0,1\}$, there exists a propositional formula $F(\mathtt{p}_1, \ldots, \mathtt{p}_n)$ s.t. for any assignment $\delta : \{\mathtt{p}_1, \ldots, \mathtt{p}_n\} \to \{0,1\}$ we get $\delta^*(F) = g(\delta(\mathtt{p}_1), \ldots, \delta(\mathtt{p}_n))$.

Definition 4.5: Tautology

(a) A propositional formula F is a tautology if $\delta \models F$ for all assignments δ (i.e. $F(\mathtt{q}_1, \ldots, \mathtt{q}_n)$ is true for any choice of truth values of the $\mathtt{q}_i$).

(b) An $\mathscr{L}$-formula φ is a tautology if there exists a propositional formula $F(\mathtt{q}_1, \ldots, \mathtt{q}_n) \in \mathscr{Fml}_{\mathscr{P}}$ that is tautology (defined previously) and also $\mathscr{L}$-formulas $\psi_1, \ldots, \psi_n$ s.t. φ equals $F_{\psi_1/\mathtt{q}_1, \ldots, \psi_n/\mathtt{q}_n}$ (by which we mean naive substitution, i.e. the word obtained by replacing every occurrence of $\mathtt{q}_i$ in F by ψ_i).

Example 4.2: Propositional tautology and formula

$F :=$ $(\mathtt{p}_1 \to (\mathtt{p}_2 \to \mathtt{p}_1))$ is a propositional tautology. Then for any $\mathscr{L}$-formulas ψ_1, ψ_2, the $\mathscr{L}$-formula $(\psi_1 \to (\psi_2 \to \psi_1))$ is universally valid.

More generally,

Lemma 4.2: Tautologies are universally valid

All $\mathscr{L}$-formulas φ are universally valid (in every $\mathscr{L}$-structure, by Lemma 4.1)

Proof: because F is tautology, no matter the truth values of $\mathcal{A} \models \psi_i[\alpha]$, $\mathcal{A} \models F(\psi_1, \ldots, \psi_n)[\alpha]$ must be true — the way $(\mathcal{A}, \alpha)$-assignment and δ^*-assignment inductively define truth of conjunction/negation statements are the same, so for any propositional formula F and $\varphi := F(\mathtt{q}_1, \ldots, \mathtt{q}_n)_{\psi_1/\mathtt{q}_1, \ldots, \psi_n/\mathtt{q}_n}$, and $\delta(\mathtt{q}_i) := 1$ iff $\mathcal{A} \models \psi_i[\alpha]$ (0 otherwise), we have that $\mathcal{A} \models \varphi[\alpha]$ is true iff $\delta^*(F) = 1$. ■

"So we have simple general method for generating universally valid formulas". Our goal then is to find **all** universally valid formulas, starting with basic known universally valid formulas (e.g. tautologies). This will lead to the development of formal proof, and ways of generating new univ. valid formulas from known ones via (two) basic rules. "Want to find small generating set of univ. valid formulas, following basic generating rules".

10/5/22 Class 5

Recall our goal is to find a "simple" set (as simple as possible, of course) of logical axioms and simple set of deduction rules, so that any universally valid sentence can be deduced from them using these rules. The idea is to pick some sentences that are "obviously"/"intuitive" universally valid sentences. And deduction rules should definitely preserve universally validity ("soundness"). Some freedom in this (not canonical), many variants in different books. Again, this brings up the theme of connecting semantic and syntax: universal validity is a semantic property, but applying formal deduction rules to fixed set of logical axioms is a syntactic process.

Along with the FO logic tautologies already discussed, which are really only concerned with $\{\neg, \wedge\}$, we also need axioms to control how to manipulate equality, functions/relations, and quantifiers syntactically, in a way consistent to the way we think equality/quantifiers should act (from our experience doing/using mathematics/FO logic).

My remarks: perhaps if one wanted, one could comb through all mathematical proofs we've done previously (on an intuitive level/without the formal rules) and isolate exactly what we used about equality and quantifiers on a syntactic level. Or perhaps more feasibly, and more in context of our eventual goal of Gödel's completeness theorem, we can define naïvely our set of logical axioms to be all universally valid formulas, then prove Gödel, and then extract only the universally valid formulas that we used in the proof (and from there extract a minimal set).

Definition 4.6: Equality axioms

The following sentences are universally valid (we will see in future while studying formal provability that these are exactly the properties we need to use about equality):

(E1) $\forall \mathtt{v}_0\ \mathtt{v}_0 = \mathtt{v}_0$ ("reflexivity")

(E2) $\forall \mathtt{v}_0 \forall \mathtt{v}_1\ (\mathtt{v}_0 = \mathtt{v}_1 \to \mathtt{v}_1 = \mathtt{v}_0)$ ("symmetry")

(E3) $\forall \mathtt{v}_0 \forall \mathtt{v}_1\ ((\mathtt{v}_0 = \mathtt{v}_1) \wedge (\mathtt{v}_1 = \mathtt{v}_2) \to \mathtt{v}_0 = \mathtt{v}_2)$ ("transitivity")

The first three axioms just express that = is an equivalence relation. Obviously universally valid because in any structure, = is interpreted as literal equality of elements of sets, which is an equivalence relation.

(E4) $\forall \mathtt{v}_1 \ldots \forall \mathtt{v}_{2n} (\bigwedge_{i=1}^{n} \mathtt{v}_i = \mathtt{v}_{i+n} \to \mathsf{f}\mathtt{v}_1 \ldots \mathtt{v}_n = \mathsf{f}\mathtt{v}_{n+1} \ldots \mathtt{v}_{2n})$ for any $\mathsf{f} \in \mathscr{F}_n^{\mathscr{L}}$ ("congruence for fns.")

(E5) $\forall \mathtt{v}_1 \ldots \forall \mathtt{v}_{2n} ((\bigwedge_{i=1}^{n} \mathtt{v}_i = \mathtt{v}_{i+n} \wedge \mathsf{R}\mathtt{v}_1 \ldots \mathtt{v}_n) \to \mathsf{R}\mathtt{v}_{n+1} \ldots \mathtt{v}_{2n})$, $\forall \mathsf{R} \in \mathscr{R}_n^{\mathscr{L}}$. ("cong. for rels.")

Again universally valid because = interpreted as literal equality of elements in set, and function $A^n \to A$ or relation/subset $\subseteq A$ acts the same on elements that are literally equal in the set. So these axioms (E1)-(E5) express that = is an equivalence relation, and a congruence with respect to functions and relations.

Remark: what if we didn't want to embed literal equality in axioms of structure? Not possible, since structure can not distinguish between literal equality and just congruence/equivalence relation.

Definition 4.7: Quantifier axioms

Let φ, ψ be arbitrary $\mathscr{L}$-formulas:

(Q1) for all variables $\mathtt{x}$ not free in φ, (Q1) is the formula $\forall \mathtt{x}(\varphi \to \psi) \to (\varphi \to \forall \mathtt{x}\psi)$;

(Q2) for all variables $\mathtt{x}$ and every term t, (Q2) is the formula $\varphi_{t/\mathtt{x}} \to \exists \mathtt{x}\varphi$;

(Q3) for all variables $\mathtt{x}$, (Q3) is the formula $\exists \mathtt{x}\varphi \leftrightarrow \neg\forall \mathtt{x}\neg\varphi$.

Note (Q1)-(Q3) depend on φ, ψ, ranging over all φ, ψ there are in fact infinitely many of each quantifier axiom.

Lemma 4.3: Quantifier axioms are universally valid

The quantifier axioms (Q1)-(Q3) defined above are universally valid.

Proof: we prove in reverse order of definition:

(Q3) clear from the definition of the abbreviation $\forall \mathsf{v}_i$ of $\neg\exists\neg\mathsf{v}_i$ (the negations cancel out)

(Q2) the substitution lemma (Lem. 3.1) tells us for any assignment α in arbitrary structure $\mathcal{A}\mathcal{A} \models \varphi_{t/\mathsf{x}}[\alpha] \iff \mathcal{A} \models \varphi[\alpha_{t^{\mathcal{A}}[\alpha]/\mathsf{x}}]$, which by definition means $\mathcal{A} \models \exists \mathsf{x}\varphi[\alpha]$ (with witness $t^{\mathcal{A}}[\alpha] \in A$).

(Q1) suppose $\mathsf{x} \notin \mathrm{Free}(\varphi)$, and $\mathcal{A} \models (\forall \mathsf{x}(\varphi \to \psi))[\alpha]$. Unwinding all the notation/abbreviations, this is $\mathcal{A} \models (\neg\exists\mathsf{x}\neg\neg(\varphi \land \psi))[\alpha]$, i.e. $\mathcal{A} \models (\neg\varphi \lor \psi)[\alpha_{a/\mathsf{x}}]$ for all $a \in A$. We have to show $\mathcal{A} \models (\varphi \to \forall\mathsf{x}\psi)[\alpha_{a/\mathsf{x}}]$, or equivalently $\mathcal{A} \models (\neg\varphi \lor \forall\mathsf{x}\psi)[\alpha]$; from here, it is clear that it suffices to show that assuming $\mathcal{A} \models \varphi[\alpha]$, it is true that $\mathcal{A} \models \forall\mathsf{x}\psi[\alpha]$.

Well, note that $\mathsf{x} \notin \mathrm{Free}(\varphi)$, so $\mathcal{A} \models \varphi[\alpha]$ leads to $\mathcal{A} \models \varphi[\alpha_{a/\mathsf{x}}]$ for any $a \in A$, but above we saw that $\mathcal{A} \models (\neg\varphi \lor \psi)[\alpha_{a/\mathsf{x}}]$ for all $a \in A$, and the only way for all these statements to hold or be true/satisfied is if $\mathcal{A} \models \psi[\alpha_{a/\mathsf{x}}]$ for all $a \in A$. By definition, this is exactly $\mathcal{A} \models \forall\mathsf{x}\psi[\alpha]$.

That concludes the lemma. ■

Remark: restriction of $\mathsf{x} \notin \mathrm{Free}(\varphi)$ is necessary, consider e.g. the case $\varphi, \psi := \mathsf{x}{=}\mathsf{c}$ for some constant $\mathsf{c} \in \mathscr{C}^{\mathscr{L}}$ and variable x. Then, $\forall\mathsf{x}(\varphi \to \psi)$ is universally valid, but in any structure $\mathcal{A}$ with base set A with $\geqslant 2$ elements, $\mathcal{A} \not\models \varphi \to \forall\mathsf{x}\psi$.

5 Lecture 5

We have just concocted some axioms that are "self-evident"/should hold in all FO structures, and we put them all together in the following definition. (*My remark (added 10/30/22):* that these few axioms "suffice" for a "good" formal proof system is justified *a postieri* by Gödel's completeness theorem; one can imagine proving Gödel's theorem without having defined a formal proof system, but instead keeping careful track of all syntactic manipulations used during the proof (syntactic manipulations corresponding to semantically true things, of course), and then defining a formal proof system to be exactly what was needed to prove Gödel.):

Definition 5.1: Logical axioms

The logical axioms are the (disjoint) union of the sets of

- first-order tautologies (obtained through propositional tautologies via substituting formulas in place of propositional variables);
- equality axioms (E1)-(E5) (universally valid first-order sentences expressing that equality is equivalence relation respecting functions and relations);
- and quantifier axioms (Q1)-(Q3) (universally valid formulas expressing basic properties of $\exists, \forall$).

Remark: we don't talk about computability in 220A, but the key feature of these axioms are that they are recursive.

Definition 5.2: Deduction rules

- Modus Ponens (MP): from φ and $\varphi \to \psi$, one deduces/can deduce ψ.
- Generalization (Gen): from φ, one can deduce $\forall \mathsf{x} \varphi$ (for any variable x).

My Remark/Warning (10/17/22): (Gen) is really so our proof system can talk about $\mathscr{L}$-formulas instead of just $\mathscr{L}$-sentences, even though basically everything we work with normally in math is a sentence, not formula with free variables (we always automatically mean to put universal quantifiers in front, if there are "free variables", which is exactly what (Gen) allows us to do syntactically). The extremely subtle point (at least to me), which I cleared up today with Clark (on my lunch break; even asking Artem at beginning of class he couldn't explain where I went wrong) is that (Gen) is NOT that $T \vdash_{\mathscr{L}} (\varphi \to \forall \mathsf{x} \varphi)$, but instead that $T \vdash_{\mathscr{L}} \varphi$ implies $T \vdash_{\mathscr{L}} \forall \mathsf{x} \varphi$ (for any $\mathscr{L}$-theory T, using notation from below). Basically (MP) gives us that implication at the level of formal syntax leads to implication at the meta-level (w.r.t. a given "background reference frame" T); but NOT THE OTHER WAY AROUND!

Remark: as promised at the beginning of class, all universally valid $\mathscr{L}$-formulas may be obtained by finite deduction/formal proof from logical axioms, establishing perfect correspondence between semantic truth, and syntactic provability — this is a corollary of Gödel's completeness theorem.

Definition 5.3: Theories

An $\mathscr{L}$-theory T is a set of $\mathscr{L}$-sentences.

Remark: some literature requires theory to be consistent; we do not.

Definition 5.4: Formal proofs and provability

- Let φ be an $\mathscr{L}$-formula, and T be $\mathscr{L}$-theory. A formal proof of φ in T (of length n) is finite sequence of $\mathscr{L}$-formulas $\varphi_1, \ldots, \varphi_n$ s.t. $\varphi_n = \varphi$, and s.t. for all $i \leqslant n$, EITHER $\varphi_i \in T$, OR φ_i is logical axiom, OR φ can be deduced by one of deduction axioms: from (MP) applied to φ_j, φ_k for $j, k < i$; or from (Gen) applied to φ_j for $j < i$.
- We say that φ is provable in T, denoted $T \vdash_{\mathscr{L}} \varphi$, if there exists a formal proof of φ in T.
- We write $\varnothing \vdash_{\mathscr{L}} \varphi$, or even just $\vdash_{\mathscr{L}} \varphi$, if φ is provable in the empty theory (i.e. from the logical axioms and two deduction rules alone).

Returning to purely semantic properties, we define

Definition 5.5: Models and logical consequence

- An $\mathscr{L}$-structure $\mathcal{A}$ is a ($\mathscr{L}$-)model of T, denoted $\mathcal{A} \models T$, if $\mathcal{A} \models \varphi$ (no α specified because φ is sentence) for all $\varphi \in T$.
- A sentence (or formula) φ is a logical consequence of T, denoted $T \models \varphi$, if for every ($\mathscr{L}$-)model $\mathcal{A}$ of T, $\mathcal{A} \models \varphi$ (for formulas $\varphi(\mathsf{x}_1, \ldots, \mathsf{x}_n)$, instead need $\mathcal{A} \models \forall \mathsf{x}_1 \ldots \forall \mathsf{x}_n \varphi$).

As with univ. validity (Lem. 4.1), $T \models \varphi$ is indep. of $\mathscr{L}$ (as long as φ can be written in $\mathscr{L}$).

Remark: a priori, $\vdash_{\mathscr{L}}$ depends on language $\mathscr{L}$ in which φ, T are considered, because if $\mathscr{L}' \supseteq \mathscr{L}$, there are more $\mathscr{L}'$-formal proofs then $\mathscr{L}$-formal proofs. But later, Gödel completeness will tell us $T \models \varphi \iff T \vdash_{\mathscr{L}} \varphi$, and we just said that $\models$ does not depend on language $\mathscr{L}$, and so a postieri formal provability of formula does not depend on language (again as long as φ can be written in $\mathscr{L}$). For now though, we must stick with the subscript.

10/7/22 Class 6

Example 5.1: A formal proof

If φ, ψ are $\mathscr{L}$-sentences, then $T := \{\varphi, \psi\}$ is $\mathscr{L}$-theory, and $T \vdash_{\mathscr{L}} (\varphi \wedge \psi)$. Formal proof:

- $\varphi_0 := \varphi$ ($\in T$)
- $\varphi_1 := \varphi \to (\psi \to (\varphi \wedge \psi))$ (1st order tautology)
- $\varphi_2 := \psi \to (\varphi \wedge \psi)$ (apply (MP) to φ_0, φ_1)
- $\varphi_3 := \psi$ ($\in T$)
- $\varphi_4 := (\varphi \wedge \psi)$ (apply (MP) to φ_3, φ_4)

Example 5.2: Some more formal proofs

If φ is $\mathscr{L}$-formula, x, y variables, and t $\mathscr{L}$-term,

(a) $\varnothing \vdash_{\mathscr{L}} \forall \mathsf{x} \varphi \to \varphi_{t/\mathsf{x}}$. Formal proof:
  - $\varphi_0 := \neg \varphi_{t/\mathsf{x}} \to \exists \mathsf{x} \neg \varphi$ (which is (Q2) on $\neg \varphi$)
  - $\varphi_1 := (\varphi_{t/\mathsf{x}} \to \exists \mathsf{x} \neg \varphi) \to (\neg \exists \mathsf{x} \neg \varphi \to \varphi_{t/\mathsf{x}}$ (tautology, coming from "contrapositive" proposi-

tional tautology: $(\neg P \to Q) \to (\neg Q \to P))$

- $\varphi_2 := \forall \mathsf{x}\varphi \to \varphi_{t/\mathsf{x}}$ (apply (MP) to φ_0, φ_1).

(b) if y does not occur in φ, then $\varnothing \vdash_{\mathscr{L}} (\forall \mathsf{y}\varphi_{\mathsf{y}/\mathsf{x}} \to \forall \mathsf{x}\varphi)$. Formal proof:

- $\varphi_0 := (\forall \mathsf{y}\varphi_{\mathsf{y}/\mathsf{x}} \to \varphi)$ (which is (a) applied to $\varphi_{\mathsf{y}/\mathsf{x}}$, where $[\varphi_{\mathsf{y}/\mathsf{x}}]_{\mathsf{x}/\mathsf{y}}$ is literally equal to φ since y does not occur in φ — more rigorously, when we say "use (a) applied to χ", we mean to repeat the formal proof of (a) again based around χ at the beginning of the current formal proof.)
- $\varphi_1 := \forall \mathsf{x}(\forall \mathsf{y}\varphi_{\mathsf{y}/\mathsf{x}} \to \varphi)$ (apply (Gen) to φ_0)
- $\varphi_2 := \forall \mathsf{x}(\forall \mathsf{y}\varphi_{\mathsf{y}/\mathsf{x}} \to \varphi) \to (\forall \mathsf{y}\varphi_{\mathsf{y}/\mathsf{x}} \to \forall \mathsf{x}\varphi)$ (by (Q1), since indeed x does not occur freely in $\forall \mathsf{y}\varphi_{\mathsf{y}/\mathsf{x}}$).
- $\varphi_3 := (\forall \mathsf{y}\varphi_{\mathsf{y}/\mathsf{x}} \to \forall \mathsf{x}\varphi)$ (apply (MP) to φ_1, φ_2).

(c) $\varnothing \vdash_{\mathscr{L}} \forall \mathsf{x}\varphi \to \varphi$. This is a special case of (a); take $t = \mathsf{x}$.

Important part is logical axioms and deduction rules computable/recursive, so computer can check if a given string is a formal proof. Universal validity might be easier for us check, for e.g. example $\forall \mathsf{x}\varphi \to \varphi_{t/\mathsf{x}}$ above, but formal proof easier for computer to do!

Lemma 5.1: Soundness lemma

Let T be $\mathscr{L}$-theory, φ an $\mathscr{L}$-formula; then $T \vdash_{\mathscr{L}} \varphi \implies T \models \varphi$ (i.e. formal syntacti provability $\implies$ logical/semantic consequence!)

Proof: we know logical axioms are all univ. valid (last time, showed tautologies univ. valid, (E1)-(E5) univ. valid, and (Q1)-(Q3) univ valid). We also know $[T \models \varphi$ and $T \models \varphi \to \psi] \implies [T \models \varphi]$ (for all $\mathscr{L}$-structures $\mathcal{A}$ s.t. $\mathcal{A} \models T$, then $\mathcal{A} \models \varphi$ and $\mathcal{A} \models \varphi \to \psi$ i.e. $\mathcal{A} \models \neg(\varphi \wedge \neg\psi)$, so must have $\mathcal{A} \models \psi$). Hence both deduction rules (MP), (Gen) preserve property of being logical consequence of a theory, so lemma is finished by induction on lenght of formal proof of φ from T. ■

Definition 5.6: Properties of $\mathscr{L}$-theories

- An $\mathscr{L}$-theory T is inconsistent if there is $\mathscr{L}$-sentence φ s.t. $T \vdash_{\mathscr{L}} \varphi$ and $T \vdash_{\mathscr{L}} \neg\varphi$.
- An $\mathscr{L}$-theory T is consistent if it is not inconsistent.
- An $\mathscr{L}$-theory T is complete if it is consistent, and for any $\mathscr{L}$-sentence φ, either $T \vdash_{\mathscr{L}} \varphi$ or $T \vdash_{\mathscr{L}} \neg\varphi$.

Remark: an $\mathscr{L}$-theory T is inconsistent $\iff T \vdash_{\mathscr{L}} \varphi$ for any $\mathscr{L}$-formula φ. *Pf:* let φ be arbitrary $\mathscr{L}$-sentence and ψ be $\mathscr{L}$-sentence s.t. $T \vdash_{\mathscr{L}}$ both $\varphi, \neg\phi$. Then (first copying the formal proofs of $\psi, \neg\psi$ from the logical axioms and axioms of T) we have the formal proof

- $\varphi_0 := (\neg\psi \to (\psi \to \varphi))$ (1st order tautology $\neg P \to (P \to Q)$)
- $\varphi_1 := \psi \to \varphi$ (apply (MP) to φ_0 and $\neg\psi$)
- $\varphi_2 := \varphi$ (apply (MP) to φ_1 and ψ)

so indeed $T \vdash_{\mathscr{L}} \varphi$.

Definition 5.7: Complete theory of structure

Let $\mathcal{A}$ be $\mathscr{L}$-structure. Then the set $\mathrm{Th}(\mathcal{A}) \coloneqq \{\varphi \text{ is an } \mathscr{L}\text{-sentence} : \mathcal{A} \models \varphi\}$ is a theory, called the <u>(complete) theory of $\mathcal{A}$</u>, which is moreover complete (consistent by soundness lemma, and complete because in a model every sentence is either true or false, by the equality/membership/existence oracles of a given set — cf. discussion in Lec. 2).

Remark: so every $\mathscr{L}$-structure produces complete theory. Can we go backwards, i.e. from complete (consistent) theory T get an $\mathscr{L}$-structure $\mathcal{A}$ s.t. $T = \mathrm{Th}(\mathcal{A})$? This is Gödel's completeness theorem.

Another remark: fundamental nature of proof is its finitary nature, i.e. if $T \vdash_{\mathscr{L}} \varphi$, then there exists a finite subset $T_0 \subseteq T$ s.t. $T_0 \vdash_{\mathscr{L}} \varphi$ (obvious for fixed φ, since formal proof is finite, so just take T_0 to be all (finitely many) sentences of T used in the finite proof). This is the heart of the compactness theorem, which we will see later. This fundamental remark has the following corollaries:

Corollary 5.1: Some faces of syntactic compactness

- If all finite subsets of $\mathscr{L}$-theory T are consistent $\mathscr{L}$-theories, then T is also consistent.
- If $(T_i)_{i \in I}$ is a nested family of consistent $\mathscr{L}$-structures (i.e. $T_i \subseteq T_j$ or $T_j \subseteq T_i$ for all $i, j \in I$), then $T \coloneqq \bigcup_{i \in I} T_i$ is also consistent.

Proof: because finite proof $\{\varphi_n\}_{n=1}^N$ of inconsistency ϕ and $\neg\phi$ must already be contained in one of the T_i (indeed only finitely many axioms of T occuring in the proofs of $\phi, \neg\phi$, which by definition of union means axioms lie in some $T_{i_1}, \ldots, T_{i_l}$, and nested condition means some $T_i \coloneqq T_{i_j}$ contains all other $T_{i_1}, \ldots, T_{i_l}$), $\{\varphi_n\}_{n=1}^N$ is a proof that $T_i \vdash_{\mathscr{L}}$ both $\phi, \neg\phi$; contradiction. ■

10/10/22 Class 7

"Very useful in arguing about formal proofs":

Lemma 5.2: Deduction lemma

Let T be $\mathscr{L}$-theory, χ an $\mathscr{L}$-sentence, φ an $\mathscr{L}$-formula; then $T \cup \{\chi\} \vdash_{\mathscr{L}} \varphi \iff T \vdash_{\mathscr{L}} (\chi \to \varphi)$.

Proof: note $T \vdash_{\mathscr{L}} (\chi \to \varphi)$ implies that $T \cup \{\chi\} \vdash_{\mathscr{L}} \varphi$ immediately by (MP) ($\varphi_0, \ldots, \varphi_n$ is formal proof of $\varphi_n \coloneqq (\chi \to \varphi)$, then take $\varphi_{n+1} \coloneqq \chi$ in $T \cup \{\chi\}$, and then $\varphi_{n+2} = \varphi$ by (MP)).

Now ($\Longleftarrow$): let $\varphi_0, \ldots, \varphi_n$ be formal proof of φ in $T \cup \{\chi\}$. By induction on $i \leqslant n$, we prove $T \vdash_{\mathscr{L}} (\chi \to \varphi_i)$:

- if $\varphi_i = \chi$, clear (tautology).
- if $T \vdash_{\mathscr{L}} \varphi_i$ (in particular when φ_i is logical axiom or element of T), clear (by (MP) and tautology $(\varphi_i \to (\chi \to \varphi_i))$).
- if φ_i is deduced by (MP) applied to φ_j, φ_k for $j, k < i$, i.e. $\varphi_k = (\varphi_j \to \varphi_i)$, it is enough to use (MP) plus tautology $((\chi \to \varphi_j) \wedge (\chi \to (\varphi_j \to \varphi_i))) \to (\chi \to \varphi_i)$, since T already proves $(\chi \to \varphi_i)$ and $(\chi \to \varphi_k)$ by induction.
- if φ_i is deduced by (Gen) from φ_j for $j < i$, i.e. $\varphi_i = \forall \mathsf{x} \varphi_j$, then we have by induction hypothesis $T \vdash_{\mathscr{L}} (\chi \to \varphi_j)$, and applying (Gen) to that yields $T \vdash_{\mathscr{L}} \forall \mathsf{x}(\chi \to \varphi_j)$. As χ is sentence, formula $\forall \mathsf{x}(\chi \to \varphi) \to (\chi \to \forall \mathsf{x} \varphi)$ is (Q1) (x does not occur freely in χ because χ sentence $\rightsquigarrow$ no free variables).

So by (MP), we get $T \vdash_{\mathscr{L}} (\chi \to \varphi_i)$.

Thus we have "lifted up" our proof $\varphi_0, \ldots, \varphi_n$ of φ in $T \cup \{\chi\}$ to a proof $(\chi \to \varphi_0), \ldots, (\chi \to \varphi_n)$ of $(\chi \to \varphi)$ in T, and we are done. ■

Corollary 5.2: Sentence provable iff negation inconsistent

Let T be $\mathscr{L}$-theory and φ $\mathscr{L}$-sentence; then $T \vdash_{\mathscr{L}} \varphi \iff T \cup \{\neg\varphi\}$ is inconsistent.

Proof: ($\Longleftarrow$): suppose $T \cup \{\neg\varphi\}$ is inconsistent. Because inconsistent theory proves everything, $T \cup \{\neg\varphi\} \vdash_{\mathscr{L}} \varphi$, so by deduction lemma, $T \vdash_{\mathscr{L}} (\neg\varphi \to \varphi)$. Well $(\neg\varphi \to \varphi) \to \varphi$ is tautology, so by (MP), $T \vdash_{\mathscr{L}} \varphi$.

($\Longrightarrow$): if $T \vdash_{\mathscr{L}} \varphi$, same proof holds in $T \cup \{\neg\varphi\}$, so $T \cup \{\chi\} \vdash_{\mathscr{L}} \varphi$. But also obviously $T \cup \{\chi\} \vdash_{\mathscr{L}} \neg\varphi$, so inconsistent. ■

6 Lecture 6

Before we embark on Gödel, we need the following technical lemma. It tells us that adding new constants in the language does not increase the provability power of the language. We do this by showing that new constants can be "emulated" by variables (which are available in the original language).

My remark: recall mentioned several lectures ago that $T \vdash_{\mathscr{L}}$ independent of $\mathscr{L}$, which is "trivial" if one knows Gödel completeness theorem $T \vdash_{\mathscr{L}} \varphi \iff T \models \varphi$. Here, in our journey to prove Gödel, must prove baby version of "indep. of $\mathscr{L}$", namely indep. of $\mathscr{L}$ plus added constants.

Lemma 6.1: Simulation of constants by variables

Let $\mathscr{L}$ be a language, ψ be an $\mathscr{L}$-formula, T an $\mathscr{L}$-theory, and c be an arbitrary (constant) symbol not already in $\mathscr{L}$. Then for a variable $\mathsf{x} \in \mathscr{V}$ and the new constant c, the following are equivalent (t.f.a.e.): ① $T \vdash_{\mathscr{L}} \psi$; ② $T \vdash_{\mathscr{L}\cup\{\mathsf{c}\}} \psi_{\mathsf{c}/\mathsf{x}}$; ③ $T \vdash_{\mathscr{L}\cup\{\mathsf{c}\}} \psi$.

Proof: ① $\implies$ ③: clearly, any $\mathscr{L}$-proof is also a $\mathscr{L} \cup \{\mathsf{c}\}$-proof.

③ $\implies$ ②: if $T \vdash_{\mathscr{L}\cup\{\mathsf{c}\}} \psi$, then by (Gen) $T \vdash_{\mathscr{L}\cup\{\mathsf{c}\}} \forall\mathsf{x}\psi$. Recall from last lecture that $\varnothing \vdash_{\mathscr{L}\cup\{\mathsf{c}\}} (\forall\mathsf{x}\varphi \to \varphi_{t/\mathsf{x}})$ for all formulas φ and terms t; in particular take $t = \mathsf{c}$. So by (MP) we get $T \vdash_{\mathscr{L}\cup\{\mathsf{c}\}} \psi_{\mathsf{c}/\mathsf{x}}$.

The hardest implication is ② $\implies$ ①. Somehow we need to convert a proof of $\psi_{\mathsf{c}/\mathsf{x}}$ in $\mathscr{L} \cup \{\mathsf{c}\}$ to a proof of ψ in $\mathscr{L}$. *Notation:* we use tildes when referring to formulas in $\mathscr{L} \cup \{\mathsf{c}\}$. The idea is to show that replacing the constant c in an $\mathscr{L} \cup \{\mathsf{c}\}$-formula by a variable y (not occurring in the formula) preserves many logical properties of the formula (preserving being logical axiom; or being obtained from (MP) or (Gen) applied to similarly emulated formulas): more precisely/neatly, if $\tilde{\varphi}, \tilde{\psi}, \tilde{\chi}$ are $\mathscr{L} \cup \{\mathsf{c}\}$-formulas with no occurrence of some variable y, and $\varphi := \tilde{\varphi}_{\mathsf{y}/\mathsf{c}}$:= the word obtained by substituting every occurrence of c in $\tilde{\varphi}$ by y (similarly/analogously defining ψ, χ), we **Claim:**

- φ is $\mathscr{L}$-formula — follows by straightforward induction on $\mathrm{ht}(\varphi)$.
- $\varphi_{\mathsf{c}/\mathsf{y}}$ is equal (as strings) to $\tilde{\varphi}$ — cf. Example 3.2, when we said that if θ is formula, variable x, term s not containing any variables occuring in θ, then $\theta_{s/\mathsf{x}}$ is naive simultaneous "find and replace"; indeed c term contains no variables so we can apply this example.
- if $\tilde{\varphi}$ is equality axiom/tautology/quantifier axiom, then so is φ — proof by casework. "Only non-trivial case is (Q2)": in which case $\tilde{\varphi}$ is of the form $(\tilde{\delta}_{t/\mathsf{x}} \to \exists\mathsf{x}\tilde{\delta})$, where $\tilde{\delta}$ is $\mathscr{L} \cup \{\mathsf{c}\}$-formula. We want to show that $\varphi := (\tilde{\delta}_{t/\mathsf{x}} \to \exists\mathsf{x}\tilde{\delta})_{\mathsf{y}/\mathsf{c}} = ([\tilde{\delta}_{t/\mathsf{x}}]_{\mathsf{y}/\mathsf{c}} \to \exists\mathsf{x}[\tilde{\delta}]_{\mathsf{y}/\mathsf{c}})$. Indeed $[\tilde{\delta}_{t/\mathsf{x}}]_{\mathsf{y}/\mathsf{c}}$ equals $[\tilde{\delta}_{\mathsf{y}/\mathsf{c}}]_{t/\mathsf{x}}$, proof by induction on $\mathrm{ht}(\tilde{\delta})$.
- if $\tilde{\varphi}$ obtained using (MP) from $\tilde{\psi}, \tilde{\chi}$, then φ obtained using (MP) from ψ, χ — casework, like in previous bullet.
- if $\tilde{\varphi}$ obtained using (Gen) from $\tilde{\psi}$, then φ obtained using (MP) from ψ — casework, like in previous bullet.

We are finally ready for ② $\implies$ ①: let $\tilde{\varphi}_0, \ldots, \tilde{\varphi}_m := \psi_{\mathsf{c}/\mathsf{x}}$ be a formal proof of $\psi_{\mathsf{c}/\mathsf{x}}$ in language $\mathscr{L} \cup \{\mathsf{c}\}$. Choose variable y not occuring in any sentence of the formal proof, and $\mathsf{y} \neq \mathsf{x}$ (we can do this because infinitely many variables, and only finitely many in $\tilde{\varphi}_0, \ldots, \tilde{\varphi}_m$). By **Claim** above, $\varphi_0, \ldots, \varphi_m = [\psi_{\mathsf{c}/\mathsf{x}}]_{\mathsf{y}/\mathsf{c}}$ is formal proof of $[\psi_{\mathsf{c}/\mathsf{x}}]_{\mathsf{y}/\mathsf{c}}$, which is equal to $\psi_{\mathsf{y}/\mathsf{x}}$ (by again Example 3.2). By (Gen), $T \vdash_{\mathscr{L}} \forall\mathsf{y}\psi_{\mathsf{y}/\mathsf{x}}$. Indeed y not in ψ, so Ex.5.2(b),(c) from Lec. 5 (plus (MP) twice) produces

$T \vdash_{\mathscr{L}} \psi$. ■

Reflections on the proof: it is long tedious analysis to check everything carefully, but intuitively c and y are not related to anything else in the formula (and don't have any syntactic baggage, as opposed to function/relation symbols or equality/quantifier symbols), and so replacing one by the other should have no impact on syntax (and hence formal proofs) besides superficially making one of the symbols look a little different. Semantically speaking, once a model and assignment are fixed, both constants/variables both act like placeholders for elements of the base set, so from a semantic perspective it also makes sense that $\left[\mathcal{A} \models \psi \text{ for any } \mathscr{L} \cup \{\mathsf{c}\}\text{-model } \mathcal{A} \text{ and any } \mathcal{A}\text{-assignment } \alpha\right]$ happens iff $\left[\mathcal{A} \models \psi_{\mathsf{c}/\mathsf{x}} \text{ for any for any } \mathscr{L} \cup \{\mathsf{c}\}\text{-model } \mathcal{A} \text{ and any } \mathcal{A}\text{-assignment } \alpha\right]$ happens — and this is by definition the statement $T \models \psi \iff T \models \psi_{\mathsf{c}/\mathsf{x}}$.

Corollary 6.1: Adding arbitrary constants doesn't impact provability

Let $\mathscr{L}$ be a language, ψ be an $\mathscr{L}$-formula, T an $\mathscr{L}$-theory, and $\mathscr{C}$ be an arbitrary set of (constant) symbols s.t. $\mathscr{C} \cap \mathscr{L} = \varnothing$. Then $T \vdash_{\mathscr{L}} \psi \iff T \vdash_{\mathscr{L} \cup \mathscr{C}} \psi$.

Proof: any fixed $\mathscr{L}$-formula ψ and $\mathscr{L} \cup \{\mathsf{c}\}$-proof of ψ involves only finite number of elements of $\mathscr{C}$, say $\mathsf{c}_1, \ldots, \mathsf{c}_n$, i.e. we know $T \vdash_{\mathscr{L} \cup \{\mathsf{c}_1,\ldots,\mathsf{c}_n\}} \psi$, so applying Lemma 6.1 ③ $\implies$ ① n times (more rigorously use induction on n) we have $T \vdash_{\mathscr{L} \cup \{\mathsf{c}_1,\ldots,\mathsf{c}_n\}} \psi \implies T \vdash_{\mathscr{L} \cup \{\mathsf{c}_1,\ldots,\mathsf{c}_{n-1}\}} \psi \implies \ldots \implies T \vdash_{\mathscr{L} \cup \{\mathsf{c}_1\}} \psi \implies T \vdash_{\mathscr{L}} \psi$. And of course $T \vdash_{\mathscr{L}} \psi \implies T \vdash_{\mathscr{L} \cup \{\mathsf{c}\}} \psi$ by same trivial 5-word argument from Lemma 6.1 ① $\implies$ ③. ■

10/12/22 Class 8

Lemma 6.2: Existential witnesses are consistent

Let T be $\mathscr{L}$-theory, $\varphi(\mathsf{x})$ be $\mathscr{L}$-formula, $\mathsf{c} \in \mathscr{L}$ a constant symbol NOT occuring in $T \cup \{\varphi(\mathsf{x})\}$. Assume T is consistent. Then, $T \cup \{\exists \mathsf{x}\varphi \to \varphi_{\mathsf{c}/\mathsf{x}}\}$ is consistent $\mathscr{L}$-theory.

Proof: recall end of last lecture, corollary of deduction lemma (Cor. 5.2): T $\mathscr{L}$-theory, φ $\mathscr{L}$-sentence, then $T \vdash_{\mathscr{L}} \varphi \iff T \cup \{\neg\varphi\}$ is inconsistent. So suppose f.s.o.c. $T \cup \{\exists \mathsf{x}\varphi \to \varphi_{\mathsf{c}/\mathsf{x}}\}$ inconsistent. Then aforementioned corollary tells us $T \vdash_{\mathscr{L}} \neg(\exists \mathsf{x}\varphi \to \varphi_{\mathsf{c}/\mathsf{x}})$, which expanding out abbreviations is exactly $(\exists \mathsf{x}\varphi \land \neg\varphi_{\mathsf{c}/\mathsf{x}})$. In particular, we have ⓘ $T \vdash_{\mathscr{L}} \exists \mathsf{x}\varphi$, and ⓘⓘ $T \vdash_{\mathscr{L}} \neg\varphi_{\mathsf{c}/\mathsf{x}}$. By simulation of constants lemma above (Lem. 6.1) ② $\iff$ ③, $T \vdash_{\mathscr{L}} \neg\varphi_{\mathsf{c}/\mathsf{x}} \iff T \vdash_{\mathscr{L}} \neg\varphi$, so by (Gen), $T \vdash_{\mathscr{L}} \forall \mathsf{x}\neg\varphi$. But recall (Q3): $\varnothing \vdash_{\mathscr{L}} (\exists \mathsf{x}\varphi \leftrightarrow \neg\forall \mathsf{x}\neg\varphi)$, so by ⓘ and (MP) we get $T \vdash_{\mathscr{L}} \neg\forall \mathsf{x}\neg\varphi$, and thus T is inconsistent; contradiction. ■

My remark 11/8/22: once we motivate Gödel completeness theorem — from more basic versions, e.g. propositional completeness theorem or 0th-order logic completeness theorem which are easier to prove), where in the course of proving these baby completeness theorems (at least in Terry's sketches) we already use baby versions of deduction lemma to phrase completeness theorem in model-existence form — and try proving (the special "admit Henkin witnesses, complete, deductively closed" case of) Gödel using Henkin witnesses (we use (E1)-(E5) in well-definedness of our model constructed "using syntax", and (Q2) in final "existential statement in $T^+ \iff$ exists constant witness", all very natural uses), we see that we will need to prove Lemma 6.2.

We already have experience with "lifting formal proofs" from (baby versions + full version of) deduction lemma (full version in particular requires us to use (Q1)), so defining logical axioms to be all the axioms we've needed so far (tautologies, (E1)-(E5), (Q1)-(Q2)), *amazingly* we can indeed prove

Lemma 6.2 by more complicated "proof lifting argument" (more complicated because have to "lift" logical axioms too, unlike in deduction lemma proof) — there is a bit of chicken-egg problem, because to prove that lemma need to already have defined formal proof system since we look at all those cases in the bullet points, but I sort-of want to motivate definition of formal proof system using this lemma; the issue is sidestepped because I already motivated the logical axioms in my formal proof system with previously mentioned major theorems.

We are now ready to state the Gödel completeness theorem:

> *Theorem 6.1: Gödel completeness theorem*
>
> Let T be $\mathscr{L}$-theory, φ be $\mathscr{L}$-sentence. Then, $T \models \varphi \iff T \vdash_{\mathscr{L}} \varphi$.

Remark: again, because $\models$ indep. of language $\mathscr{L}$, a postieri we see $\vdash_{\mathscr{L}}$ indep. of $\mathscr{L}$, so once we finish proof of Gödel, we may drop the subscript $\mathscr{L}$ and just write $T \vdash \varphi$.

Jas asked if there was purely syntactic proof that formal provability is independent of adding new symbols. Artem says not sure, but would be tricky: just adding constants alone took us a decent amount of work, and proof of Gödel is "highly nonconstructive" (uses weak form of axiom of choice).

Observe that we already know the ($\Longleftarrow$) direction (soundless lemma of last lecture); the difficult direction is ($\Longrightarrow$), where we need to exhibit/construct a formal proof. Recall by definition (Def. 5.5) $T \not\models \varphi \iff T \cup \{\neg\varphi\}$ hs a model. And by corollary of deduction lemma $T \nvdash_{\mathscr{L}} \varphi \iff T \cup \{\neg\varphi\}$ is consistent. So Gödel completeness is equivalent to the following

> *Theorem 6.2: Gödel completeness theorem, model existence formulation*
>
> An $\mathscr{L}$-theory T has a model $\iff$ it is consistent.

Remark: again as in Gödel original formulation, LHS is semantic notion, and RHS is syntactic/formal notion, and Gödel is the bridge between the realms.

7 Lecture 7

We begin the *Proof* of Thm. 6.2. Again, we have an easy direction accomplished by soundness: if T has a model, then it is consistent: if f.s.o.c. $T \vdash_{\mathscr{L}} \varphi, \neg\varphi$, then by soundness $T \vdash_{\mathscr{L}} \varphi \implies T \models \varphi$ and $T \vdash_{\mathscr{L}} \neg\varphi \implies T \models \neg\varphi$, so since there exists $\mathcal{A} \models T$, then $\mathcal{A} \models \varphi$ and $\mathcal{A} \models \neg\varphi$, which is impossible by def. of satisfaction of sentences — i.e. the oracles of set theory forbid this.

So the difficult task now is: given a consistent $\mathscr{L}$-theory T, build a model. "Ingenious trick (of Henkin) is to build a model out of syntax, i.e. add a lot of constant symbols, and somehow establish relations (formulas) on them so that they behave correctly". In slightly more detail, Henkin's method is to define an expansion T^+ of T in some language $\mathscr{L}^+ \supseteq \mathscr{L}$ with "better properties", namely the following property:

Definition 7.1: Henkin witnesses

Let $\mathscr{L}$ language, $\mathscr{C}$ a set of constant symbols s.t. $\mathscr{L} \cup \mathscr{C} = \varnothing$. We say a $\mathscr{L} \cup \mathscr{C}$-theory T^+ admits Henkin witnesses in $\mathscr{C}$ if for any $\mathscr{L} \cup \mathscr{C}$-formula $\varphi(\mathsf{x})$ there exists $\mathsf{c} \in \mathscr{C}$ s.t. $(\exists \mathsf{x} \varphi \rightarrow \varphi_{\mathsf{c}/\mathsf{x}}) \in T^+$.

So T^+ contains sentences that force $\mathscr{C}$ to witness all existential formulas/ensure all existential formulas have a witness, i.e. element (given by interpretation of constant symbol) satisfying φ.

Example 7.1: Model produces complete theory with Henkin witnesses

If $\mathcal{A}$ is $\mathscr{L}$-structure, and $\{a_{\mathsf{c}} : \mathsf{c} \in \mathscr{C}\}$ is an enumeration (possibly not injective, i.e. $\mathsf{c} \neq \mathsf{c}'$ but $a_{\mathsf{c}} = a_{\mathsf{c}'}$) of the base set A by set of constants $\mathscr{C}$. Then we denote by $\mathcal{A}^+$ the $\mathscr{L} \cup \mathscr{C}$-structure obtained from $\mathcal{A}$ by interpreting c by a_{c}.

Then $\mathrm{Th}(\mathcal{A}^+)$ (which by definition is complete and deductively closed) admits Henkin witnesses in $\mathscr{C}$: this is because $\mathcal{A} \models \exists \mathsf{x} \varphi$ means there is $a \in A$ s.t. $\mathcal{A} \models \varphi(a)$, where $a = a_{\mathsf{c}}$ for some (i.e. at least one) $\mathsf{c} \in \mathscr{C}$, so then $\mathcal{A}^+ \models \varphi(\mathsf{c}^{\mathcal{A}^+}) \iff \mathcal{A}^+ \models \varphi_{\mathsf{c}/\mathsf{x}}$ (by Def. 3.2), and hence $\mathrm{Th}(\mathcal{A}^+) \cup \{\exists \mathsf{x} \varphi\} \vdash \varphi_{\mathsf{c}/\mathsf{x}}$, which is equivalent to $\mathrm{Th}(\mathcal{A}^+) \vdash (\exists \mathsf{x} \varphi \rightarrow \varphi_{\mathsf{c}/\mathsf{x}})$ by the deduction lemma.

To summarize: given $\mathscr{L}$-structure and bunch of constants naming elements of this structure, adding these constants to language $\mathscr{L}$, the complete theory of expansion $\mathrm{Th}(\mathcal{A}^+)$ is automatically "nice". This exact idea of encoding a model purely syntactically using new constants and the $\mathscr{L}^+$-theory $\mathrm{Th}(\mathcal{A}^+)$ we will see again when we discuss the diagram method.

"This example is very straightforward way of attaining (complete, deductively closed) theory that admits Henkin witness. We will in fact show that every such theory can be attained in this manner, i.e. there is a 1-1 correspondence between complete (and deductively closed) theories with Henkin witnesses and $\mathscr{L}$-structures. In other words, this correspondence proves Gödel (Thm. 6.2) for the special case of complete (and deductively closed) theories with Henkin witnesses in some set of constants $\mathscr{C}$. The last step is to reduce any arbitrary theory to a complete theory admitting Henkin witnesses (this is the nonconstructive step)."

10/17/22 Class 9 (got through half of Prop. last time, but redid from beginning)

Proposition 7.1: Correspondence of structures and complete theories admitting Henkin witnesses

Any complete $\mathscr{L} \cup \mathscr{C}$-theory T^+ which admits Henkin witnesses in $\mathscr{C}$ has model $\mathcal{A}^+$ consisting of constants of $\mathscr{C}$, i.e. with base set $A^+ = \{\mathsf{c}^{\mathcal{A}^+} : \mathsf{c} \in \mathscr{C}\}$.

Proof: first **Note:** that replacing T^+ by the set $\hat{T}^+$ of $\mathscr{L} \cup \mathscr{C}$-sentences φ s.t. $T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \varphi$ doesn't change assumptions, because all sentences required by "admitting Henkin witnesses in $\mathscr{C}$" are already in $T^+ \subseteq \hat{T}^+$; and T^+ is already complete, i.e. already proves either φ or $\neg\varphi$, so the same holds for any larger set than T^+ (that T^+ still proves, so as to avoid inconsistency), e.g. $\hat{T}^+$. By this **Note**, we may assume T^+ is deductively closed, i.e. $T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \varphi \iff \varphi \in T^+$.

Given $\mathsf{c}, \mathsf{d} \in \mathscr{C}$, we define the notation $\mathsf{c} \sim \mathsf{d} \;:\Longleftrightarrow\; \mathsf{c{=}d} \in T^+ \iff T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{c{=}d}$. It follows from the equality axioms (E1)-(E3) that $\sim$ is an equivalence relation (i.e. satisfies (E1)-(E3)). We notate equivalences classes by $a_{\mathsf{c}} := \mathsf{c}/\sim$, and $A^+ := \{a_{\mathsf{c}} : \mathsf{c} \in \mathscr{C}\}$. Then $\mathcal{A}^+ \neq \varnothing$ (because $\mathscr{C} \neq \varnothing$, implicitly true in assumption that T^+ "admits Henkin witnesses in $\mathscr{C}$"; more explicitly take sentence $\exists \mathsf{v}_1(\mathsf{v}_1{=}\mathsf{v}_1)$ which is witnessed by some $\mathsf{c} \in \mathscr{C}$).

So we have defined what we think the base set should be; we now have to interpret constant/function/relation symbols in $\mathcal{A}$ consistent with how T^+ thinks of those symbols. We can now define $\mathscr{L} \cup \mathscr{C}$-structure $\mathcal{A}^+$ as follows:

- for a constant symbol d (i.e. $\mathsf{d} \in \mathscr{C}^{\mathscr{L}} \cup \mathscr{C}$), define $\mathsf{d}^{\mathcal{A}^+} := a_{\mathsf{c}}$ for $\mathsf{c} \in \mathscr{C}$ if $\mathsf{c{=}d} \in T^+$. (Such c always exists because $\mathsf{d{=}d} \rightarrow \exists \mathsf{x}\, \mathsf{x{=}d} \in T^+$ by (Q2), and by (E1) and (MP) we get $\exists \mathsf{x}\, \mathsf{x{=}d} \in T^+$, and by defiinition of "admit Henking witnesses in $\mathscr{C}$", must exist $\mathsf{c} \in \mathscr{C}$ s.t. $\exists \mathsf{x}\, \mathsf{x{=}d} \rightarrow \mathsf{c{=}d} \in T^+$, and by (MP) again $\mathsf{c{=}d} \in T^+$ as wanted). Moreover by equality axioms (E1)-(E3)/equivalence class shenanigans, $\mathsf{d}^{\mathcal{A}^+}$ doesn't depend on the choice of c.
- for $\mathsf{R} \in \mathscr{R}_n^{\mathscr{L}}$, set $(a_{\mathsf{c}_1}, \ldots, a_{\mathsf{c}_n}) \in \mathsf{R}^{\mathcal{A}^+} \;:\Longleftrightarrow\; \mathsf{Rc}_1 \ldots \mathsf{c}_n \in T^+$. Well-defined by (E5), and perhaps also using (E3).
- for $\mathsf{f} \in \mathscr{F}_n^{\mathscr{L}}$, set $\mathsf{f}^{\mathcal{A}^+}(a_{\mathsf{c}_1}, \ldots, a_{\mathsf{c}_n}) = a_{\mathsf{c}_0} \;:\Longleftrightarrow\; \mathsf{fc}_1 \ldots \mathsf{c}_n{=}\mathsf{c}_0 \in T^+$. Well-defined by (E4),(E3), and also T^+ admits Henkin witnessses in $\mathscr{C}$ and is deductively closed $\rightsquigarrow \mathsf{f}^{\mathcal{A}^+}$ is defined on all of A^n and is a function. More explicitly, for all $\mathsf{c}_1, \ldots, \mathsf{c}_n \in \mathscr{C}$, defining the formula $\varphi(\mathsf{y}) := \mathsf{y{=}fc}_1 \ldots \mathsf{c}_n$ and term $t := \mathsf{fc}_1 \ldots \mathsf{c}_n$, we have $\varnothing \vdash_{\mathscr{L}\cup\mathscr{C}} \varphi_{t/\mathsf{y}} \implies \varnothing \vdash_{\mathscr{L}\cup\mathscr{C}} \exists \mathsf{y}\varphi)$ by (Q2), so by admitting Henkin witnesses there is $\mathsf{c}_0 \in \mathscr{C}$ s.t. $T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{c}_0{=}\mathsf{fc}_1 \ldots \mathsf{c}_n$, hence $\mathsf{f}^{\mathcal{A}^+}$ is defined on all of A^n.
  ▶ Well-defined because if $\mathsf{c}'_1, \ldots, \mathsf{c}'_n \in \mathscr{C}$ are s.t. $a_{\mathsf{c}_i} = a_{\mathsf{c}'_i} \implies T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{c}_i{=}\mathsf{c}'_i$ for $i \in [n]^+$, and $T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{c}_0{=}\mathsf{fc}_1 \ldots \mathsf{c}_n$ and also $T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{c}'_0{=}\mathsf{fc}'_1 \ldots \mathsf{c}'_n$, then by (E4) $T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{c}'_0{=}\mathsf{fc}_1 \ldots \mathsf{c}_n$, and by (E3) $T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{c}_0{=}\mathsf{c}'_0 \implies a_{\mathsf{c}_0} = a_{\mathsf{c}'_0}$).

(Artem's response to a question of Mark's from last time): ***Remark:*** so far we have not used completeness of T^+ (deductive closedness just so we can interchange "$\ldots \in T^+$" and "$T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \ldots$", mainly for ease of notation because the former is easier to write than the latter — i.e. replace all instances of "$\ldots \in T^+$" by "$T^+ \vdash_{\mathscr{L}\cup\mathscr{C}} \ldots$" and then we will also not have used deductive closure either), only used existence of Henkin witnesses. Only use completeness in one important place, will point out when we get there.

We have finished defining the $\mathscr{L}\cup\mathscr{C}$-structure $\mathcal{A}^+$, so all we have left is to show $\mathcal{A}^+$ actually models T^+. **Observe** that:

(I) if t is $\mathscr{L}\cup\mathscr{C}$-term without variables (i.e. compositions of functions evaluated at constants) and $\mathsf{c}\in\mathscr{C}$, then $t^{\mathcal{A}^+} = a_{\mathsf{c}} \iff t{=}\mathsf{c}\in T^+$;

(II) and if ψ is $\mathscr{L}\cup\mathscr{C}$-sentence, then $\mathcal{A}^+\models\psi \iff \psi\in T^+$.

In other words, satisfaction/truth in structure $\mathcal{A}^+$ is completely controlled by which sentences are $\in T^+$; in particular (II) implies $\mathcal{A}^+$ is desired model in Prop., because (II) means that $\mathcal{A}^+\models T^+$. Now, proof of **Observations**:

(I) induction on term-height $\mathrm{ht}(t)$ (the base case $t \coloneqq \mathsf{f}\mathsf{c}_1\ldots\mathsf{c}_n$ is exactly by definition of $\mathsf{f}^{\mathcal{A}^+}$ above), in particular using (E4),(E3) during the induction step: (recall term t has no variables so no need for $\mathcal{A}$-assignment α) $a_{\mathsf{c}} = [\mathsf{f}t_1\ldots t_n]^{\mathcal{A}^+} \coloneqq \mathsf{f}^{\mathcal{A}^+}(t_1^{\mathcal{A}^+},\ldots,t_n^{\mathcal{A}^+}) \eqqcolon f^{\mathcal{A}^+}(a_{\mathsf{c}_1},\ldots,a_{\mathsf{c}_n})$, and induction hypothesis says $t_i^{\mathcal{A}^+} \eqqcolon a_{\mathsf{c}_i} \iff t_i{=}\mathsf{c}_i\in T^+$, so (E4) implies $T^+\vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{f}t_1\ldots t_n{=}\mathsf{f}\mathsf{c}_1\ldots\mathsf{c}_n$, and base case/def. of $f^{\mathcal{A}^+}$ gives $T^+\vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{f}\mathsf{c}_1\ldots\mathsf{c}_n{=}\mathsf{c}$, so (E3) puts these together to get the desired $T^+\vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{f}t_1\ldots t_n{=}\mathsf{c}$.

(II) induction on formula-height $\mathrm{ht}(\psi)$. Recall from previous lectures that $\mathrm{ht}(\psi)=\mathrm{ht}(\psi_{\overline{\mathsf{s}}/\overline{\mathsf{x}}})$.
- if $\psi = t_1{=}t_2$ for terms t_1,t_2, then (II) follows from (I).
- if $\psi=\mathsf{R}t_1\ldots t_n$ then $t_i^{\mathcal{A}^+}$ is some element of $\mathcal{A}^+$, say a_{c_i} for some $\mathsf{c}_i\in\mathscr{C}$. By (I), $t_i{=}\mathsf{c}_i\in T^+$, which allows us to use (E5):

$$\mathcal{A}^+\models\psi \underset{\text{def. of }\mathcal{A}^+}{:\iff} T^+\vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{R}\mathsf{c}_1\ldots\mathsf{c}_n \underset{\text{(E5)}}{\iff} T^+\vdash_{\mathscr{L}\cup\mathscr{C}} \mathsf{R}t_1\ldots t_n.$$

- if $\psi=(\varphi_1\wedge\varphi_2)$, follows by induction hypothesis on φ_1,φ_2.
- if $\psi=\neg\varphi$, then
$$\mathcal{A}^+\models\psi \iff \mathcal{A}^+\not\models\varphi \underset{\substack{\text{induction}\\ \text{hypo.}}}{\iff} \varphi\notin T^+ \underset{\substack{T^+\text{ complete}\\ \text{deduc. closed}}}{\iff} \psi\in T^+$$

  This is the only place where completeness is necessary. Without it, we can define the same structure, but no way to prove it models the theory.
- if $\psi=\exists\mathsf{x}\varphi$, then **Note:** $\psi\in T^+ \iff$ there exists $\mathsf{c}\in\mathscr{C}$ s.t. $\varphi_{\mathsf{c}/\mathsf{x}}\in T^+$ because we assume T^+ admits Henkin witness in $\mathscr{C}$ so ($\Longrightarrow$) follows by (MP), and for ($\Longleftarrow$), (Q2) says $(\varphi_{\mathsf{c}/\mathsf{x}}\to\exists\mathsf{x}\varphi)$ universally valid, so by (MP) and deductive closure can conclude $T^+\vdash_{\mathscr{L}\cup\mathscr{C}}\exists\mathsf{x}\varphi \iff \psi\coloneqq\exists\mathsf{x}\varphi\in T^+$, so then

$$\mathcal{A}^+\models\psi \underset{\text{def. of }\models}{\iff} \text{there is some element } a_{\mathsf{c}}\in A^+ \text{ (i.e. there is some } \mathsf{c}\in\mathscr{C}) \text{ s.t. } \mathcal{A}^+\models\varphi[a_{\mathsf{c}}]$$
$$\underset{\substack{\text{substitution}\\ \text{lemma, Lem. 3.1}}}{\iff} \mathcal{A}^+\models\varphi_{\mathsf{c}/\mathsf{x}} \text{ for some } \mathsf{c}\in\mathscr{C} \underset{\substack{\text{induction}\\ \text{hypo.}}}{\iff} \varphi_{\mathsf{c}/\mathsf{x}}\in T^+ \text{ for some } \mathsf{c}\in\mathscr{C} \underset{\textbf{Note}}{\iff} \psi\in T^+.$$

This concludes the proof of the proposition. ■

Thus, to finish the proof of the Gödel completeness theorem, we just have to reduce all theories to the special case addressed in Prop.; continued in Lec. 8 below.

EDIT 11/27/22: we defined admitting Henkin witnesses in $\mathscr{C}$ as "every $\mathscr{L}\cup\mathscr{C}$-formula $\varphi(\mathsf{x})$ with at most one free variable has an associated $\mathscr{L}\cup\mathscr{C}$-sentence $\exists\mathsf{x}\varphi\to\varphi_{\mathsf{c}/\mathsf{x}}$ in the theory, for some $\mathsf{c}\in\mathscr{C}$".

But looking through the proofs (only appears/plays a role in the very last bullet point of Prop. 7.1, and Lemma 6.2, which should be changed to "if T is consistent $\mathscr{L}$-theory and $\exists \mathsf{x} \varphi(\mathsf{x})$ is an existential sentence in T and $\mathsf{c} \in \mathscr{L}$ is a constant symbol that does not occur in T, then $T \cup \{\varphi(\mathsf{c})\}$ is a consistent $\mathscr{L}$-theory"), a perhaps more intuitive definition is "every existential $\mathscr{L} \cup \mathscr{C}$-sentence $\exists \mathsf{x} \varphi(\mathsf{x})$ in the theory has an associated $\mathscr{L} \cup \mathscr{C}$-sentence $\varphi_{\mathsf{c}/\mathsf{x}} =: \varphi(\mathsf{c})$ in the theory, for some $\mathsf{c} \in \mathscr{C}$" (although maybe a little harder to work with, since have to deal with the precondition of "if the formula is existential"). Obviously the book/class definition implies my definition, but I don't think the converse is true.

8 Lecture 8

Lemma 8.1: Adding Henkin witnesses

For any consistent $\mathscr{L}$-theory T, there exists a set of constant symbols $\mathscr{C}$ disjoint from $\mathscr{L}$, and for $\mathscr{L}^+ := \mathscr{L} \cup \mathscr{C}$ there is $\mathscr{L}^+$-theory T^+ containing T that moreover admits Henkin witnesses in $\mathscr{C}$.

Artem: "This is one of those proofs in mathematics that are called 'just do it' proofs. We don't have any algorithm, but we just add things and hope we don't get any contradictions."

Proof: for any $\mathscr{L}$-formula $\varphi(\mathsf{x})$ (by which recall means $\mathscr{L}$-formula with $\leqslant 1$ free variable, with that variable named x)

"So we 'just did it', added constants and said that they obey the Henkin witness sentence. Of course, we can't just say that, we must check that we didn't do anything stupid, wasn't just wishful thinking." So, we must prove the following claim.

10/19/22 Class 10

- **Claim:** this $\mathscr{L}_1$-theory T_1 is consistent.

This concludes the proof of Lemma 8.1. ■

Lemma 8.2: Completion

Any consistent $\mathscr{L}$-theory admits a completion, i.e. is contained in some complete $\mathscr{L}$-theory.

Proof: Zorn! ■

Interlude

Fleshed out proof(s) of Lemma

9 Lecture 9

Limbo

10/21/22 Class 11

"Theorem 0" of model theory; purely semantic, and its importance elevates it to "Fundamental Theorem of Model Theory".

> *Theorem 9.1: Compactness*
>
> For any language $\mathscr{L}$ and any $\mathscr{L}$-theory T, if every finite subset of T has a model, then T has a model.

Proof: recall (Corollary in Lec. 5) that T consistent $\iff$ any finite subset of T is consistent, and Gödel completeness theorem says that theory has model $\iff$ consistent. ■

Remark: originally due to Gödel for countable $\mathscr{L}$, Malcev for $\mathscr{L}$ of arbitrary cardinality. As for the name, we will see later that this corresponds to literal topological compactness of a topological space associated with the theory, namely the space of types. Also, we will later see purely semantic proof with ultrafilters, without the need for Gödel completeness.

> *Definition 9.1: Substructures*
>
> For $\mathscr{L}$-structures $\mathcal{A}, \mathcal{B}$, we say $\mathcal{A} \subseteq \mathcal{B}$, read "$\mathcal{A}$ is a substructure of $\mathcal{B}$" to mean that the base sets $A \subseteq B$, and that for every
>
> - $\mathsf{c} \in \mathscr{C}^{\mathscr{L}}$, $\mathsf{c}^{\mathcal{A}} = \mathsf{c}^{\mathcal{B}}$ (as elements of B, in fact of A since LHS $\mathsf{c}^{\mathcal{A}}$ must be in A);
> - $\mathsf{f} \in \mathscr{F}^{\mathscr{L}}_n$, $\mathsf{f}^{\mathcal{A}} \equiv \mathsf{f}^{\mathcal{B}}\restriction_{A^n}$ as functions $A^n \to A$ (in particular A must be closed under $\mathsf{f}^{\mathcal{B}}$);
> - $\mathsf{R} \in \mathscr{R}^{\mathscr{L}}_n$, $\mathsf{R}^{\mathcal{A}} = \mathsf{R}^{\mathcal{B}} \cap A^n$ as subsets of A^n.
>
> In other words, $\mathcal{A} \subseteq \mathcal{B}$ if and only if the identity/inclusion map $\iota : A \hookrightarrow B$ is a embedding/model homomorphism (as defined in Def. 1.4).

Remark: if $\mathcal{A} \subseteq \mathcal{B}$, then for any q.f. formula $\varphi(\mathsf{x}_1, \dots, \mathsf{x}_n)$ and parameters $a_1, \dots, a_n \in A$, we have $\mathcal{A} \models \varphi(a_1, \dots, a_n) \iff \mathcal{B} \models \varphi(a_1, \dots, a_n)$ (this is 220Ahw1p7(3); the proof is to observe for variables/constants/functions/relations $\mathsf{x}^{\mathcal{A}}(a) := a =: \mathsf{x}^{\mathcal{B}}(a)$, $\mathsf{c}^{\mathcal{A}} := \mathsf{c}^{\mathcal{B}}$, $\mathsf{f}^{\mathcal{A}} :\equiv \mathsf{f}^{\mathcal{B}}\restriction_{\mathcal{A}}$, and $\mathsf{R}^{\mathcal{A}} := \mathsf{R}^{\mathcal{B}} \cap A^{\mathsf{arity(R)}}$, where the first 3 imply for all terms $t(\overline{\mathsf{x}})$, $t^{\mathcal{A}} \equiv t^{\mathcal{B}}\restriction_{A}$ in the sense of functions $A^{|\overline{\mathsf{x}}|} \to \mathcal{A}$; from here, induction on formula-height with base case equality/relations, e.g. $\mathcal{A} \models [t_1{=}t_2](\overline{a}) \iff t_1^{\mathcal{A}}(\overline{a}) = t_2^{\mathcal{A}}(\overline{a}) \iff t_1^{\mathcal{B}}(\overline{a}) = t_2^{\mathcal{B}}(\overline{a}) \iff \mathcal{B} \models [t_1{=}t_2](\overline{a})$, and inductive step considering only negation/conjunction is trivial) — note that more generally for any embedding $h : A \hookrightarrow B$ we have $\mathcal{A} \models \varphi(\overline{a}) \iff \mathcal{B} \models \varphi(h(\overline{a}))$ for all q.f. formulas φ (injectivity of h important (in fact necessary) because above proof proves $h \circ t^{\mathcal{A}} \equiv t^{\mathcal{B}}\restriction_{h(A)} \circ h : A^{|\overline{\mathsf{x}}|} \to A$, and hence e.g. analogously to above we get $h(t_1^{\mathcal{A}}(\overline{a})) = h(t_2^{\mathcal{A}}(\overline{a})) \iff t_1^{\mathcal{B}}(h(\overline{a})) = t_2^{\mathcal{B}}(h(\overline{a}))$, but the equivalence $h(t_1^{\mathcal{A}}(\overline{a})) = h(t_2^{\mathcal{A}}(\overline{a})) \iff t_1^{\mathcal{A}}(\overline{a}) = t_2^{\mathcal{A}}(\overline{a})$ requires injectivity of h). However, not necessarily on all $\mathscr{L}$-formulas: consider $\mathcal{B} := (\mathbb{N}, <) \supseteq (\mathbb{N} \setminus \{1\}, <) =: \mathcal{A}$ and sentence $\varphi(\mathsf{v}_1, \mathsf{v}_2) := \exists \mathsf{v}_3\ \mathsf{v}_1{<}\mathsf{v}_3{<}\mathsf{v}_2$; clearly $\mathcal{A} \models \neg\varphi(0, 2)$ but $\mathcal{B} \models \varphi(0, 2)$. This motivates stronger relationships between structures:

Definition 9.2: Elementary equivalent

We say $\mathscr{L}$-structures $\mathcal{M}, \mathcal{N}$ are <u>elementary equivalent</u> to mean they satisfy the same $\mathscr{L}$-sentences, i.e. $\mathrm{Th}(\mathcal{M}) = \mathrm{Th}(\mathcal{N})$ (recall from Def. 5.7 $\mathrm{Th}(\mathcal{N}) := \{\varphi \text{ an } \mathscr{L}\text{-sentence} : \mathcal{N} \models \varphi\}$).

Definition 9.3: Elementary substructure

If $\mathcal{M} \subseteq \mathcal{N}$ are $\mathscr{L}$-structures, we say $\mathcal{M}$ is <u>elementary substructure</u> of $\mathcal{N}$ (and $\mathcal{N}$ is <u>elementary extension</u> of $\mathcal{M}$), denoted $\mathcal{M} \preceq \mathcal{N}$, to mean that for every $\mathscr{L}$-formula $\varphi(\mathsf{x}_1, \ldots, \mathsf{x}_n)$ and every tuple $(a_1, \ldots, a_n) \in M^n$, $\mathcal{M} \models \varphi(a_1, \ldots, a_n) \iff \mathcal{N} \models \varphi(a_1, \ldots, a_n)$. If $\mathcal{M} \preceq \mathcal{N}$ and $\mathcal{M} \subsetneq \mathcal{N}$, we write $\mathcal{M} \prec \mathcal{N}$.

Perhaps a more intuitive notion (one that is not concerned with exactly what elements are in the base sets) is that of

Definition 9.4: Elementary embedding

For $\mathscr{L}$-structures $\mathcal{M}, \mathcal{N}$, a map $h : M \hookrightarrow N$ s.t. for all $\mathscr{L}$-formulas $\varphi(\bar{\mathsf{x}})$ and parameters $\bar{a} \in M^n$, $\mathcal{M} \models \varphi(\bar{a}) \iff \mathcal{N} \models \varphi(h(\bar{a}))$ is called an elementary embedding. (Compare to above *Remark* about how embeddings satisfy this "$\iff$" only for q.f. formulas.)

Remark: note that the injectivity assumption on h is redundant, since taking $\varphi(\mathsf{x}, \mathsf{y}) := \mathsf{x}{=}\mathsf{y}$, we have that $a = b$ in $M \iff \mathcal{M} \models \varphi(a, b) \iff \mathcal{N} \models \varphi(h(a), h(b)) \iff h(a) = h(b)$ in N. Anyways, this embedding $h : \mathcal{M} \hookrightarrow \mathcal{N}$ induces an isomorphism $\mathcal{M} \xrightarrow{\cong} \mathcal{N}' := \mathrm{im}(h)$, where $\mathcal{N}' \preceq \mathcal{N}$ (since the embedding is elementary).

These definitions relate to each other in the following ways:

-
-
- $\mathcal{M} \subseteq \mathcal{N}$ and $\mathcal{M} \equiv \mathcal{N} \not\Longrightarrow \mathcal{M} \preceq \mathcal{N}$ (i.e. the converse to the first bullet point does not hold. Or rephrased: $\mathcal{M}, \mathcal{N}$ with $M \subseteq N$, who interpret constant/function/relation symbols exactly the same on all M-valued inputs/parameters/assignments, and satisfy the same $\mathscr{L}$-sentences, does NOT imply that they satisfy the same $\mathscr{L}$-formulas, even when restrictied to only M-valued inputs/parameters/assignments).

The next result gives sufficient (and necessary) condition for $\mathcal{M} \preceq \mathcal{N}$.

Theorem 9.2: Tarski-Vaught test

Let $\mathcal{M}, \mathcal{N}$ be $\mathscr{L}$-structures with $\mathcal{M} \subseteq \mathcal{N}$. Assume that for any $\mathscr{L}$-formula $\varphi(\mathsf{x}_0, \ldots, \mathsf{x}_n)$ and tuple $\bar{a} := (a_1, \ldots, a_n) \in M^n$, if there exists $b_0 \in N$ s.t. $\mathcal{N} \models \varphi(b_0, a_1, \ldots, a_n)$, then there exists $a_0 \in M$ s.t. $\mathcal{M} \models \varphi(a_0, \ldots, a_n)$. Then $\mathcal{M} \preceq \mathcal{N}$.

Remark: "dragging down witnesses in N to M one quantifier at a time suffices". Converse of above statement is true ... (see 2022 black hardcover notebook for overcomplicated proof). Tarski-Vaught test arises very naturally out of trying to prove $\mathcal{M} \models \varphi(\bar{a}) \iff \mathcal{N} \models \varphi(\bar{a})$ by induction on $\mathrm{ht}(\varphi)$:

Proof (of Thm. 9.2): ■

Theorem 9.3: Downwards Löwenheim-Skolem (↓-LS)

Let $\mathcal{M}$ be an $\mathscr{L}$-structure, and $A \subseteq M$ a subset of the base set M of $\mathcal{M}$. Assuming that $|M| \geqslant |\mathscr{L}|$, then there exists an elementary substructure $\mathcal{M}_0 \preceq \mathcal{M}$ s.t. $A \subseteq M_0$, and $|M_0| = \max\{|A|, |\mathscr{L}|\}$.

Idea: start with A, add existential witnesses as required by Tarski-Vaught, and keep growing substructures inductively (countably infinite many times). Then show result doesn't grow too big.

10/24/22 Class 12

Proof (of Thm. 9.3): first note: if $|A| < |\mathscr{L}|$, we may enlarge the subset A to $\hat{A}$ so that $|\hat{A}| = |\mathscr{L}|$, where indeed proving the theorem as stated for $\hat{A}$ proves the theorem as stated for A. Thus, in what follows, we may assume $|A| \geqslant |\mathscr{L}|$.

Notation: if we have a non-empty set $B \subseteq M$, recall from 220Ahw1p2 we denote by $\tilde{B} \subseteq M$ the base set of the substructure $\langle B\rangle_{\mathcal{M}}$ generated by B defined by $\langle B\rangle_{\mathcal{M}} := \bigcap_{\mathcal{M}' \subseteq \mathcal{M}, B \subseteq M'} \mathcal{M}'$ (using that intersection of substructures of $\mathcal{M}$ are also substructures of $\mathcal{M}$). We furthermore know an explicit description of $\tilde{B} = \{t^{\mathcal{M}}(b_1, \ldots, b_n) : t(\mathsf{x}_1, \ldots, \mathsf{x}_n) \in \mathcal{T}^{\mathscr{L}}, b_1, \ldots, b_n \in B\}$.

Observation: the explicit description of $\tilde{B}$ gives us a surjection $\mathcal{T}^{\mathscr{L}} \times \bigcup_{n\in\mathbb{N}} B^n \twoheadrightarrow \tilde{B}$; also obvious surjection $\tilde{B} \twoheadrightarrow B$ by fixing term $t(\mathsf{x}) := \mathsf{x}$ and plugging in/assigning any $b \in B$ to t (indeed $B \subseteq \tilde{B}$). If $|B| \geqslant |\mathscr{L}| \geqslant \aleph_0$, then the cardinality $|\mathcal{T}^{\mathscr{L}} \times \bigcup_{n\in\mathbb{N}} B^n| \geqslant |\tilde{B}|$ is (note also $\bigcup_{n\in\mathbb{N}} \mathscr{L}^n \twoheadrightarrow \mathcal{T}^{\mathscr{L}}$)

$$|\tilde{B}| \leqslant \left|\mathcal{T}^{\mathscr{L}} \times \bigcup_{n\in\mathbb{N}} B^n\right| = |\mathcal{T}^{\mathscr{L}}| + \left|\bigcup_{n\in\mathbb{N}} B^n\right| \leqslant \left|\bigcup_{n\in\mathbb{N}} \mathscr{L}^n\right| + \left|\bigcup_{n\in\mathbb{N}} B^n\right| \leqslant |\mathbb{N}| \cdot \max_{n\in\mathbb{N}} |\mathscr{L}^n| + |\mathbb{N}| \cdot \max_{n\in\mathbb{N}} |B^n|$$
$$= \aleph_0 \cdot |\mathscr{L}| + \aleph_0 \cdot |B| = \max\{\aleph_0, |\mathscr{L}|, |B|\} = |B| \leqslant |\tilde{B}|$$

where we use that $|\bigsqcup_{i\in I} \kappa_i| \leqslant |\bigsqcup_I (\max_{i\in I} \kappa_i)|$ due to obvious surjection $\twoheadleftarrow$; and that $\lambda \times \kappa \twoheadrightarrow \bigsqcup_\lambda \kappa$ (mapping (α, β) to the β in the αth copy of κ) hence $|\bigsqcup_\lambda \kappa| \leqslant \lambda \cdot \kappa$; and that $\kappa \cdot \lambda = \max\{\kappa, \lambda\} = \kappa + \lambda$ for infinite cardinals — see ProofWiki, which uses that for infinite cardinals, $\kappa \cdot \kappa = \kappa$ (MSE shows furthermore order-isomorphic); also Clark recommended Enderton's *Elements of Set Theory*: see pg. 162 on arithmetic of infinite cardinals for proofs (using Zorn's lemma) of $\kappa \cdot \kappa = \kappa$ and subsequently $\kappa + \lambda = \max\{\kappa, \lambda\} = \kappa \cdot \lambda$.

Actual Construction: we can now define inductively a substructure starting with A, and ensuring closure under existential witnesses, as required by Tarski-Vaught. First let $\mathcal{A}_0 := \langle A\rangle_{\mathcal{M}}$, and let A_0 be the base set of $\mathcal{A}_0$, i.e. $A_0 := \tilde{A}$. The **Observation** and the assumption $|A| \geqslant |\mathscr{L}|$ tell us $|A_0| = |A|$.

Assuming $A_i \subseteq M$ is defined (with $|A_i| \geqslant |\mathscr{L}|$), we define A_{i+1} as follows:

- for any $\mathscr{L}$-formula $\varphi(\mathsf{x}_0, \ldots, \mathsf{x}_n)$ and n-tuple $\overline{a} \in A_i^n$, if $\mathcal{M} \models \varphi(b_0, \overline{a})$ for some witness $b_0 \in M$, we choose one such b_0 and denote it by $c(\varphi, \overline{a}) \in M$;
- then we set $B_{i+1} := A_i \cup \{c(\varphi, \overline{a}) : \varphi(\mathsf{x}_0, \ldots, \mathsf{x}_n) \in \mathcal{F}ml^{\mathscr{L}}, \overline{a} \in A_i^n, n \in \mathbb{N}^+\}$. Note that surjection $A_i \cup \mathcal{F}ml^{\mathscr{L}} \times \bigcup_{i=1}^{\infty} A_i \twoheadrightarrow B_{i+1}$ implies (using that $|A_i| \geqslant |\mathscr{L}| = |\mathcal{F}ml^{\mathscr{L}}| \geqslant \aleph_0$) $|A_i| = |A_i| + |\mathcal{F}ml^{\mathscr{L}}| \cdot \aleph_0 \cdot |A_i| \geqslant |B_{i+1}|$, moreover equal because $B_{i+1} \supseteq A_i$.
- finally, we set $A_{i+1} := \tilde{B}_{i+1} :=$ the base set of $\langle B_{i+1}\rangle_{\mathcal{M}}$ (as above, $A_{i+1} \supseteq \tilde{B}_{i+1} \supseteq B_{i+1} \supseteq A_i$, and by **Observation**, $|A_{i+1}| := |\tilde{B}_{i+1}| = |B_{i+1}| = |A_i|$).

Let $M_0 := \bigcup_{i=1}^{\infty} A_i$. Then $M_0 \supseteq A$ (non-empty because we assumed $|A| \geqslant |\mathscr{L}|$), and M_0 also contains all $\mathsf{c}^{\mathcal{M}}$ (for any $\mathsf{c} \in \mathscr{C}^{\mathscr{L}}$) and closed under all $\mathsf{f}^{\mathcal{M}}$ for all $\mathsf{f} \in \mathscr{F}^{\mathscr{L}}$ (already $B_2 \subseteq A_2$ was closed under constants/functions, since $\mathsf{c}^{\mathcal{M}}$ witness to $\varphi(\mathsf{x}_0) := \mathsf{x}_0\mathsf{=c}$; and for all $\overline{a} \in M^n, \mathsf{f} \in \mathscr{F}_n^{\mathscr{L}}$, we know $\mathsf{f}^{\mathcal{M}}(\overline{A}) \in M$ is unique witness $c(\varphi, \overline{a})$ to $\varphi(\mathsf{x}_0, \dots, \mathsf{x}_n) := \mathsf{x}_0\mathsf{=f}\mathsf{x}_1 \dots \mathsf{x}_n$.)

Thus M_0 is the base set of the substructure $\mathcal{M}_0 := \langle M_0 \rangle_{\mathcal{M}}$. Now, to apply Tarski-Vaught to show $\mathcal{M}_0 \preceq \mathcal{M}$: let $\varphi(\mathsf{x}_0, \dots, \mathsf{x}_n)$ be $\mathscr{L}$-formula, $\overline{a} := (a_1, \dots, a_n) \in M_0^n$ and $b_0 \in M$ s.t. $\mathcal{M} \models \varphi(b_0, \overline{a})$. Then we know there exists $N \in \mathbb{N}$ s.t. the parameters $a_1, \dots, a_n \in A_N$ (because M_0 is union of increasing sets A_i). By construcution of A_{N+1} there exists $c_0 \in A_{n+1} \subseteq M_0$ s.t. $\mathcal{M} \models \varphi(c_0, \overline{a})$. Hence by Tarski-Vaught, $\mathcal{M}_0 \preceq \mathcal{M}$. And finally, $|M_0| = |\bigcup_{i=1}^{\infty} A_i| = \aleph_0 \cdot \max_{i \in \mathbb{N}} |A_i| = \max\{\aleph_0, |A|\} = |A|$ (because recall we found during construction that $|A_i| = \dots = |A_0| = |A|$). ■

Question: Mark: example of formula not reached by A_1? Artem: each step add new parameters, so need to meet Tarski-Vaught test for formulas involving those parameters, so abstractly this iteration is necessary. Can try to think up explicit example later... Artem tried to give example with group $(\mathbb{R}, +)$ and $A = \{1\}$ but $\mathscr{L}_{\text{group}}$ structure makes $\mathcal{A}_0$ already $(\mathbb{Z}, +)$; I also thought of σ-algebra (Lebesgue msble. sets in $\mathbb{R}, ^{\complement}, \cup, \cap$) and $A := \{(a, b) : a, b \in \mathbb{R}\}$, but again whole algebra generated by closure under $\cup, \cap, ^{\complement}$ in $\mathcal{A}_0$ (σ-algebra has $\bigcup_1^{\infty}, \bigcap_1^{\infty}$ as functions taking infinite arguments, not allowed; and I think even allowing that, all Borel sets still reached by $\mathcal{A}_0$)... after class Artem suggested $(\mathbb{R}, <)$ satisfies dense linear ordering, so taking $A = \mathbb{Z}$, no function symbols so $\mathcal{A}_0 = (\mathbb{Z}, +)$, and we can add Tarski-Vaught witnesses of $\varphi(x, y) := (\mathtt{x<y} \to \exists \mathtt{z}\; x\mathtt{<z<}y)$ to get $\frac{1}{2}\mathbb{Z}$; but there are more sentences (saying there exist two/three/four/etc. distinct elements between x, y), so still already by A_1 we get elementary $\mathscr{L}_{\text{order}}$-substructure of $(\mathbb{R}, <)$ (at the very least still get dense linear ordering, e.g. $\bigcup_{n=1}^{\infty} \frac{1}{2^n}\mathbb{Z}$)... harder than originally thought... EDIT 10/26/22: my misunderstanding here was that Tarski-Vaught only need witness to single variable formulas (the rest are paremeters), so no need to witness "there exist two/three/four distinct elements between", like I suggested above.

Answer (10/26/22): dense linear orders without endpoints, with $\mathcal{M} := (\mathbb{R}, <) \models \text{DLO}$. We haven't discussed this yet, but this theory has quantifier elimination QE. $A_0 = \{0\}$. Then no function symbols means $\tilde{B} = B$, so $B_1 = A_0$ again. Look at A_2. Any formula $\psi(\mathsf{x}_0, \mathsf{x}_1, \dots, \mathsf{x}_n)$, where $\mathsf{x}_1, \dots, \mathsf{x}_n$ parameters from previous level, namely all 0. From QE we know what formulas look like: some Boolean combination of (in)equalities, $\bigwedge \bigvee \mathsf{x}_1 \mathsf{<} \mathsf{x}_j$. We thus have to add at least one element > 0 and < 0, and that's all we have to add. Each step we only have to add finitely many elements, but any model of DLO is infinite, so need infinitely many steps. What about sentence expressing exists two distinct elements between two parameters (like I suggested in previous paragraph above)? Then can't just add one element? Artem: in QE we actually have sentence logically equiv. to q.f. formula with one free variable; not true that sentence necessarily logically equiv. to q.f. sentence.

10 Lecture 10

Recall last time, proved downward Löwenheim-Skolem theorem (↓LS): $\mathscr{L}$ first-order language, $\mathcal{M}$ be $\mathscr{L}$-structure, $A \subseteq M$ subset; then exists elementary substructure $\mathcal{M}_0 \preceq \mathcal{M}$ s.t. $A \subseteq M_0$ and $|M_0| = |A| + |\mathscr{L}| = \max\{|A|, |\mathscr{L}|\}$.

Example 10.1: Skolem "paradox"

If ZFC ($\mathscr{L}_{\text{set}}$-theory) is consistent, then it admits a countable model $\mathcal{V}$ (immediate corollary of ↓LS because $\mathscr{L}_{\text{set}}$ consists of one nonlogical symbol $\in$, so $|\mathscr{L}_{\text{set}}| = |\mathscr{Fml}^{\mathscr{L}_{\text{set}}}| = \aleph_0$; or suffices to use Henkin construction from Gödel's completeness theorem). But there are elements (sets) in the model $\mathcal{V}$ satisfying first-order formula expressing uncountability, i.e. $\nexists$ element of model which represents bijection between $\mathbb{N}$ and the set. That is, outside of $\mathcal{V}$, we know (can prove on the outside, i.e. prove formally in the meta-theory ZFC) that V is countable (and hence if $\mathcal{V}$ is transitive, all elements of V are also countable) via some witnessing bijection (encoded as a set), but none of these witnessing bijections are elements of $\mathcal{V}$ so from the perspective of $\mathcal{V}$, inside $\mathcal{V}$ there are no witnessing bijections.

Example 10.2: $\mathbb{R}$ elementarily equivalent to non-Archimedean model

Let $\mathcal{R} := (\mathbb{R}; 0, 1, +, \cdot, <)$ "ordered field of reals" be $\mathscr{L}_{\text{ord.ring}}$-structure. Then exists $\mathcal{R}' \equiv \mathcal{R}$ that is non-Archimedean, meaning that there exists an "infinitessimal" element $\epsilon > 0$ s.t. $n \cdot \epsilon := \underbrace{\epsilon + \ldots + \epsilon}_{n \text{ times}} < 1$ for any $n \in \mathbb{N}$.

Proof: ■ Here we just used compactness, like 220Ahw2p6; ↓LS used in next example.

Example 10.3: $\mathbb{R}$ elementarily equivalent to non-complete model

Proof: ■ *Note:* completeness $\implies$ Archimedean, since $\{n \cdot \epsilon : n \in \mathbb{N}\}$ bounded from above by $\leqslant 1$ but if s is LUB, then all $n \cdot \epsilon \leqslant s$ implies all $n \cdot \epsilon \leqslant s - \epsilon$, so $s - \epsilon$ is lesser upper bound than s; contradiction. So the first example alone aleady shows exists $\mathcal{R}' \equiv \mathcal{R}$ that is not complete.

More notes: so we saw it is possible to get an elementary substructure of $\mathcal{R}$ that is not complete, but conversely if we know $\mathcal{R}'$ is complete, then $\mathcal{R}' \simeq \mathcal{R}$ (as ordered fields) since all complete ordered fields are isomorphic (as ordered fields) to $\mathbb{R}$ (at least within the context of the background ZFC-metatheory; see also more MSE comments).

Ultimately, these 2 examples show us limitations of what is expressible in first-order formulas, since we know from real anlaysis/by def. of $\mathbb{R}$ that $\mathcal{R}$ satisifes completeness and Archimedean property, so these 2 results tell us Archimedean-ness and completeness are not expresssible by any collection of 1st-order sentences.

Remark: recall we discussed definable sets. Suppose $\mathcal{M} \preceq \mathcal{M}'$ and $D \subseteq M^n$ is a definable set (with parameters), i.e. $D = \varphi(\mathcal{M}, \overline{b}) := \{(a_1, \ldots, a_n) \in M^n : \mathcal{M} \models \varphi(\overline{a}, \overline{b})\}$ for fixed tuple $\overline{b}$ and fixed $\varphi(\overline{\mathsf{x}}, \overline{\mathsf{y}})$ an $\mathscr{L}$-formula. There is a canonical way to view D as a definable set in $\mathcal{M}'$ (Artem: "not canonical in terms of category, just canonical like... it feels... canonical"): define canonical extension $D' := \varphi(\mathcal{M}', \overline{b})$, which then satisfies $D' \cap M^n = D$ by elementarity. Note D' does not depend on choice of $\varphi(\overline{\mathsf{x}}, \overline{b})$ (OK, here's the "canonicity"), by elementarity. This allows us to view definable sets as not in one specific model, but in all elementary extensions of some model simultaneously.

10/26/22 Class 13

Definition 10.1: Axiomatizability

Let P be property that an $\mathscr{L}$-structure may or not satisfy. We say P is axiomatizable (resp. finitely axiomatizable) if there exists $\mathscr{L}$-theory T (resp. $\mathscr{L}$-sentence φ) s.t. for any $\mathscr{L}$-sentence $\mathcal{M}$, the property P is satisfied by $\mathcal{M} \iff \mathcal{M} \models T$ (resp. $\mathcal{M} \models \varphi$). In this case, we say T (resp. φ) axiomatizes property P.

Artem: "property" of $\mathscr{L}$-structures defined informally... corresponds to "class" of $\mathscr{L}$-structures but don't want to define class.

Proposition 10.1: Finite axiomatizable iff P and its negation both axiomatizable

Let P be a property of $\mathscr{L}$-structures. Then P is finitely axiomatizable iff P and its negation are both axiomatizable.

Proof: ($\Longrightarrow$): trivial: if φ axiomatized P, then $\neg\varphi$ axiomatizes negation of P.

($\Longleftarrow$): main tool, compactness. ■

Example 10.4: Axiomatizable and finite axiomatizable properties

(1) For any prime number p, fields of characteristic p are finitely axiomatizable in $\mathscr{L}_{\text{ring}}$ language, because property of being field axiomatized by φ_{field} (the conjunction of all the field axioms), and characteristic p axiomatized by $\varphi_p := \underbrace{1 + \ldots + 1}_{p \text{ times}} = 0$; so $(\varphi_{\text{field}} \wedge \varphi_p)$ is sentence that axiomatizes property of being field of char. p.

(2)

(3)

Diagram Method

General method of building/constructing elementary extension. ↓LS proof gave general method for finding elementary substructure (start with seed set A, grow upwards), i.e. going up instead of down. Towards this end, we formulate some syntactic condition — consistency of certain expansion. "So we do something silly: add a new constant for every element of model." Consider language $\mathscr{L}_{\mathcal{M}}$

Definition 10.2: Complete diagram of $\mathcal{M}$

awef

Similarly we can define

Definition 10.3: Simple diagram of $\mathcal{M}$

The simple diagram of $\mathcal{M}$, denoted $\Delta(\mathcal{M}) \subseteq D(\mathcal{M})$ consists of all $\mathscr{L}_{\mathcal{M}}$-sentences $\varphi(c_{m_1}, \ldots, c_{m_n})$ where φ is q.f. $\mathscr{L}$-formula, and again $\overline{m} \in M^n$ s.t. $\mathcal{M} \models \varphi(\overline{m})$.

In other words we have defined two ways to associate a theory to every $\mathscr{L}$-structure. These definitions are justified by the following propositions:

Proposition 10.2: (Reducts of) models of complete diagram correspond to elementary extensions

($\mathscr{L}$-reducts of) $\mathscr{L}_{\mathcal{M}}$-models of $D(\mathcal{M})$ correspond, up to $\mathscr{L}$-iso., to elementary extensions of $\mathcal{M}$.

Proof: ■

Proposition 10.3: (Reducts of) models of simple diagram correspond to arbitrary extensions

($\mathscr{L}$-reducts of) $\mathscr{L}_{\mathcal{M}}$-models of $\Delta(\mathcal{M})$ correspond, up to $\mathscr{L}$-iso., to arbitrary extensions of $\mathcal{M}$.

Proof: ■

What is the point? To build elementary extension of $\mathcal{M}$, equivalent to building model of $D(\mathcal{M})$, and the compactness theorem tells us it suffices to build a model for any finite subset of $D(\mathcal{M})$! That is the "diagram method".

11 Lecture 11

Theorem 11.1: Upwards Löwenheim-Skolem (↑LS)

Let $\mathcal{M}$ be an infinite $\mathscr{L}$-structure, and let κ be cardinal s.t. $\kappa \geqslant |M| + |\mathscr{L}| = \max\{|M|, |\mathscr{L}|\}$. Then, there exists an elementary extension of cardinality κ.

Proof: Idea: building model of $D(\mathcal{M})$ of large cardinality easy; just add lots of constant symbols, and have many sentences expressing all the constants distinct. Finite portion of this theory always satisfiable because $\mathcal{M}$ itself (infinite by assumption) definitely has arbtrarily many finite distinct elements.

Rigorously, first note: enought to construct .. ■

Corollary 11.1: Löwenheim-Skolem

For $\mathscr{L}$-theory T with infinite model $\mathcal{M}$, we know T has model $\mathcal{M}_\kappa$ of cardinality κ for any $\kappa \geqslant |\mathscr{L}|$ (indeed, could consider only elementary substructure/extension of $\mathcal{M}$, which in particular by Lec. 9 $\rightsquigarrow \mathcal{M}_\kappa \equiv \mathcal{M}$).

Artem: "It's magic... at no point did we actually do anything [infinite, I guess] but somehow we got models of every possible cardinality". What LS tells us about FO logic: "Among infinite sizes, size is completely irrelevant (at least in the eyes of FO logic)".

Limbo (only def. of conservative extension used in Class 14)

Expansions by Definition

When studying $\mathscr{L}$-structures, may be useful to enrich language with new constants/functions/realtions. Here we describe how to do this (formally) without changing expressibility of language.

Definition 11.1: Notation for "exists unique"

$\exists!\mathsf{x}\psi$ (read: "there exists unique x s.t. ψ") is an abbr. for the formula $\exists\mathsf{x}(\psi \wedge \forall\mathsf{y}(\psi_{\mathsf{y}/\mathsf{x}} \rightarrow \mathsf{x}{=}\mathsf{y}))$, where y is a variable distinct from x (don't need $\mathsf{y} \notin \mathrm{Free}(\psi)$ or anything like that, because we defined substitution of terms into variables carefully — recall (Def. 3.2) the main idea is that if there is "variable collision", then we rename the bound variable to something that won't cause "collisions").

Definition 11.2: Expansion by definition

Let T be $\mathscr{L}$-theory, and suppose we have a larger language $\mathscr{L}' \supseteq \mathscr{L}$. Assume that

- for any n-ary relation symbol $\mathsf{R} \in \mathscr{L}' \setminus \mathscr{L}$, there is some associated $\mathscr{L}$-formula $\varphi_{\mathsf{R}}(\mathsf{x}_1, \ldots, \mathsf{x}_n)$ (no further specifications);
- for any n-ary function symbol $\mathsf{f} \in \mathscr{L}' \setminus \mathscr{L}$, there is some associated $\mathscr{L}$-formula $\varphi_{\mathsf{f}}(\mathsf{x}_0, \ldots, \mathsf{x}_n)$ s.t. $T \models \forall\mathsf{x}_1 \ldots \forall\mathsf{x}_n \exists!\mathsf{x}_0\varphi_{\mathsf{f}}$

- for any constant symbol $\mathsf{c} \in \mathscr{L}' \setminus \mathscr{L}$, there is some associated $\mathscr{L}$-formula $\varphi_{\mathsf{c}}(\mathsf{x}_0)$ s.t. $T \models \exists! \mathsf{x}_0 \varphi_{\mathsf{c}}$

Then, the $\mathscr{L}'$-theory

$$\begin{aligned} T' := T &\cup \{\forall \mathsf{x}_1 \ldots \forall \mathsf{x}_n (\varphi_{\mathsf{R}}(\mathsf{x}_1, \ldots, \mathsf{x}_n) \leftrightarrow \mathsf{R}\mathsf{x}_1 \ldots \mathsf{x}_n) : \mathsf{R} \in \mathscr{L}' \setminus \mathscr{L} \text{ relation symbol}\} \\ &\cup \{\forall \mathsf{x}_1 \ldots \forall \mathsf{x}_n \, \varphi_{\mathsf{f}}(\mathsf{f}\mathsf{x}_1 \ldots \mathsf{x}_n, \mathsf{x}_1, \ldots, \mathsf{x}_n) : \mathsf{f} \in \mathscr{L}' \setminus \mathscr{L} \text{ function symbol}\} \\ &\cup \{\varphi_{\mathsf{c}}(\mathsf{c}) : \mathsf{c} \in \mathscr{L}' \setminus \mathscr{L} \text{ constant symbol}\} \end{aligned}$$

is called an expansion by definition of T. Intuitively, we added new symbols but really these are "shorthand" for relations/functions/constants already "expressed" by $\mathscr{L}$-formulas

Definition 11.3: Equivalence in a theory

Formulas $\varphi_i(\mathsf{x}_1, \ldots, \mathsf{x}_n)$ $(i = 1, 2)$ are called equivalent in theory T if $T \vdash \forall \mathsf{x}_1 \ldots \forall \mathsf{x}_n (\varphi_1 \leftrightarrow \varphi_2)$, and are called logically equivalent if equivalent in the empty theory $T = \varnothing$.

Lemma 11.1: Logically equivalent to height $\leqslant 1$

Any formula is logically equivalent to a formula with all terms of height $\leqslant 1$. (Recall Lec. 1: term is of the form constant symbol/variable symbol/function evaluation, and height of term is (maximum) number of nested function calls)

Proof: *Idea:* a term $t := \mathsf{f}(\mathsf{f}_1(\bar{\mathsf{x}}_1), \ldots, \mathsf{f}_k(\bar{\mathsf{x}}_k), \text{constants and variables}))$ can be "replaced" by something like $\mathsf{f}(\mathsf{y}_1, \ldots, \mathsf{y}_k, \text{constants and variables}) \wedge \bigwedge_1^k \mathsf{y}_i{=}\mathsf{f}_i(\bar{\mathsf{x}}_i)$. (Obviously at face value this is nonsense since what I just wrote isn't a term — it contains conjunctions for instance — but this idea of taking nested functions outside via conjunction is key.)

Rigorously: for formula φ, let $M(\varphi)$ denote maximal height of a term contained in φ.
Ooops, maks eas osie aoig aois ■

12 Lecture 12

10/28/22 Class 14

Recall from 220Ahw1p9 that every $\mathscr{L}$-formula φ is logically equivalent to a formula in prenex normal form, i.e. $\varphi \rightarrow \mathsf{Q}_1\mathsf{x}_1 \ldots \mathsf{Q}_n\mathsf{x}_n\psi(\mathsf{x}_1, \ldots, \mathsf{x}_n)$ where $\mathsf{Q}_i \in \{\exists, \forall\}$ and ψ is q.f. formula. It turns out ...

"Studies in psychology people only understand 3 quantifier complexity (e.g. uniform continuity)... after that the brain breaks. Usually bad sign that formulas get really complex, means the theory can not be fully understood." Notorious example is Peano arithmetic $\subseteq \mathrm{Th}(\mathbb{N}, +, \cdot)$. Famous result of Gödel there is formula of complexity n not equivalent to formula of complexity $< n$. Many bad things happen in PA (Gödel incompleteness, Tarski's undefinability of truth, etc.). But many theories, suffice to understand formulas without quantifiers, which makes everything much easier!

Theorem 12.1: Equivalent definitions of quantifier elimination

Let T be any $\mathscr{L}$-theory, $n \in \mathbb{N}^+$, and fix any $\mathscr{L}$-formula $\varphi(\mathsf{x}_1, \ldots, \mathsf{x}_n)$. Then, t.f.a.e.:

(1) φ is equivalent in T to some q.f. $\mathscr{L}$-formula $\psi(\mathsf{x}_1, \ldots, \mathsf{x}_n)$.

(2) for any two models $\mathcal{M}, \mathcal{N}$ of T, any common $\mathscr{L}$-substructure $\mathcal{A} \subseteq \mathcal{M}, \mathcal{N}$, and all $\overline{a} \in A^n$, we have that $\mathcal{M} \models \varphi(\overline{a}) \iff \mathcal{N} \models \varphi(\overline{a})$. It is perhaps more natural to instead say "common substructure up to isomorphism", since the specific base sets don't matter, merely the relationship between the elements as encoded by the $\mathscr{L}$-structures in question. Ultimately it doesn't matter because although my reformulation looks stronger at face-value, if we have $\mathcal{A}, \mathcal{A}'$ substructures of $\mathcal{M}, \mathcal{M}'$ resp. and $\mathcal{A} \cong \mathcal{A}'$, then we can find $\widetilde{\mathcal{M}}, \widetilde{\mathcal{M}}'$ isomorphic to $\mathcal{M}, \mathcal{M}'$ resp. s.t. $\widetilde{\mathcal{M}}, \widetilde{\mathcal{M}}'$ have base sets that actually literally contain the base set A, thus reducing to (2) as written.

(3) *(my own addition)* for any $\mathscr{L}$-structure $\mathcal{A}$, and extensions $\mathcal{M}, \mathcal{N}$ of $\mathcal{A}$ that model T, we have that $\mathcal{M}, \mathcal{N}$ agree on the truth value of $\varphi(\overline{a})$ for all parameters $\overline{a} \in A^n$.

In other words, this theorem tells us the only way truth of φ could be "independent" of the model of T — i.e. $\varphi(a_1, \ldots, a_n)$ true in all models of T in which a copy of the elements $a_1, \ldots, a_n$, related to each other in some fixed predetermined way (as encoded by the $\mathscr{L}$-structure $\mathcal{A}$), exist — is if φ is (logically equivalent in T to) a q.f. formula.

Remark: what if $n = 0$, i.e. φ is a sentence? Can consider φ as a formula $\varphi(\mathsf{x})$ and apply Thm to get a q.f. formula $\psi(\mathsf{x})$, which is equivalent in T to $\varphi(\mathsf{x})$. Example: if $\varphi :\equiv \exists\mathtt{y(y=y)}$ (which is provably from the empy theory, see Lec. 5), then we can take $\psi(\mathsf{x}) :\equiv \mathtt{x=x}$ because indeed $\varnothing \vdash \forall\mathtt{x(}\exists\mathtt{y(y=y)} \leftrightarrow \mathtt{x=x)}$.

Note also if $\mathscr{L}$ has no constant symbols, then there are no quantifier free $\mathscr{L}$-sentences (just by syntactic rules of sentence formation). In this case, (in the course of the proof of Thm) when we assert the existence of q.f. formula ψ equiv to sentence φ (in the sense of the Rmk. above), we allow ψ to have a single free variable.

Another way to address this isue, is to always require languages to possess logical (constant) symbols for `True` and `False`.

My remark: maybe the hope that quantifier elimination is even possible in lots of cases can be first motivated by e.g. https://mathoverflow.net/questions/33563/intuition-for-model-theoretic-proof-of-the-nu
Lemma 12.1 also makes QE seem a lot more manageable, i.e. only need to focus on $\varphi :\equiv \exists \mathsf{v}_0(\text{q.f. fml})$. The semantic version of this, i.e. Cor. 12.1 is also easy to check in e.g. ACF (single variable polys behave

very well in ACF). Maybe proof of Thm. 12.1 is conceptually easier if consider only $\varphi := \exists \mathsf{v}_0(\text{q.f. fml})$? It's very interesting that the syntactic side, reducing to the one-variable-existence case is trivial; but on the semantic side, such a reduction (I'm thinking even in just ACF case) is non-trivial. E.g. we know from all this that it should be possible to go from [$K \subseteq L$ ACFs, φ Boolean combination of polynomial equations and $\overline{a} \in K^n \implies$ every solution $b \in L$ s.t. $L \models \varphi(b, \overline{a})$ can be pushed down to $c \in K$ s.t. $K \models \varphi(c, \overline{a})$] to the same result but for general φ, or even just $\varphi = \exists \overline{\mathsf{x}} \psi(\overline{\mathsf{x}})$ for q.f. ψ, which is all I need for Nullstellensatz (4/20/25).

No me gusta tampoco (not clear a priori to only focus on parameters common from common substructure - why not common $\mathcal{A} \to \mathcal{M}, \mathcal{A} \to \mathcal{N}$? or common set instead of substructure? — WAIT substructure is just common set that contains interpretations of constant symbols, and ensure that $\mathcal{M}, \mathcal{N}$ function and relation symbols' interpretations are compatible on the common set)**:::** *My remark:* I think that the semantic version ((2) in Thm above) of φ equiv. in T to q.f. formula to be more intuitive/easier to motivate.

Consider how we use this Thm in Prop. 12.1(1): we use it to show model-completeness of a theory T: all models of T agree on truth of sentences. So (2) in Thm above can be thought of as a natural strengthening of that notion to "model-completeness relative to common substructures (that expand the language with new parameters that one can substitute into formal variable symbols)": where now all models agree on truth of formulas with parameters, where in order to make sense of this we need to draw the parameters from (and thus consider only models extending) arbitrary common substructures $\mathcal{A}$ (i.e. expand the language $\mathscr{L}$ with constant symbols from a substructure, and models agree on truth of sentences in that expanded language; I guess we should also ask that the models model the expanded theory $T \cup D(\mathcal{A})$ the complete diagram of the chosen substructure $\mathcal{C}$).

No me gusta::: In 220C, I remember Merk showing something in representation theory was true in any alg. closed field (char. 0) iff true in $\mathbb{C}$ iff ture in $\overline{\mathbb{Q}}$ using only tools of algebra (or something along these lines). As I mention in this lecture notes, extension preserve truth values of q.f. formulas ("trivial observation"), may lead to wondering if it is possible for a non q.f. formula to also preserve truth values under arbitrary extension and substructure (as long as the substructure contains the necessary constants/parameters).

Clue2: maybe playing around in algebra/alg. geom., somehow figure out that only definable sets are Boolean combinations of vanishing sets of polynomials $\rightsquigarrow$ quantified formulas may be equivalent to q.f. formulas. Maybe proof in ACF case without model theory helps provide hints for how to approach this Thm in general model theory case? I am not so compelled by this Clue2.

My warning: in my handwritten notes (yellow stickynote 9/27/22?), I mistakenly thought (2) above is equivalent to "for all models $\mathcal{A} \models T$ and all $\overline{a} \in A^n$, $\varphi(\overline{a})$ is sasifed in $\mathcal{A}$ iff satisfied in any extension of $\mathcal{A}$". This is not right, because (2) does not need $\mathcal{A}$ to be model of T; it says substructure, not elementary substructure! Pretty sure (3) as I've written above is OK though.

Proof (of Thm. 12.1): (1) $\implies$ (2): any formula $\varphi(\overline{\mathsf{x}})$ is equivalent in T to a q.f. formula $\psi(\overline{\mathsf{x}})$, so for any $\mathscr{L}$-substructure and $\mathscr{L}$-models $\mathcal{A} \subseteq \mathcal{M}, \mathcal{N} \models T$ and any $\overline{a} \in A^{\bullet}$, we have $\mathcal{M}, \mathcal{N} \models \forall \overline{\mathsf{x}}(\varphi \leftrightarrow \psi) \implies$ for any $\overline{a} \in A^{\bullet}$,

$$\mathcal{M} \models \varphi(\overline{a}) \iff \mathcal{M} \models \psi(\overline{a}) \iff \mathcal{A} \models \psi(\overline{a})$$

where we used that ψ q.f. means any extension/substructure preserves truth value (220Ahw1p7.1).

The same is true for $\mathcal{N}$ in place of $\mathcal{M}$, so indeed

$$\mathcal{M} \models \varphi(\bar{a}) \iff \mathcal{M} \models \psi(\bar{a}) \iff \mathcal{A} \models \psi(\bar{a}) \iff \mathcal{N} \models \psi(\bar{a}) \iff \mathcal{N} \models \varphi(\bar{a}).$$

---

(2) $\implies$ (1): our goal is to find a q.f. formula $\chi^\star$ s.t. "$T \vdash (\chi^\star \leftrightarrow \varphi)$". We define the "haystack" Γ, in which we search for our desired needle $\chi^\star$: $\Gamma(\bar{\mathsf{x}}) := \{\chi(\mathsf{x}_1, \ldots, \mathsf{x}_n) \text{ q.f. } \mathscr{L}\text{-fmls.} : T \models \forall \mathsf{x}_1 \ldots \mathsf{x}_n (\varphi \to \chi)\}$ ("the set of all q.f. $\mathscr{L}$-formulas $\chi(\bar{\mathsf{x}})$ with $\leqslant n$ free variables that are implied in T by φ", i.e. the set of all q.f. $\chi(\bar{\mathsf{x}})$ s.t. $\forall \mathcal{M} \models T$ and $\forall$ 'satisfiers' $\bar{s} \in M^n$ s.t. $\mathcal{M} \models \varphi(\bar{s})$, also satisfy $\mathcal{M} \models \chi(\bar{s})$), so that our goal can be rephrased as finding some $\chi^\star(\bar{\mathsf{x}}) \in \Gamma(\bar{\mathsf{x}})$ to conversely imply φ in T, or equivalently (by compactness) $T \cup \Gamma(\bar{\mathsf{x}}) \models \varphi(\bar{\mathsf{x}})$.

We introduce new constants in place of $\bar{\mathsf{x}}$ ("simulating $\bar{\mathsf{x}}$") because it's easier to work with those to create models — in particular we defined theories to consist of sentences, not formulas: let $\mathsf{c}_1, \ldots, \mathsf{c}_n$ be new pairwise distinct constant symbols, and define $\Gamma(\bar{\mathsf{c}}) := \{\chi(\mathsf{c}_1, \ldots, \mathsf{c}_n) : \chi \in \Gamma(\bar{\mathsf{x}})\}$ (i.e. the substitution of c_i for x_i in the formulas χ), which is then a theory in the language $\mathscr{L}' := \mathscr{L} \cup \{\mathsf{c}_1, \ldots, \mathsf{c}_n\}$.

- **Claim:** the $\mathscr{L}'$-theory $T \cup \Gamma(\bar{\mathsf{c}})$ proves/models $\models \varphi(\bar{\mathsf{c}})$ (a $\mathscr{L}'$-sentence). ($T \cup \Gamma(\bar{\mathsf{c}})$ has a $\mathscr{L}'$-model, because by assumption (2) in Thm, we can assume we have an $\mathscr{L}$-model $\mathfrak{N} \models T$ satisfying $\mathfrak{N} \models \varphi(\bar{s})$ (for 'satisfiers' $\bar{s} \in \mathfrak{N}^n$), which implies $\mathfrak{N} \models \chi(\bar{s})$ for all $\chi(\bar{\mathsf{x}}) \in \Gamma(\bar{\mathsf{x}})$, by def. of Γ. Thus, interpreting $\mathsf{c}_i^{\mathfrak{N}'} := s_i \in \mathfrak{N}$, we get $\mathfrak{N}' \models T \cup \Gamma(\bar{\mathsf{c}})$.)
- *Proof (spread out in following ►):* suppose not, $\not\models \varphi(\bar{\mathsf{c}})$. Then, by the completeness theorem,
- $\exists$ an $\mathscr{L}'$-model $\mathcal{M}' \models T \cup \Gamma(\bar{\mathsf{c}}) \cup \{\neg\varphi(\bar{\mathsf{c}})\}$.
- Let $\mathcal{A}' := \langle \mathsf{c}_1^{\mathcal{M}'}, \ldots, \mathsf{c}_n^{\mathcal{M}'} \rangle_{\mathcal{M}'}$ be the $\mathscr{L}'$-substructure of $\mathcal{M}'$ generated by the interpretations of the new constant symbols $\mathsf{c}_1^{\mathcal{M}'}, \ldots, \mathsf{c}_n^{\mathcal{M}'} \in \mathcal{M}'$ — the $\langle \bullet \rangle_{\mathcal{M}'}$ notation defined at end of Lec9 ('Class12').
- Let A denote the base set of $\mathcal{A}'$. In particular, note that
- all $a_i := \mathsf{c}_i^{\mathcal{M}'} \in A$.
- The STRATEGY, is to build an $\mathscr{L}'$-model $\mathcal{N}^* \models T \cup \{\varphi(\bar{\mathsf{c}})\}$ that has an $\mathscr{L}'$-substructure isomorphic to $\mathcal{A}'$. This reaches a contradiction, because we will get two $\mathscr{L}$-structures $\mathcal{N} := \mathcal{N}^* \restriction_{\mathscr{L}}$, $\mathcal{M} := \mathcal{M}' \restriction_{\mathscr{L}}$ both modelling T, and both containing (up to iso.) the common $\mathscr{L}$-substructure $\mathcal{A} := \mathcal{A}' \restriction_{\mathscr{L}}$, in particular both containing $a_i := \mathsf{c}_i^{\mathcal{M}'} \in A$, which satisfy $\mathcal{N} \models \varphi(\bar{a})$ but $\mathcal{M} \models \neg\varphi(\bar{a})$; contradiction to assumption (2) in Thm.
  We can ensure the $\mathscr{L}'$-model $\mathcal{N}^*$ contains (an isomorphic copy of) $\mathcal{A}'$ by asking $\mathcal{N}^*$ to also model the simple diagram (of the $\mathscr{L}$-reduct) $\Delta(\mathcal{A}' \restriction_{\mathscr{L}})$, because by definition of simple diagram Δ, any $\mathscr{L}'$-structure $\mathcal{N}^*$ that models $\Delta(\mathcal{A}' \restriction_{\mathscr{L}})$ contains an isomorphic copy of $\mathcal{A}'$ as an $\mathscr{L}'$-substructure.
- Let us refresh ourselves on simple diagrams, by making the exercise observation that our haystack ("substituting $\bar{\mathsf{c}}/\bar{\mathsf{x}}$ in the set of all q.f. $\mathscr{L}$-formulas $\chi(\bar{\mathsf{x}})$ that are implied in T by φ") $\Gamma(\bar{\mathsf{c}}) \subseteq \Delta(\mathcal{A}' \restriction_{\mathscr{L}})$, the simple diagram of the $\mathscr{L}$-reduct $\mathcal{A}' \restriction_{\mathscr{L}}$ — which recall (from end of Lec. 10) is the set of all $\mathscr{L}_A$-sentences $\phi(\mathsf{c}_{m_1}, \ldots, \mathsf{c}_{m_?})$ (with ϕ a q.f. $\mathscr{L}$-fml.) satisfied in $\mathscr{L}_A$-str.: $\langle A; \ldots \rangle$ with interpretations of $\mathscr{L}$-symbols coming from $\mathcal{A}' \restriction_{\mathscr{L}}$, and new constant symbols $\mathsf{c}_m \in \mathscr{L}_A \setminus \mathscr{L}$ interpreted to $m \in A$. Our current c_i would be written in the old notation as $\mathsf{c}_{(\mathsf{c}_i^{\mathcal{M}'})}$, giving an inclusion $\mathscr{L}' \hookrightarrow \mathscr{L}_A$. The containment ("$\Gamma \subseteq \Delta$") is true because

$$\mathcal{M}' \models \Gamma(\bar{\mathsf{c}}) \text{ (q.f. } \mathscr{L}'\text{-fmls.!) and } \mathcal{A}' \subseteq \mathcal{M}' \implies \mathcal{A}' \models \Gamma(\bar{\mathsf{c}}) \iff \mathcal{A}' \restriction_{\mathscr{L}} \models \Gamma(\mathsf{c}_1^{\mathcal{M}'}, \ldots, \mathsf{c}_n^{\mathcal{M}'}).$$

  Henceforth, the notation $\Delta(\mathcal{A}')$ means $\Delta(\mathcal{A}' \restriction_{\mathscr{L}})$ (typecheckers recall: it's a set of q.f. $\mathscr{L}_A$-sentences).
- As outlined above, the proof hinges upon the following **Subclaim** ("constructing $\mathcal{N}^*$ opposing $\mathcal{M}'$"):

– **Subclaim:** the theory $\Sigma := T \cup \Delta(\mathcal{A}') \cup \{\varphi(\bar{\mathsf{c}})\}$ has an $\mathscr{L}_A$-model.

  * Suppose not. Then every $\mathscr{L}_A$-model of $T \cup \Delta(\mathcal{A}')$ satisfies $\neg\varphi(\bar{\mathsf{c}})$, so by Gödel completeness (or Cor. 5.2 of Deduction Lemma in Lec. 5) $T \cup \Delta(\mathcal{A}') \vdash \neg\varphi(\bar{\mathsf{c}})$.

  * Now, every element $a \in A$ can be written as an $\mathscr{L}'$-term (without free variables) $t_a^{\mathcal{M}'}(\bar{\mathsf{c}})$ (from the explicit description of base sets of generated structures from Lec. 9: the generated substructure $\mathcal{A}' := \langle a_1, \ldots, a_n \rangle_{\mathcal{M}'} = \bigcap_{\{a_1,\ldots,a_n\} \subseteq \mathfrak{N} \subseteq \mathcal{M}'} \mathfrak{N}$ has base set

$$\left\{ \text{interpretations } t^{\mathcal{M}'}(\bar{a}) : \mathscr{L}'\text{-terms } t(\bar{\mathsf{x}}) \in \mathscr{T}^{\mathscr{L}'}, \bar{a} \in \{a_1, \ldots, a_n\}^{?} \right\}.$$

  For example for $\mathscr{L} = \mathscr{L}_{\text{ring}}$ and $T = \text{ACF}$, substructure = subring, and subring generated by $a_1, \ldots, a_n$ in ACF $\mathcal{M}'$ has base set given by all finite combinations of $a_1, \ldots, a_n$ by $+, -, \cdot$ = interpretations of $\mathscr{L}_{\text{ring}}$ terms).

  * Denote by $\Delta_{\bar{\mathsf{c}}}(\mathcal{A}')$ the set of q.f. $\mathscr{L}'$-sentences in $\Delta(\mathcal{A}')$. Note $\Gamma(\bar{\mathsf{c}}) \subseteq \Delta_{\bar{\mathsf{c}}}(\mathcal{A}') \subseteq \Delta(\mathcal{A}')$.

  * We have that the $\mathscr{L}_A$-theory $T \cup \Delta(\mathcal{A}')$ is a conservative expansion of $\mathscr{L}'$-theory $T \cup \Delta_{\bar{\mathsf{c}}}(\mathcal{A}')$ (recall Lec. 11, means that both theories exactly agree on provability of an $\mathscr{L}'$-sentence), because it is essentially an expansion by definition (meaning the only new sentences in larger theory are those literally *defining* the new relation/function/constant symbols in the larger language in terms of formulas in the smaller language):

    · $\mathscr{L}_A \supseteq \mathscr{L}'$ only adds more constant symbols, and so the $\mathscr{L}_A$-theory $U := T \cup \Delta_{\bar{\mathsf{c}}}(\mathcal{A}') \cup \{\mathsf{c}_a{=}t_a(\bar{\mathsf{c}}) : a \in A\}$ (i.e. in notation of Lec. 11, $\varphi_{\mathsf{c}_a} := \mathsf{x}_0{=}t_a(\bar{\mathsf{c}})$) is exp. by def. of $T \cup \Delta_{\bar{\mathsf{c}}}(\mathcal{A}')$ (hence conservative by Prop. of Lec. 11).

    · And all $\mathscr{L}_A$-sentences $\varphi(\mathsf{c}_{a_1}, \ldots, \mathsf{c}_{a_n})$ in $\Delta(\mathcal{A}')$ are equivalent in U to $\varphi(t_{a_1}(\bar{\mathsf{c}}), \ldots, t_{a_n}(\bar{\mathsf{c}}))$ hence (by Computation at end of Lec. 11) $T \cup \Delta(\mathcal{A}')$ is conservative exp. of U , and stacking conservative extensions remain conservative exp., so indeed $T \cup \Delta(\mathcal{A}')$ is conservative exp. of $T \cup \Delta_{\bar{\mathsf{c}}}(\mathcal{A}')$.

  * Then in particular (using def. of conservative exp.), $T \cup \Delta_{\bar{\mathsf{c}}}(\mathcal{A}') \vdash \neg\varphi(\bar{\mathsf{c}})$.

  * To summarize, although may have terms involving constants not in $\bar{\mathsf{c}} = (\mathsf{c}_1, \ldots, \mathsf{c}_n) = (\mathsf{c}_{a_1}, \ldots, \mathsf{c}_{a_n})$ in $\Delta(\mathcal{A}')$ in the proof $T \cup \Delta(\mathcal{A}') \vdash \neg\varphi(\bar{\mathsf{c}})$, we use that all such terms can be written in terms of $\bar{\mathsf{c}}$, so the smaller theory already proves it: $T \cup \Delta_{\bar{\mathsf{c}}}(\mathcal{A}') \vdash \neg\varphi(\bar{\mathsf{c}})$.

  * By compactness, there exists q.f. $\mathscr{L}$-formulas $\xi_1(\bar{\mathsf{x}}), \ldots, \xi_k(\bar{\mathsf{x}})$ s.t. all $\xi_i(\bar{\mathsf{c}}) \in \Delta_{\bar{\mathsf{c}}}(\mathcal{A}')$ and $T \cup \{\xi_i(\bar{\mathsf{c}})\}_{i=1}^k \vdash \neg\varphi(\bar{\mathsf{c}})$, which is iff (by Deduction Lemma of Lec. 5, and trivial conjunction step)

$$T \models \underbrace{\bigwedge_{i=1}^k \xi_i(\bar{\mathsf{c}})}_{=:\xi(\bar{\mathsf{c}})} \to \varphi(\bar{\mathsf{c}}).$$

  I point out that obviously, $\Delta_{\bar{\mathsf{c}}}(\mathcal{A}') \models \xi(\bar{\mathsf{c}})$. (We will later contradict this.)

  * However, note that the constant symbols $\mathsf{c}_1, \ldots, \mathsf{c}_n$ don't appear in T, nor in $\varphi(\bar{\mathsf{x}})$ nor $\xi(\bar{\mathsf{x}})$. Thus, using Simulation of Contants by Variables lemma (Lec. 6), we get that $T \models (\xi(\bar{\mathsf{x}}) \to \neg\varphi(\bar{\mathsf{x}}))$, and by (Gen), $T \models \forall\bar{\mathsf{x}}(\xi(\bar{\mathsf{x}}) \to \neg\varphi(\bar{\mathsf{x}}))$.

  * Taking contrapositive, $T \models \forall\bar{\mathsf{x}}(\varphi(\bar{\mathsf{x}}) \to \neg\xi(\bar{\mathsf{x}}))$.

  * As all $\xi_i(\bar{\mathsf{x}})$ were q.f., $\neg\xi_i(\bar{\mathsf{x}})$ is too, so $\neg\xi(\bar{\mathsf{x}}) \in \Gamma(\bar{\mathsf{x}})$, so $\neg\xi(\bar{\mathsf{c}}) \in \Gamma(\bar{\mathsf{c}}) \subseteq \Delta(\mathcal{A}')$.

  * This contradicts what we said earlier: $\Delta_{\bar{\mathsf{c}}}(\mathcal{A}') \vdash \xi(\bar{\mathsf{c}}) \implies \Delta(\mathcal{A}') \vdash \xi(\bar{\mathsf{c}})$. This is a genuine contradiction because we know $\Delta(\mathcal{A}')$ is consistent (it's a subset of the complete diagram, which is Th(a model), hence always complete $\implies$ consistent).

– So, we have shown the **Subclaim**, i.e. $\Sigma := T \cup \Delta(\mathcal{A}') \cup \{\varphi(\bar{c})\}$ has a model.

10/31/22 Class 15

Tidying up/summarizing and fulfilling our promises:

- Returning to the original **Claim:** now we know the $\mathcal{L}'$-theory Σ has a $\mathcal{L}'$-model $\mathcal{N}^*$, and the $\mathcal{L}$-reduct $\mathcal{N} := \mathcal{N}^*\restriction_{\mathcal{L}}$ contains an isomorphic copy $\mathcal{B}'$ of $\mathcal{A}'$, because in particular $\mathcal{N}^*$ is model of $\Delta(\mathcal{A}') \subseteq \Sigma$ (and any model of simple diagram of a structure contains an isomorphic copy of that structure as substructure; also going to $\mathcal{L}$-reduct only throws away irrelevant interpretations in this case).

  So up to identification of $\mathcal{B}'$ and $\mathcal{A}'$, we have constructed two models $\mathcal{M} := \mathcal{M}'\restriction_{\mathcal{L}}$ and $\mathcal{N}$ of T, containing a common substructure $\mathcal{A} := \mathcal{A}'\restriction_{\mathcal{L}}$, s.t. if we set $a_i := \mathsf{c}_i^{\mathcal{M}'}$, then defining $\bar{a} := (a_i)_1^n$ we get $\mathcal{N} \models \varphi(\bar{a})$ ("a sentence in theory Σ"), but recall from first line of **Claim** that by definition $\mathcal{M}' \models \neg\varphi(\bar{a})$, which contradicts property (2) in the theorem. So, the **Claim:** $T \cup \Gamma(\bar{c}) \models \varphi(\bar{c})$ is proven.

Finally, by compactness, there exists $\xi_1(\bar{\mathsf{c}}), \ldots, \xi_m(\bar{\mathsf{c}}) \in \Gamma(\bar{\mathsf{c}})$ s.t. (similar to before)

$$T \models \bigwedge_{i=1}^{n} \xi_i(\bar{\mathsf{c}}) \to \varphi(\bar{\mathsf{c}}), \text{ and again as before, } T \models \forall \bar{\mathsf{x}} (\underbrace{\textstyle\bigwedge_{i=1}^{n} \xi_i(\bar{\mathsf{x}})}_{=:\xi(\bar{\mathsf{x}})} \to \varphi(\bar{\mathsf{x}})).$$

From definition of $\Gamma(\bar{\mathsf{c}})$, we know $T \models \forall \bar{\mathsf{x}}(\varphi \to \xi_i)$ for $i \in [m]^+$, so we have $T \models \forall \bar{\mathsf{x}}(\xi(\bar{\mathsf{x}}) \leftrightarrow \varphi(\bar{\mathsf{x}}))$ where all ξ_i are q.f. $\mathcal{L}$-formulas $\rightsquigarrow \xi$ is q.f. $\mathcal{L}$-formula. This ends the proof of Thm. 12.1. ■

Definition 12.1: Theory admitting quantifier elimination

Let T be $\mathcal{L}$-theory. We say T admits quantifier elimination (QE) (in language $\mathcal{L}$) if every $\mathcal{L}$-formula is equivalent in T to a q.f. $\mathcal{L}$-formula.

Above we found characterization of a single formula with n free variables being equivalent in T to a single q.f. formula with n free variables; it turns out to check a whole theory is QE, it suffices to look at formulas with just one free variable:

Lemma 12.1: Syntactic criterion (in fact characterization) for theory to have QE

Assume that for every q.f. formula φ and variable x (φ is truly arbitrary, can be in any number of variables, doesn't even have to mention x, though in that case $\exists \mathsf{x}\varphi$ trivially logically equivalent to the q.f. formula φ), there exists a q.f. formula ψ s.t. the formula $\exists \mathsf{x}\varphi$ is equivalent in T to ψ. Then T admits QE. Obviously converse also holds.

Proof: just induction on formula-height.

More explicitly, let ψ, ψ' be two formulas equivalent in T, denoted $\psi \sim_T \psi'$. Since (I proved this in Lemma in 220Ahw1p9) also $\neg\psi \sim_T \neg\psi', \exists \mathsf{x}\psi \sim_T \exists \mathsf{x}\psi'$ hence $\forall \mathsf{x}\psi \sim_T \forall \mathsf{x}\psi'$, and $\chi \wedge \psi \sim_T \chi \wedge \psi'$ (for any formula χ), we can argue by induction on $\mathrm{ht}(\psi)$: suffice to consider only formulas in prenex normal form $\varphi := \mathsf{Q}_1\mathsf{x}_1 \ldots \mathsf{Q}_n\mathsf{x}_n\chi$ (once we show those equiv. in T to q.f. formula, use transitivity of equivalence relation $\sim_T$), the formula $\mathsf{Q}_2\mathsf{x}_2 \ldots \mathsf{Q}_n\mathsf{x}_n\chi$ is smaller height so by induction hypothesis $\sim_T$ to q.f. formula ψ, so then $\varphi \sim_T \mathsf{Q}_1\mathsf{x}_1\psi$. By assumption of this lemma, there is q.f. formula $\sim_T$ to

$\mathsf{Q}_1\mathsf{x}_1\psi$ and hence to φ by transitivity. ■

This allows us to simplify the semantic criterion for theories with QE (I call it "q.f. Tarski-Vaught for QE"):

Corollary 12.1: Semantic criterion (in fact characterization) for theory to have QE

Let T be $\mathscr{L}$-theory, and assume [for any $\mathcal{M}, \mathcal{N} \models T$, substructure $\mathcal{A} \subseteq \mathcal{M}, \mathcal{N}$, q.f. $\mathscr{L}$-formula $\varphi(\mathsf{x}_0, \ldots, \mathsf{x}_n)$, and $\overline{a} \in A^n$; there existing $b_0 \in M$ s.t. $\mathcal{M} \models \varphi(b_0, \overline{a}) \implies$ there exists $c_0 \in N$ s.t. $\mathcal{N} \models \varphi(c_0, \overline{a})$]. Then, T has QE. Like with Thm. above, may be more natural to rephrase this with $\mathcal{A}$ common substructure *up to isomorphism* of $\mathcal{M}, \mathcal{N}$. Converse also holds, see *Remark*.

Proof: let $\chi := \exists \mathsf{x}_0 \varphi$. By assumption of this corollary, $\mathcal{M} \models \chi(\overline{a}) \iff \mathcal{N} \models \chi(\overline{a})$ for all $\overline{a} \in A^n$. This is for all models $\mathcal{M}, \mathcal{N} \models T$ and all common substructures $\mathcal{A} \subseteq \mathcal{M}, \mathcal{N}$, so Thm. 12.1 tells us χ is equivalent in T to some q.f. formula. This is for all q.f. φ and all variables x_0, so by the previous lemma (Lemma 12.1), T has QE. ■

Remark: converse also holds: by previous Thm. 12.1, condition (2) holds for any $\mathscr{L}$-formula, in particular $\exists \mathsf{x}_0 \varphi$. Note this criterion much easier to check than Thm. 12.1 condition (2), since we deal with only formulas with one existential quantifier and q.f. subformula, instead of arbitrary formula.

The hope is many "natural" structures in mathematics have QE (in minimal/natural language — "nonsense fact": always possible to extend language so theory admits QE in that language, so we must aim for "minimal/natural" languages), since then we can understand definability well.

11/4/22 we finished up this last Prop. and Cor.

Proposition 12.1: Characterizations of $\equiv$ and $\preceq$ for theories with QE

Let T be a theory with QE. Then for models $\mathcal{M}, \mathcal{N} \models T$,

(1) if there is a common substructure $\mathcal{A} \subseteq \mathcal{M}, \mathcal{N}$, then $\mathcal{M} \equiv \mathcal{N}$;

(2) substructure is equivalent to elementary substructure: $\mathcal{M} \subseteq \mathcal{N} \iff \mathcal{M} \preceq \mathcal{N}$.

Proof: (1) here is a special case of (1) $\implies$ (2) in Thm. 12.1; indeed, any sentence φ is equivalent in T to a q.f. formula $\psi(\mathsf{x})$, so for $\mathcal{A} \subseteq \mathcal{M}, \mathcal{N} \models T$ and any $a \in A$, we have $\mathcal{M}, \mathcal{N} \models \forall \mathsf{x}(\varphi \leftrightarrow \psi(\mathsf{x})) \implies$ for any $a \in A$,

$$\mathcal{M} \models \varphi \iff \mathcal{M} \models \psi(a) \iff \mathcal{A} \models \psi(a)$$

where we used that ψ q.f. means any extension/substructure preserves truth value (220Ahw1p7.1). The same is true for $\mathcal{N}$ in place of $\mathcal{M}$, so indeed

$$\mathcal{M} \models \varphi \iff \mathcal{M} \models \psi(a) \iff \mathcal{A} \models \psi(a) \iff \mathcal{N} \models \psi(a) \iff \mathcal{N} \models \varphi.$$

As for (2), it is a direct consequence of (converse of) "q.f. Tarski-Vaught for QE" Cor. 12.1 (taking $\mathcal{A}$ in the statement of Cor 12.1 to be $\mathcal{M}$), and the actual Tarski-Vaught test. ■

Corollary 12.2: For theories with QE, "minimal" substructure up to iso. implies complete

For T with QE, if there exists some "minimal" $\mathscr{L}$-structure $\mathcal{A}$ in the sense that for any $\mathcal{M} \models T$, $\mathcal{A} \subseteq \mathcal{M}$ (or more precisely there exists substructure $\mathcal{A}' \subseteq \mathcal{M}$ s.t. $\mathcal{A}' \simeq \mathcal{A}$), then T is complete theory.

Proof: by (1) in Prop. 12.1 above, and 220Ahw2p1 all models of T elementarily equivalent $\iff$ T complete. ■

Remark: I should emphasize that $\mathcal{A}$ in Prop. and Cor. above are ARBITRARY $\mathscr{L}$-structures; it doesn't have to model T! Thus for languages $\mathscr{L}$ admitting super simple $\mathscr{L}$-structures (e.g. languages with no function/constant symbols, like $\mathscr{L}_{\text{ord}}$), these propositions can be used with jaw-dropping effect, e.g. in 220Ahw3p2.2.

13 Lecture 13

11/2/22 Class 16

Our aim is to understand Th($\mathbb{C}$); in particular, we prove the celebrated theorem of Tarski-Chevalley: Th($\mathbb{C}$) admits QE (more generally for an alg. closed field). "Particularly important and nice example of theory from algebra". Recall first $\mathscr{L}_{\text{ring}} := \{\text{logical symbols}\} \cup \{\mathtt{0}, \mathtt{1}, \mathtt{-}, \mathtt{+}, \cdot\}$, and axioms of fields can be written as one $\mathscr{L}_{\text{ring}}$-sentence φ_{field} formed by conjuncting the [identity/inverse/commutativity/associative] axioms of addition and multiplication and the distributative law and $\neg\mathtt{0}=\mathtt{1}$. (For example, $\forall \mathtt{x}(\neg(\mathtt{x}=\mathtt{0}) \rightarrow (\mathtt{x}\cdot\mathtt{y}=\mathtt{1}))$.) We define property $P =$ "is a field" of $\mathscr{L}_{\text{ring}}$-strs. to be the property axiomatized by φ_{field}.

Definition 13.1: ACF

The theory of algebraically closed fields is the $\mathscr{L}_{\text{ring}}$-theory, denoted by ACF, consisiting of φ_{field} and, for each $n \in \mathbb{N}^+$, a sentence ρ_n ("rho" for "root") expressing that any polynomial of degree n has a root, i.e. $\rho_n := \forall \mathtt{z}_0 \ldots \mathtt{z}_{n-1} \exists \mathtt{x}(\mathtt{x}^n + \mathtt{z}_{n-1}\mathtt{x}^{n-1} + \ldots + \mathtt{z}_0 = \mathtt{0})$ (we are using abbreviation $\mathtt{x}^n := \underbrace{\mathtt{x} \cdot \ldots \cdot \mathtt{x}}_{n \text{ times}}$). (ACF is infinite axiomatization of property "is an algebraically closed field", a property of $\mathscr{L}_{\text{ring}}$-structures)

Example 13.1: $\mathbb{C} \models ACF$

By the fundamental theorem of algebra, $\mathbb{C}$ is a model of ACF.

To prove Tarski-Chevalley, we will use the following facts from abstract algebra:

Definition 13.2: Algebraic over, algebraic closure

- For a subring $A \subseteq$ field K, an element $a \in A$ called algebraic over A if it is root of some non-zero polynomial with coefficients in A (i.e. $p(a) = 0$ for some $p \in A[x]$).
- If A is integral domain, an algebraic closure of A is an algebraiclly closed field $K \supseteq A$ s.t. any element of K is algebraic over A.

Fact 13.1: Facts about alg. closed fields

Let A be an integral domain. Then,

(1) there exists an algebraic closure of A;

(2) if K, K' are two alg. closures of A, then exists (ring/field/$\mathscr{L}_{\text{ring}}$-structure)-isomorphism $\iota : K \xrightarrow{\cong} K'$ extending id_A, i.e. s.t. $\iota\restriction_A := \text{id}_A$;

(3) assuming that A is a subring of an alg. closed field L, then $A^{\text{alg}}_L := \{b \in L : b \text{ algebraic over } A\}$ is a field, and moreover an alg. closure of A (in particular is itself alg. closed);

(4) letting $\mathbb{F}_p^{\text{alg}}$ be algebraic closure of field $\mathbb{F}_p$ of p elements (p a prime number), then $\mathbb{F}_p^{\text{alg}}$ is increasing union of finite subfields F_N, $N \in \mathbb{N}^+$ — more precisely, for all $k \in \mathbb{N}^+$, the set of roots of the polynomial $x^{p^k} - x$ in $\mathbb{F}_p^{\text{alg}}$ is $=:$ subfield $\mathbb{F}_{p^k}$ (with p^k elements), and $\bigcup_{k \in \mathbb{N}^+} \mathbb{F}_{p^k} = \mathbb{F}_p^{\text{alg}}$, where moreover $\mathbb{F}_{p^k} \subseteq \mathbb{F}_{p^l} \iff k|l$; thus it suffices to take $F_N := \mathbb{F}_{p^{N!}}$;

(5) any algebraically closed field is infinite;

(6) and finally, letting $K \subseteq L$ be a field extension with K alg. closed field and $b \in L \setminus K$; then b is not algebraic over K (i.e. $K_L^{\text{alg}} = K$).

Theorem 13.1: Tarski-Chevalley

The theory ACF admits QE.

Proof: first note that a $\mathscr{L}_{\text{ring}}$-structure of a field is a subring (in order for restriction maps of interpretations of $\cdot, +, - \in \mathscr{L}_{\text{ring}}$ be well-defined, must be closed under those operations). By the semantic criterion for QE from Lec. 12, it is enough to prove

- **Goal:** if K, L are alg. closed fields (i.e. 2 models of theory ACF), and A is common subring of K, L (i.e. a common substructure), and $\varphi(\mathsf{x}_0, \ldots, \mathsf{x}_n)$ is q.f. $\mathscr{L}_{\text{ring}}$-formula, then for any $\overline{a} \in A^n$, [there existing $b \in L$ s.t. $L \models \varphi(b, \overline{a})$] $\implies$ there exists $c \in K$ s.t. $K \models \varphi(c, \overline{a})$.

  Note: because we quantify/range over all models K, L of ACF, the **Goal** is actually equivalent to putting " $\iff$ " in place of " $\implies$ " (by symmetry).

We can now make several reductions:

- **Claim1:** observe that if $A' \supseteq A$ is also common subring of K, L, it suffices to prove **Goal** for just A' since "$\forall \overline{a} \in A'$" $\rightsquigarrow$ "$\forall \overline{a} \in A$". The key idea is that by Fact (2) and (3), K and L both contain alg. closures of A, denoted F_K and F_L resp., that are isomorphic via some map ι that is identity on A, so up to identification of elements of K, L by ι, we can assume K, L contain a common alg. closed subfield $A' := F_L \simeq F_K$ (without identifying elements by ι, this is the statement $\overline{a} \in F_L$, prove $\exists b \in L$ s.t. $L \models \varphi(b, \overline{a})$ implies $\exists c \in K$ s.t. $K \models \varphi(c, \iota(\overline{a}))$). That is to say, we have reduced to proving **Goal1**: the original **Goal** in the special case of *alg. closed* $A \subseteq K, L$.
- **Claim2:** in fact, it suffices to prove **Goal1** ("alg. closed $A_0 \subseteq K, L$") in the super-special case "$A := K \subseteq L$" (this is now **Goal2**), since we can just apply the super-special case twice (once for "$A := A_0 \subseteq K$"; and once for "$A := A_0 \subseteq L$") to get: for all $\overline{a} \in A_0^n$, $\exists b \in L$ s.t. $L \models \varphi(b, \overline{a}) \iff \exists a_0 \in A_0$ s.t. $A_0 \models \varphi(a_0, \overline{a}) \iff \exists c \in K$ s.t. $K \models \varphi(c, \overline{a})$, where the two ends of this iff-chain being equivalent is exactly **Goal1** for the (abitrarily) chosen/given A_0.

  We have simplified the semantic parts of the goal, but we can also simplify the syntactic part of the goal:

- **Claim (general observation about q.f. formulas in any language):** any q.f. formula like φ is logically equiv. to a formula in disjunction normal form, i.e. a formula of the form $\bigvee_i \bigwedge_j \chi_{i,j}$ where each $\chi_{i,j}(\mathsf{x}_0, \ldots, \mathsf{x}_n)$ is either atomic or negated atomic (I also write this as $\bigvee_i \bigwedge_j \pm\chi_{i,j}$ for atomic $\chi_{i,j}$). In this form, φ is satsified iff at least one of its disjunctions is satisfied, so it suffices to prove **Goal2** in only the case that φ is a conjunction $\bigwedge$ of atomic or negated atomic formulas.

We are now done with all the reductions. In $\mathscr{L}_{\text{ring}}$, any atomic formula is equivalent to $p(\overline{\mathsf{x}})\texttt{=}0$ for some polynomial p in variables $\overline{\mathsf{x}}$ and with *integer* coefficients. Thus, we have boiled everything down to showing:

- **FinishLine:** given an alg. closed field $K \subseteq L$ and an $\mathscr{L}_{\text{ring}}$-formula of the form $\bigwedge_{i=1}^{n} P_i(\overline{\mathsf{x}})\texttt{=}0 \wedge \bigwedge_{j=1}^{m} \neg Q_j(\overline{\mathsf{x}})\texttt{=}0$ (with $P_i, Q_j \in \mathbb{Z}[\overline{\mathsf{x}}]$); THEN $\forall \overline{a} \in K^n$, if [$\exists b \in L$ s.t. on $(b, \overline{a})$, all P_i simultaneously vanish and all Q_j simultaneously do not vanish $\left\{\begin{smallmatrix}\text{all } P_i(b,\overline{a})=0\\ \text{all } Q_j(b,\overline{a})\neq 0\end{smallmatrix}\right]$, then $[\exists c \in K$ s.t. on $(c,\overline{a})$, $\left\{\begin{smallmatrix}\text{all } P_i(c,\overline{a})=0\\ \text{all } Q_j(c,\overline{a})\neq 0\end{smallmatrix}\right]$.

Indeed, if *some* $P_i(\mathsf{x}_0, a_1, \ldots, a_n) \in K[\mathsf{x}_0]$ is not the zero polynomial, everything is now super easy, because then for any $b \in L$ a root of the single variable polynomial $P_i(\mathsf{x}_0, \overline{a}) \in K[\mathsf{x}_0]$, b is algebraic over K, but K being algebraically closed tells us that actually b was already $\in K$ (see Fact (6)). I.e. as long as *some* $P_i(\mathsf{x}_0, \overline{a}) \in K[\mathsf{x}_0]$ is non-zero, then any $b \in L$ that satisfies the precondition $\left\{\begin{smallmatrix}\text{all } P_i(b,\overline{a})=0\\ \text{all } Q_j(b,\overline{a})\neq 0\end{smallmatrix}\right.$ automatically forces $b \in K$ and hence $c := b \in K$ satisfies the postcondition.

Thus the only remaining case is is $\varphi = \bigwedge_{j=1}^{m} \neg Q_i(\overline{\mathsf{x}}){=}0$. By the assumed existence of $b \in L$ s.t. all $Q_j(b, \overline{a}) \neq 0$ (the LHS precondition), we get that each polynomial $Q_i(\mathsf{x}_0, \overline{a}) \in K[\mathsf{x}_0]$ is non-zero, hence has only a finite number of roots (single variable polys have $\leqslant$ deg many roots). But Fact (5) says K is infinite, meaning we can find $c \in K$ that is not a root of any $Q_j(\mathsf{x}_0, \overline{a})$ for $j \in 1..m$. And then indeed $K \models \varphi(c, \overline{a})$. ■

11/4/22 Class 17

In particular, because we alreaady know the form of all q.f. formulas (up to logical equivalence),

Corollary 13.1: Definable sets in ACF

In $K \models \mathrm{ACF}$, the definable sets (with or without parameters) are precisely the the constructible sets, i.e. sets given by Boolean combinations of polynomial equations with coefficients in K.

Proof: by QE, all $\mathscr{L}_{\text{ring}}$-formulas are equivalent in ACF to q.f. formulas, which are logically equivalent to Boolean combinations of polynomial equations with coefficiens in K. ■

We have shown ACF has QE, but because of different characteristics, ACF is not complete. But we will show that specifying the characteristic, ACF_p becomes complete for all p prime or $p = 0$.

Definition 13.3: Theory ACF_p

For p a prime number $\in \mathbb{N}_+$, let ACF_p be the theory ACF union the sentence $\chi_p := \{\underbrace{\mathsf{1+\ldots+1}}_{p \text{ times}}\mathsf{=0}\}$. For $p = 0$, let ACF_0 denote $\mathrm{ACF} \cup \{\neg\chi_p : p \text{ prime number}\}$

Theorem 13.2: ACF_p complete

Let p be an prime ($\in \mathbb{N}$), or $p = 0$. Then ACF_p is complete.

Proof: ■

In fact, $\mathrm{Th}(\mathbb{C})$ is decidable — 220B concept: because axioms of $\mathrm{Th}(\mathbb{C})$ are recursive (i.e. just φ_{field}, sentences about the characteristic, and χ_n for all $n \in \mathbb{N}^+$), can enumerate all $\mathscr{L}$-words (by length and lexicographically), and can check if each is a formal proof of ϕ or $\neg\phi$, and if so, does it only use axioms of $\mathrm{Th}(\mathbb{C})$ (given ϕ_n in the formal proof, only finitely many axioms of $\mathrm{Th}(\mathbb{C})$ up to a given length, so because axioms recursive we can loop over all such finitely many to see if there's a match), so in principle given any $\mathscr{L}_{\text{ring}}$-sentence φ this algorithm will in finite time output either $\mathbb{C} \models \varphi$ or $\mathbb{C} \models \neg\varphi$.

Very important historically. Mark: if this proven after Gödel incompleteness, why expect $\mathbb{C}$ to be decidable? Artem: of course Gödelian phenomenon ZFC, PA, $\mathrm{Th}(\mathbb{Q}, +, \cdot)$ (famous theorem of Julia Robinson), "things are somehow as bad as possible". But somehow crossing threshold of "tameness", things are suddenly nicer: $\mathbb{C}, \mathbb{R}, \mathbb{Q}_p$ (all these "most tame"), then Presburger arithmetic, free groups (long and difficult proof, thousands of pages). "Model theory is the geography of tame mathematics" "Why should it be that such sharp dichotomy between tame/wild behavior exists? Not clear why nature is like this, but it somehow is." As particular example of wild theory mentioned above, $\mathrm{Th}(\mathbb{Q}, +, \cdot)$ can encode integers by Diophantine equations. There is algorithm for degree 2 polynomials (in arbitrary number of variables) having solutions in $\mathbb{Q}$; no such algorithm for degree 4; still open for degree 3 (entering into the realm of elliptic curves, number theory, Fermat last theorem).

Corollary 13.2: Lefschetz principle

Let φ be $\mathscr{L}_{\text{ring}}$-structure. Then, t.f.a.e.:

(1) $\mathbb{C} \models \varphi$;

(2) there exists an ACF char. 0 field in which φ is satisfied.

(3) any field $\models \mathrm{ACF}_0$ satisfies φ.

(4) there exists $N \in \mathbb{N}$ s.t. φ is satisfied in any field $K \models \mathrm{ACF}_p$ for $p > N$

(5) there exists infinitely many primes p s.t. there exist fields $K_p \models \mathrm{ACF}_p$ in which φ is satisfied.

Proof: (1) $\iff$ (2) $\iff$ (3) immediate from completeness of theory ACF_0 (...)

(3) $\implies$ (4):
(4) $\implies$ (5):
(5) $\implies$ (3): ■

14 Lecture 14

11/7/22 Class 18

We want to understand $\mathrm{Th}(\mathbb{Z}; 0, 1, +, -, <)$. Aim for today: give natural/"minimal" axioms for this theory to prove QE/understand definable subsets. Notice no multiplication, since Gödel incompleteness tells us not exist reasonable, i.e. recursive, axiomatization of complete theory of $Z \coloneqq (\mathbb{Z}; 0, 1, +, -, \cdot, <)$.

Preliminary exercise: for any q.f. formula $\varphi(\mathtt{x})$ in the language $\mathscr{L} \coloneqq \{\mathtt{0}, \mathtt{1}, \mathtt{+}, \mathtt{-}, \mathtt{<}\}$, there exists $N \in \mathbb{N}$ s.t. $Z \coloneqq (\mathbb{Z}; 0, 1, +, -, <) \models \varphi(n)$ for all $n > N$, or $Z \models \neg\varphi(n)$ for all $n > N$ (i.e. for every q.f. formula, the set of solutions in this structure is either final or cofinal). But the formula $\exists \mathtt{y}\ \mathtt{x}{=}\mathtt{y}{+}\mathtt{y}$, "x is divisible by 2" defines set $2\mathbb{Z}$, and similarly for numbers > 2 ("the set of integer multiples of 3", etc.), so the exercise shows that these formulas can not be equivalent to a q.f. formula in original language.

We talked about adding predicate for every formula in the theory, "adding new names for every definable set of the structure" and we gain QE, but that's stupid. We want to find some "minimal" structure. The famed theorem of Presburger (student of Tarski, proved one theorem and then quit mathematics) is that these "multiples of k" sentences are all the predicates one needs to add to the language. Towards this end, define $\mathscr{L}_{\text{Presb.}} \coloneqq \{\mathtt{0}, \mathtt{1}, \mathtt{+}, \mathtt{-}, \mathtt{<}, \mathtt{P}_1, \mathtt{P}_2, \ldots\}$ and expand structure to $\tilde{Z} \coloneqq (\mathbb{Z}; 0, 1, +, -, <, \mathbb{Z}, 2\mathbb{Z}, 3\mathbb{Z}, \ldots)$, i.e. $\mathtt{P}_i^{\tilde{Z}} \coloneqq i \cdot \mathbb{Z}$.

One can verify that that $\tilde{Z}$ satisfies the following $\mathscr{L}_{\text{Presb.}}$-sentences:

(1) axioms of additive abelian groups;

(2) axioms expressing that $\mathtt{<}$ is linear ordering;

(3) addition respects ordering, i.e. $\forall \mathtt{x}, \mathtt{y}, \mathtt{z}(\mathtt{x}{<}\mathtt{y} \to \mathtt{x}{+}\mathtt{z}{<}\mathtt{y}{+}\mathtt{z})$;

(4) discreteness of linear ordering, i.e. $\mathtt{0}{<}\mathtt{1} \wedge \neg\exists \mathtt{y}(\mathtt{0}{<}\mathtt{y}{<}\mathtt{1})$;

(5) division with remainder, i.e. for all fixed $n \in \mathbb{N}^+$, $\forall \mathtt{x} \exists \mathtt{y} \bigvee_{0 \leqslant r < n} \mathtt{x}{=}n\mathtt{y}{+}r\mathtt{1}$, which is abbreviation for $\mathscr{L}_{\text{Presb.}}$-sentence $\forall \mathtt{x} \exists \mathtt{y}(\mathtt{x} = \underbrace{\mathtt{y}{+}\ldots{+}\mathtt{y}}_{n \text{ times}}{+}\underbrace{\mathtt{1}{+}\ldots{+}\mathtt{1}}_{0 \text{ times}} \vee \ \ldots \ \vee \mathtt{x} = \underbrace{\mathtt{y}{+}\ldots{+}\mathtt{y}}_{n \text{ times}}{+}\underbrace{\mathtt{1}{+}\ldots{+}\mathtt{1}}_{n-1 \text{ times}})$

(6) predicate symbols express divisibility by n, i.e. for all fixed $n \in \mathbb{N}^+$, $\forall \mathtt{x}(\mathtt{P}_n\mathtt{x} \leftrightarrow \exists \mathtt{y}\ \mathtt{x}{=}\underbrace{\mathtt{y}{+}\ldots{+}\mathtt{y}}_{n \text{ times}})$

These $\mathscr{L}_{\text{Presb.}}$-sentences together are known as the Presburger arithmetic axioms, denotes PrA. "In practice, you just keep adding stuff and hope it captures everything. Now I'm just telling you the answers beforehand."

Theorem 14.1: Presburger arithmetic has QE and is complete

PrA has QE and is complete.

Proof: recall q.f. formulas logically equivalent to disjunctive normal form $\varphi \coloneqq \bigvee \bigwedge \pm\psi_i$ for atomic ψ_i, and reduction in Lec. 13 using "q.f. Tarski-Vaught for QE" $\equiv$ "semantic criterion for QE" says we only need to focus on $\varphi \coloneqq \bigwedge \pm\psi_i$ (i.e. can ignore outer disjunction), and main Thm. of Lec. 13 equates semantic and syntactic criterion for QE for any single formula, so it suffices to show $\exists \mathtt{y}\varphi(\bar{\mathtt{x}}, \mathtt{y})$ is PrA-equivalent to a q.f. formula for any $\varphi(\mathtt{x}_1, \ldots, \mathtt{x}_n, \mathtt{y}) = \bigwedge \pm\psi_i$ a conjunction of atomic and negated atomic formulas. Or purely syntactically, one can observe $\exists \mathtt{y}(\psi \vee \chi)$ logically equiv. to $\exists \mathtt{y}\psi \vee \exists \mathtt{y}\chi$, so we really can ignore $\bigvee$.

Note furthermore that atomic $\mathscr{L}_{\text{Presb.}}$-formulas have the form $t_1{=}t_2$, $t_1{<}t_2$, or $\mathtt{P}_N(t)$. The only function symbols we have are $\mathtt{+}, \mathtt{-}$, so using the additive abelian group axioms we can rewrite (rearrange so $\mathtt{y}$ isolated) the atomic formulas (in logically equiv. manner) to have the form $m\mathtt{y}{=}t(\bar{\mathtt{x}})$, $m\mathtt{y}{<}t(\bar{\mathtt{x}})$, or

$\mathtt{P}_N(m\mathtt{y}+t(\bar{\mathtt{x}}))$. Again, my is just an abbreviation for $\underbrace{\mathtt{y}+\ldots+\mathtt{y}}_{m\text{ times}}$.

- **Claim1:** we may assume that each conjunct of φ is of the following form ($m, N \in \mathbb{N}^+$, and $t(\bar{\mathtt{x}})$ some $\mathscr{L}_{\text{Presb.}}$-term): $m\mathtt{y}=t(\bar{\mathtt{x}})$, $m\mathtt{y}<t(\bar{\mathtt{x}})$, $m\mathtt{y}>t(\bar{\mathtt{x}})$, and $\mathtt{P}_N(m\mathtt{y}+t(\bar{\mathtt{x}}))$ (again all things that look like multiplication are just abbreviation for $\underbrace{\bullet+\ldots+\bullet}_{\#\text{ times}}$).

  <pf>: indeed every conjunct $\neg(m\mathtt{y}=t(\bar{\mathtt{x}}))$ can be replaced by (i.e. is logically equivalent to, denoted $\sim_\varnothing$) $(m\mathtt{y}<t(\bar{\mathtt{x}}) \vee m\mathtt{y}>t(\bar{\mathtt{x}}))$, and therefore

$$\exists \mathtt{y}\Big(\bigwedge(\ldots) \wedge (\varphi_1(\bar{\mathtt{x}},\mathtt{y}) \vee \varphi_2(\bar{\mathtt{x}},\mathtt{y}))\Big) \quad \sim_\varnothing \quad \exists \mathtt{y}\Big(\bigwedge(\ldots) \wedge \varphi_1(\bar{\mathtt{x}},\mathtt{y}))\Big) \vee \exists \mathtt{y}\Big(\bigwedge(\ldots) \wedge \varphi_2(\bar{\mathtt{x}},\mathtt{y})\Big).$$

  Recall we already reduced the case of multiple disjuncts to handling one disjunct at a time. Similarly, the negation $\neg\mathtt{P}_N(m\mathtt{y}+t(\bar{\mathtt{x}}))$ can be replaced by /is logically equiv. to

$$\mathtt{P}_N(m\mathtt{y}+t(\bar{\mathtt{x}})+1) \vee \ldots \vee \mathtt{P}_N(m\mathtt{y}+t(\bar{\mathtt{x}})+(n-1)\cdot 1),$$

  using the Presburger arithmetic axiom of division with remainder. </pf>

- **Claim2:** we may also assume that the variable $\mathtt{y}$ appears in every conjunct.
  <pf>: indeed, conjunctions in which $\mathtt{y}$ does not appear may be eliminated, because of the following univ. valid sentence:

$$\exists \mathtt{y}(\psi(\bar{\mathtt{x}}) \wedge \theta(\bar{\mathtt{x}},\mathtt{y})) \leftrightarrow \psi(\bar{\mathtt{x}}) \wedge \exists \mathtt{y}\theta(\bar{\mathtt{x}},\mathtt{y}).$$

  So finding q.f. PrA-equivalent formula to LHS (of above $\leftrightarrow$) is iff finding q.f. PrA-equivalent formula to $\exists \mathtt{y}\theta(\bar{\mathtt{x}},\mathtt{y})$ because $\psi(\bar{\mathtt{x}})$ is already q.f. (by starting assumption that $\varphi(\bar{\mathtt{x}},\mathtt{y}) := \psi(\bar{\mathtt{x}}) \wedge \theta(\bar{\mathtt{x}},\mathtt{y})$ is conjunction of atomic/negated atomic formulas, hence in particular q.f.). </pf>

- **Claim3:**
  <pf>: </pf>

Alright. Suppose first that $H \neq \varnothing$, and say $h' \in H$. Thus formula $\exists \mathtt{y}\varphi(\bar{\mathtt{x}},\mathtt{y})$ is logically equiv. to the formula

$$\exists \mathtt{y}\ m\mathtt{y}=t_{h'}(\mathtt{x}) \wedge$$

which is PrA-equivalent to

which is a q.f. $\mathscr{L}_{\text{Presb.}}$-formula.

$$\bigwedge_{h\in H} m\mathtt{y}=t_h(\bar{\mathtt{x}}) \wedge$$

11/9/22 Class 19

Recap: we said to show PrA has QE, it suffices to show that $\exists \mathtt{y}\varphi(\bar{\mathtt{x}},\mathtt{y})$ is PrA-equivalent to q.f. formula. "Minimal amount of work required". We have reduced it to $\varphi(\bar{\mathtt{x}},\mathtt{y})$ of the form

$$\bigwedge_{i\in I} t_i(\bar{\mathtt{x}})<m\mathtt{y} \wedge \bigwedge_{j\in J} m\mathtt{y}<t_j(\bar{\mathtt{x}}) \wedge \bigwedge_{k\in K} \mathtt{P}_{N(k)}(m\mathtt{y}+t_k(\bar{\mathtt{x}}))$$

for some finite index sets I, J, K.

For the next step, lets consider what happens in actual integers $Z \coloneqq (\mathbb{Z}; 0, 1, +, -, <)$ (instead of abstractly in arbitrary model $\mathcal{A}$). Fix $\overline{a} \in \mathbb{Z}^n$. We have a system of linear congruences, in variable $\mathtt{y}$: for fixed $k \in K$, $\mathtt{P}_{N(k)}(m\mathtt{y}\mathtt{+}t_k(\overline{a}))$ can be written in the more familiar notation $m\mathtt{y}\mathtt{+}t_k(\overline{a}) \equiv 0 \pmod{N(k)}$. The solutions (values of $\mathtt{y}$) in Z to this system (over all $k \in K$) of congruences form a union of congruence classes modulo $N \coloneqq \prod_{k\in K} N(k)$. "Really nothing" (nothing fancy like Chinese remainder theorem, which needs pairwise coprime anyways); this just says solutions in $\mathtt{y}$ of system can be checked by searching over congruence classes modulo N (fixed value, since $N(k)$ are fixed values). Also if $K = \varnothing$, take $N = 1$.

This suggests replacing existential statement in $\mathtt{y}$ by checking successively for $\mathtt{y}$ equal to $Nz, 1 + Nz, \ldots, (N-1) + Nz$. Now formally:

- **Claim:** the formula $\exists \mathtt{y} \varphi(\overline{\mathtt{x}}, \mathtt{y})$ (in its currently most reduced form) is PrA-equivalent to the formula $\theta(\overline{\mathtt{x}})$ defined as:

$$\theta(\overline{\mathtt{x}}) \coloneqq \bigvee_{r=0}^{N-1} \left(\bigwedge \mathtt{P}_{N(k)}(mr + t_k(\overline{\mathtt{x}})) \wedge \exists \mathtt{z} (\bigwedge_{i\in I} t_i(\overline{\mathtt{x}}) \mathtt{<} m(r + N\mathtt{z}) \wedge \bigwedge_{i\in I} t_j(\overline{\mathtt{x}}) \mathtt{>} m(r + N\mathtt{z})) \right).$$

Once we have this **Claim** we will have made further reduction of only considering φ with $K = \varnothing$.

`<pf>`: suppose $\mathcal{A} \coloneqq (A; \ldots) \models \mathrm{PrA}$, $\overline{a} = (a_1, \ldots, a_n) \in A^n$. We show: $\mathcal{A} \models \exists \mathtt{y} \varphi(\overline{a}, \mathtt{y}) \iff \mathcal{A} \models \theta(\overline{a})$, i.e. by definition, for all $\mathcal{A} \models \mathrm{PrA}$, $\mathcal{A} \models \forall \mathtt{x} (\mathtt{y} \varphi(\overline{a}, \mathtt{y}) \leftrightarrow \theta(\overline{a}))$.

($\implies$): let $b \in A$ s.t. $\mathcal{A} \models \varphi(\overline{a}, b)$. Division with remainder axiom in PrA tells us exists $c \in A$ and $r \in [N] \coloneqq \{0, \ldots, N-1\}$ s.t. $b = r + Nc \coloneqq \underbrace{1\mathtt{+} \ldots \mathtt{+}1}_{r \text{ times}} + \underbrace{c\mathtt{+} \ldots \mathtt{+}c}_{N \text{ times}}$. Then for every $k \in K$ (i.e. every congruence condition), we have

$$mb + t_k(\overline{a}) = m(r + Nc) + t_k(a) = mr + mNc + t_k(a),$$

which by assumption is divisible by $N(k)$, i.e. the expression is $\in N(k) \cdot \mathcal{A}$. Then, $\mathcal{A} \models \mathtt{P}_{N(k)}(mr + t_k(\overline{a}))$ (because mNc divisible by $N(k)$); and also $t_i(\overline{a}) < m(r + Nc)$ for every $i \in I$ and $m(r+Nc) < t_j(\overline{a})$ for every $j \in J$. So $\mathcal{A} \models \theta(\overline{a})$ (indeed our $r \in [N]$) with $\exists \mathtt{z}$ witnessed by our c.

($\impliedby$): assume now $\mathcal{A} \models \theta(\overline{a})$. So for some fixed $r \in [N]$ (may be more, but at least one), the rth disjunct of θ is satisfied. Let $c \in A$ s.t. $\exists \mathtt{z}(\ldots)$ holds. Take $b \coloneqq r + Nc$. Then (again using division with remainder axiom?) $\mathcal{A} \models \varphi(\overline{a}, b)$. `</pf>`

So we have finally reduced to the case that $\varphi(\overline{\mathtt{x}}, \mathtt{y})$ of the form

$$\bigwedge_{i\in I} t_i(\overline{\mathtt{x}}) \mathtt{<} m\mathtt{y} \wedge \bigwedge_{j\in J} m\mathtt{y} \mathtt{<} t_j(\overline{\mathtt{x}}),$$

for I, J finite index sets. If $I = \varnothing$ or $J = \varnothing$, we have $\mathrm{PrA} \models \exists \mathtt{y} \varphi(\overline{\mathtt{x}}, \mathtt{y})$: indeed if $I = \varnothing$, then for any $\mathcal{A} \models \mathrm{PrA}$ and $\overline{a} \in A^n$, $B = \{t_j(\overline{a}) : j \in J\} \subseteq A$ is finite, so we find $b \in A$ less than $\min B$ s.t. $mb \coloneqq \underbrace{b\mathtt{+} \ldots \mathtt{+}b}_{n \text{ times}} < B$ (there exist elements less than finitely many values). Similarly if $J = \varnothing$ can find $b > \max\{t_i(\overline{a}) : i \in I\}$ s.t. $mb >$ all such $t_i(\overline{a})$.

So we may assume $I, J \neq \varnothing$. Let $\mathcal{A} \models \mathrm{PrA}$. For each $\overline{a} \in A^n$, there is $i_0 \in I$ s.t. $t_i(\overline{a})$ is maximal among the the $t_i(\overline{a})$, $i \in I$. Similarly there is $j_0 \in J$ with $t_j(\overline{a})$ minimal among the $t_j(\overline{a})$, $j \in J$. Observe that $\exists \mathtt{y} \varphi(\overline{\mathtt{x}}, \mathtt{y})$ is the question of whether there is some multiple $m \cdot \mathtt{y}$ of m s.t. $t_{i_0}(\overline{a}) < m \cdot \mathtt{y} < t_{j_0}(\overline{a})$. And moreover every interval of $\geqslant m$ successive elements contains an element in $m \cdot A$ (from axioms of

PrA), so it suffices to search just m values $t_{i_0}(\overline{a})+1, \ldots, t_{i_0}(\overline{a})+m$.

To summarize/make rigorous, we have that $\exists y \varphi(\overline{x}, y)$ is equivalent in $\mathcal{A}$ to a disjunction over all pairs $(i_0, j_0) \in I \times J$ of a q.f. formula:

$$\bigvee_{(i_0,j_0)\in I\times J} \left(\bigwedge_{i\in I_0} t_i(\overline{x}) \leq t_{i_0}(\overline{x}) \wedge \bigwedge_{i\in I_0} t_j(\overline{x}) \geq t_{j_0}(\overline{x}) \wedge \bigvee_{r=1}^{m} (P_m(t_{i_0}(\overline{x})+r) \wedge (t_{i_0}(\overline{x})+r) < t_{j_0}(\overline{x})) \right),$$

and we are done. ■

Remark: this gives explicit algorithm for finding equivalent q.f. formulas (unlike abstract approach with ACF in Lec. 13), but "still exponential".

Corollary 14.1: PrA complete

PrA is complete

Proof: PrA has QE, and every model $\mathcal{A} \models$ PrA contains copy of $\tilde{Z}$ as substructure (generated by $1^{\mathcal{A}}$, so Cor. ??? applies ■

Remark: could in fact have continuum many copies of $\tilde{Z}$ or whatever, but somehow addition still well defined: take element of one copy, another element of another copy; their sum is well-defined, residing in perhaps yet another copy.

15 Lecture 15

16 Lecture 16

17 Lecture 17

18 Lecture 18

19 Lecture 19

20 Lecture 20

21 Lecture 21

22 Lecture 22

23 Lecture 23

24 Lecture 24

25 Lecture 25

26 Lecture 26

27 Lecture 27

28 Lecture 28