MATH 27700: MATHEMATICAL LOGIC 1 NOTES FALL 2020

BENEDICT MORRISSEY

Contents

1. Preface 1
2. Ordinal and Cardinal Arithmetic 1
3. Propositional Logic 2
3.1. Informal Propositional Logic 2
3.2. Sentences and Uniqueness of interpretation 3
3.3. Syntactic Proof 5
3.4. Semantic Truth/Proof 9
3.5. Completeness Theorem 10
3.6. Compactness Theorem 12
3.7. Boolean Algebras 13
3.8. Stone Duality 18
4. First Order Logic 18
4.1. First Order Theories 18
4.2. Models 21
4.3. Syntactic Category 23
4.4. Models as Functors 23
4.5. Ultraproducts and Compactness 23
4.6. Applications of Compactness 27
4.7. Syntactic Proof in First Order Logic 30
4.8. Henkin Models and Completeness 31
4.9. Types 34
4.10. Definability 36
4.11. Saturated Models 38
4.12. Quantifier Elimination 38
5. Appendix A: Some Alternatives 38
6. Appendix B: Category Theory 39
References 41

1. Preface

These are soome rough notes produced for the course Math/CMSC 27700: Mathematical Logic 1 at the University of Chicago in the Fall quarter of 2020. These are not polished notes, and it is reasonably likely that errors exist. Please inform me you find errors.

No results contained within are original.

2. Ordinal and Cardinal Arithmetic

Omitted. See e.g. [5].

Date: November 5, 2020.

3. Propositional Logic

Note: This is sometimes called Sentential Logic.

This is a simple form of logic. While it is limited in what it can talk about it is useful to cover first because it is a special case of and includes some of the machinery of the significantly more applicable first order logic (or second, third, ... order logic). Hence we can both use it to introduce some of the machinery, and prove analogous theorems to those about first order logic. The analogues for propositional logic are usually much easier to prove!

Idea: We have a set of symbols $\{A_i\}_{i \in I}$. Informally we should see A_i as referring to something (such as a statement) which can either be true or false.

We want to be able to assemble sentences which are finite strings consisting of these symbols, together with logical operators such as and ($\wedge$), or ($\vee$) and not ($\neg$).

- Goal 1: Say which strings are allowed. Check that each string has a unique interpretation.

We want to be able to complete reasoning of the form; If A_1, A_2 are true, then so is $A_1 \wedge A_2$.

- Goal 2: Describe a form of reasoning using sentences in propositional logic. This is a *syntactic notion of proof/truth*. We will denote this by $\Gamma \vdash \phi$, where under the assumptions Γ being a set of sentences we assume are true, and ϕ being a sentence we have proven is true using these assumptions.

There is another notion of truth, which we call the semantic notion of truth.

Namely if we have assigned each A_i as being true or false, this should also assign every sentence in the associated propositional logic as true or false. We will call a *model* of a set of sentences Γ to be a map t from sentences in the propositional logic to the set $\{0, 1\}$ (1 denotes true, while 0 denotes false), satisfying certain compatibility conditions (e.g. if $t(A_1) = t(A_2) = 1$ then $t(A_1 \wedge A_2) = 1$), and such that for $A \in \Gamma$ we have $t(A) = 1$. We then say that $\Gamma \vDash B$ if for all such models $t(B) = 1$. We call this the *semantic notion of truth.*

- Goal 3: Describe the relation between the syntactic notion of truth, and the semantic notion of truth. We will show
$$(\Gamma \vdash B) \Leftrightarrow (\Gamma \vDash B).$$
  This is the *Completeness Theorem* for propositional logic.

  We will also show a closely related result, the *Compactness Theorem* for first order logic which can be seen as an abstract version of the semantic part of the Completeness Theorem.

Finally we will describe the set (category) of theories (pairs of the set of symbols $\{A_i\}_{i \in I}$) together with a set of sentences Γ [that we can see as axioms – we are assuming these to be true]), in an algebraic fashion in terms of Boolean algebras. We can then describe models in terms of certain morphisms of Boolean algebras.

If there is time we will consider Stone duality from the viewpoint of propositional logic, where it forms a deep link between models of a theory and the theory itself.

3.1. **Informal Propositional Logic.**

We have a set of logical variables $A, B, C, ...$ each of which can be true (1) or false (0).

We have a set of logical connectives, whose truth values are defined as in table 1:

A	B	$A \wedge B$	$A \vee B$	$\neg A$	$A \rightarrow B$	$A \leftrightarrow B$	$\top$	$\bot$
1	0	0	1	0	0	0	1	0
1	1	1	1	0	1	1	1	0
0	0	0	0	1	1	1	1	0
0	1	0	1	1	1	0	1	0

Table 1. Definition of logical connectives

Remark 3.1. The equivalence of various logical connectives in the classical propositional logic we work in is not the case in various non-classical logics. One example of this is relevance logic which does not use that $\phi \to \psi$ is equivalent to $(\neg\phi) \vee \psi$, but requires some link between ϕ and ψ, in a similar way to the standard meaning of implies in non-mathematical language.

The reason for our definition of implies is primarily a matter of mathematical equivalence.

Definition 3.2. We call a formula a *tautology* if it is true for all values of the propositional variables.

Exercise 3.3. Show that the three formulas/sentences in definition 3.12 are tautologies.

3.2. **Sentences and Uniqueness of interpretation.**

Definition 3.4 (Language for Propositional Logic)**.** A Language $\mathcal{L}$ for propositional logic consists of

- A set $L = \{A_i\}_{i \in I}$. We call elements in this set *atomic formula* or *propositional variables.* This set must not contain the symbols appearing in the following bullet points, or the left or right parenthesis.
- A set of logical connectives[1], we will use $\leftrightarrow$, $\to$, $\vee$, $\wedge$, $\top$, $\perp$, $\neg$.
- Parentheses (,).

Definition 3.5 (Formulas/Sentences)**.** Consider the set of all finite strings of sequences of elements of L, logical connectives, and parentheses.

The set $Form(L)$ is the smallest subset of the above set, with the property that

- $A_i \in Form(L)$ for $A_i \in L$.
- $\top, \perp \in Form(L)$.
- If $A, B \in Form(L)$ then so are:

$$(A) \leftrightarrow (B), (A) \to (B), (A) \vee (B), (A) \wedge (B), \neg(A) \tag{3.1}$$

We call an element of the set $Form(L)$ a *formula* of $\mathcal{L}$ or a *sentence* of $\mathcal{L}$.

It is not clear that $Form(L)$ is well defined. This is because it is not clear that there is a set with the required properties that is a subset of all other sets with the required properties.

One approach to resolve this is to define a function LC from the power set of the set of finite strings of elements of L, logical connectives, and parentheses to itself.

We define $LC(S) := S \cup \{(A) \leftrightarrow (B), (A) \to (B), (A) \vee (B), (A) \wedge (B), \neg(A) | A, B \in S\}$.

It is clear that

$$\cup_{i=1}^{\infty} LC^i(L \cup \{\top, \perp\}) \tag{3.2}$$

is the desired smallest such set. One could also use this as a definition. This description will be useful in proofs.

An alternative definition of this smallest set is as the set of strings C such that there does not exist a set S satisfying the conditions of the definition with $C \notin S$.

Variant 3.6 (Language using fewer logical connectives)**.** *We could clearly use a subset of the logical connectives, and only introduce the atomic formulas, and those formed from formulas by the smaller set of logical connectives considered.*

It is clear that the unique reading lemma (lemma 3.8) will still hold for this variant.

Definition 3.7. We define a *theory in propositional logic* to be a pair (L, Γ) of a language L and a subset $\Gamma \subset Form(L)$.

Lemma 3.8 (Unique Reading Lemma propositional logic)**.** *Let $\phi \in Form(L)$, then ϕ has exactly one of the following forms (equation 3.3), and can be represented in that form in a unique way:*

$$(A) \leftrightarrow (B), (A) \to (B), (A) \vee (B), (A) \wedge (B), A, \top, \perp. \tag{3.3}$$

[1] We have some freedom to modify what these are.

We base our proof off that in [3]. We will prove this by induction on the length of the formula (number of symbols in the formula). For the base case where the formula contains a single symbol then it must by $A_i \in L$, or $\top$, or $\bot$, and the Lemma is true.

To prove the induction step we are going to use the following unique bracketing lemma.

Lemma 3.9 (Unique bracketing lemma)**.** *Let $c_+, c_i \subset \{1, ..., n\}$ be two disjoint subsets.*

There is at most one bijection $cl : c_+ \to c_-$ (which we will refer to as a bracketing) with the properties that:

- *We have $i \leq cl(i)$ for all $i \in c_+$.*
- *If $i, j \in c_+$ and $i < j < cl(i)$, then $i < j < cl(j) < cl(i)$.*

Suppose we have $\phi \in Form(L)$. Let n be the number of symbols in ϕ, and number these symbols left to right with the numbers 1 to n. Let c_+, c_- be the numbers corresponding to the open parentheses, and the closed parentheses respectively. Then:

Exercise 3.10. Show that there is a such a function cl given by mapping (the number of) a left parenthesis to the (number of the) associated right parenthesis.

Proof of lemma 3.9. We proceed by induction on n. If $n = 1$ then we must have $c_+ = c_- = \varnothing$ and the result follows.

Assume the result is true for all $n < k$, then for $n = k$:

If $1 \notin c_+$ we can "move everything down by one" and use the inductive hypothesis for $n = k - 1$. We can hence assume $1 \in c_+$.

Let $g : \{1, ..., n\} \to \mathbb{Z}$ be given by

$$g(c) = \begin{cases} 1 & \text{if } c \in c_+ \\ -1 & \text{if } c \in c_- \\ 0 & \text{else.} \end{cases}$$

Let $f(m) = \sum_{i=1}^{m} g(i)$.

We claim that if there is such a function cl, then $cl(1)$ is the smallest value of m (positive integer) such that $f(m) = 0$.

Note that if there is no such m then $|c_+| \neq |c_-|$ and as such there is no bijection.

Note that by the two conditions cl gives a bijection between $c_+ \cap \{1, , , cl(1)\}$ and $c_- \cap \{1, ..., cl(1)\}$. This means that we must have $f(cl(1)) = 0$.

Suppose that m is the positive integer such that $f(m) = 0$. Then by the first condition cl^{-1} must map $c_- \cap \{1, .., m\}$ injectively to $c_+ \cap \{1, .., m\}$. As $|c_- \cap \{1, .., m\}| = |c_+ \cap \{1, .., m\}|$ it must be a bijection. Hence $cl(1) \in \{1, ..., m\}$, and hence $cl(1) = m$.

Considering the setting where we are assigning left brackets and right brackets in a formula ϕ we can write $\phi = (\psi)\chi$ and apply the inductive hypothesis. We can in fact apply the inductive hypothesis by considering the bracketing restricting to $\{2, ..., m - 1\}$ and $\{m + 1, ..., n\}$. The result follows. ☐

We will now use lemma 3.9 to prove the unique reading lemma.

Proof of unique reading lemma (lemma 3.8). Let $\psi \in Form(L)$. By exercise 3.10 there is at least one way to pair the brackets in ψ. By the unique bracketing lemma this pairing of brackets is unique. The result follows. ☐

Remark 3.11. Having proven the unique reading lemma using careful and pedantic placing of brackets we will regress to using brackets when we feel they are needed as is done in standard human readable mathematics.

3.2.1. *Choice of Logical Connectives.* Note that the set of logical connectives we specified above is more than necessary. This is because we have the following equivalences

- $\neg\phi$ is equivalent to $\phi \to \perp$.
- $\phi \vee \psi$ is equivalent to $(\neg\phi) \to \psi$. This is in turn equivalent to $(\phi \to \perp) \to \psi$
- $\top$ is equivalent to $\neg\perp$, which by the above is equivalent to $\perp \to \perp$.
- $\phi \wedge \psi$ is equivalent to $\neg(\neg\phi \vee \neg\psi)$. Using the above we can then write this as a (long) sentence involving only the logical connectives $\to$ and $\perp$. However we will slightly simplify things by seeing it as equivalent to $\neg(\phi \to \neg\psi)$.
- $\phi \leftrightarrow \psi$ is equivalent to $(\phi \to \psi) \wedge (\psi \to \phi)$.

There are several other choices of logical connectives we could make. Adding logical connectives makes working inside propositional logic easier. Using a slimmer set of logical connectives makes proving things about propositional logic easy.

3.3. **Syntactic Proof.** There are many choices for what we mean by a proof. Choices include natural deduction, and sequent calculus. We will use a version of Hilbert calculus. The version we use is easy to prove things about, but difficult to prove things with.

We first introduce some axioms, that we will use in writing out proofs.

We first use section 3.2.1 to work in propositional logic using only the logical symbols $\to$ and $\perp$.

If we write propositional logic using only the logical connectives $\to$ and $\perp$ we can then write the axioms of propositional logic as:

Definition 3.12 (Axioms of Propositional Logic, using only the logical connectives $\to, \perp$)**.** We will refer to the following sentences as axioms of propositional logic:

- $\phi \to (\psi \to \phi)$
- $((\phi \to \perp) \to \perp) \to \phi$.
- $(\phi \to (\psi \to \chi)) \to ((\phi \to \psi) \to (\phi \to \chi))$

for any sentences ϕ, ψ, χ.

The idea is that we will be able to use these axioms together with deduction rules to deduce all tautologies/statements that are always true for any truth valuations.

You can check using truth tables that these are statements that are true for any truth valuations. It is not a. priori clear that we might not require additional axioms.

Remark 3.13. There are many other choices of axioms we could have made, that would give us the same result.

To include additional logical symbols we would need to add axioms describing their relationship to $\to$ and $\perp$.

It is not entirely clear what the axioms we would need to add are. For example to add the logical symbol $\neg$ we could add the axioms:

- $(\neg\phi) \to (\phi \to \perp)$
- $(\phi \to \perp) \to (\neg\phi)$.

It is however not immediately clear whether or not these axioms would be sufficient.

We could instead modify the below notion of Syntactic proof, by allowing us to replace any logical sentence by the version that only involves $\perp, \to$. Or we could add the axioms that for any sentence ϕ, state that ϕ is equivalent to a sentence (use the recursive definition of formulas to define a way to recursively replace all logical connectives in a formula with the connectives $\to$, $\perp$ to gain an equivalent formula) only using the logical connectives $\to$, $\perp$. Or we could just interpret any formula using other logical connectives as the equivalent formula which only uses $\to$, $\perp$.

Let Γ be a set of sentences of our language, and let ϕ be a sentence in our language.

Definition 3.14 (Syntactic Proof in propositional Logic)**.** In propositional logic a syntactic proof $\Gamma \vdash \phi$ is a finite sequence of sentences $S_1, ..., S_n$ ($n \in \mathbb{N}$) of our language, with the properties that $S_n = \phi$, and that for each $1 \leq k \leq n$ either:

- $S_k \in \Gamma$.
- S_k is one of the axioms listed in definition 3.12.
- $\exists\ i, j < k$ with the properties that S_j is $S_i \to S_k$. (Modus Ponens)

Remark 3.15. It is clear that we should consider such a series of sentences a proof. It is perhaps less clear that we should *only* consider such a series a proof. We might ask if we should have included additional axioms, and additional deduction methods. Essentially the completeness theorem for propositional logic tells us we can get all semantic consequences the above definition is sufficient.

This definition is useful in that it is relatively easy to reason about what we can prove, it is not so good if one actually wants to prove something! In this case one could use other definitions of proof, such as natural deduction.

Example 3.16. Suppose we want to prove the statement $(\neg A) \vee A$. We realize this is (in disguised form) the axiom $((A \to \perp) \to \perp) \to A$.

If on the other hand we start with $A \vee (\neg A)$ we can instead have to prove

$$B \to B$$

for $B = (A \to \perp)$.

A proof follows

$$B \to ((B \to B) \to B)$$
$$(B \to ((B \to B) \to B)) \to ((B \to (B \to B)) \to (B \to B))$$
$$(B \to (B \to B)) \to (B \to B) \text{ (Modus Ponens)}$$
$$B \to (B \to B)$$
$$B \to B \text{ (Modus Ponens)}$$

Exercise 3.17. Find a proof of the statement $A \to (\perp \to \perp)$.

To prove the completeness theorem we are going to give some properties of the set $\{\phi | \Gamma \vdash \phi\}$.

We firstly show the *deduction* theorem:

Theorem 3.18 (Deduction Theorem)**.** *We have $\Gamma, \phi \vdash \psi$ if and only if $\Gamma \vdash \phi \to \psi$.*

Proof. The if direction is clear. We will hence concentrate on the only if.

By definition of a proof, we only need (see remark 3.19) to show this for:

- ψ an axiom or an element of $\Gamma \cup \{\phi\}$.
- ψ is such that we have already show $\Gamma \vdash (\phi \to \chi)$, and $\Gamma \vdash (\phi \to (\chi \to \psi))$.

In the first case if ψ is an axiom or in Γ a proof $\Gamma \vdash (\phi \to \psi)$ is given by

$$\psi \to (\phi \to \psi)$$
$$\psi$$
$$(\phi \to \psi) \text{ (Modus Ponens).}$$

In the first case if $\psi = \phi$, then we have example 3.16 gives a proof $\Gamma \vdash \psi \to \psi$.

In the second case a proof is given by concatenating the proofs of $\phi \to \chi$ and $\phi \to (\chi \to \psi)$ with

$$\phi \to \chi$$
$$\phi \to (\chi \to \psi)$$
$$(\phi \to (\chi \to \psi)) \to ((\phi \to \chi) \to (\phi \to \psi))$$
$$(\phi \to \chi) \to (\phi \to \psi)$$
$$\phi \to \psi.$$

The result follows.

□

Remark 3.19. The purpose of this remark is to explain the reduction we made. Let $MP : \mathcal{P}(Form(L)) \to \mathcal{P}(Form(L))$ be the operation that takes a set S to the set S union with all $\phi \in Form(L)$ such that there exists $\chi \in Form(L)$ such that $\chi, \chi \to \phi \in S$.

Then by the definition of proof we have that

$$\{\psi | \Gamma \vdash \psi\} = \cup_{n \in \mathbb{N}} MP^n(\Gamma \cup \{\text{Axioms}\}).$$

Let $Pd(S) := \{\psi | S \vdash \phi\}$ for any set $S \subset Form(L)$. Let $Pd_{\phi\to}(S) := \{\psi | S \vdash \phi \to \psi\}$.

We are showing that $Pd(\Gamma \cup \{\phi\}) \subset Pd_{\phi\to}(\Gamma)$. We are doing this by inductively (on n) arguing that

$$MP^n(\Gamma \cup \{\phi\} \cup \{\text{axioms}\}) \subset Pd_{\phi\to}(\Gamma).$$

Example 3.20 (Explosion). We have that $\bot \to \phi$ for any sentence ϕ.

By the first axiom we have $\vdash (\bot \to (\phi \to \bot) \to \bot)$. We hence have that

$$\bot \vdash (\phi \to \bot) \to \bot.$$

The second axiom gives that $\bot \vdash \phi$.

Theorem 3.21. *The following hold true:*

(1) *For any Γ, and any $\phi \in Form(L)$, we have $\Gamma \vdash (\neg\phi) \vee \phi$.*
(2) *If $\Gamma_1 \subset \Gamma_2 \subset \{\phi | \Gamma \vdash \phi\}$, then $\Gamma_1 \vdash \phi$ if and only if $\Gamma_2 \vdash \phi$.*
(3) *If $\Gamma \cup \{\phi\} \vdash \chi$, and $\Gamma \cup \{\psi\} \vdash \chi$, then $\Gamma \cup \{\phi \vee \psi\} \vdash \chi$.*
(4) *$\Gamma \vdash \top$.*
(5) *$\neg(\phi \to \psi) \vdash \neg\psi, \phi$.*
(6) *$(\neg\phi) \wedge \phi \vdash \bot$.*
(7) *$\{\phi, \phi \to \psi\} \vdash \psi$.*
(8) *If $\Gamma_1 \subset \Gamma_2$ then if $\Gamma_1 \vdash \phi$ then $\Gamma_2 \vdash \phi$.*
(9) *If $\{\Gamma_i\}_{i \in I}$ is well ordered by inclusion then if $\cup_{i \in I}\Gamma_i \vdash \phi$, then there exists i such that $\Gamma_i \vdash \phi$.*

Proof. (1) In example 3.16 we showed that $\vdash (\neg\phi) \vee \phi$. This first observation under point (2) shows that we hence have $\Gamma \vdash (\neg\phi) \vee \phi$.

(2) Firstly suppose that $(S_1, ..., S_n)$ is a proof $\Gamma_1 \vdash \phi$. Then by definition it is also a proof $\Gamma_2 \vdash \phi$.

Secondly suppose that $(S_1, ...S_m)$ is a proof $\Gamma_2 \vdash \phi$. Let $S_{j_1}, ..., S_{j_n} \in \Gamma_2 \backslash \Gamma_1$. We then have that for each $1 \leq i \leq n$ there is a proof $(C_{i,1}), ..., C_{i,l_i})$ of S_{j_i}. Then

$$(C_{1,1}, ..., C_{1,l(1)-1}, C_{2,1}, ..., C_{2,l(2)-1}, ..., C_{n,l(n)-1}, S_1, ..., S_m)$$

is a proof $\Gamma_1 \vdash \phi$.

(3) We firstly reduce this to the case where ψ is $\neg\phi$.

Suppose that this is true for ϕ and $\neg\phi$. Assume (using the deduction theorem) $\Gamma \vdash \phi \to \chi, \psi \to \chi$. Let $\Gamma' := \Gamma \cup \{(\neg\phi) \to \chi\}$. Note that $\Gamma', \phi \vdash \chi$ and $\Gamma', \neg\phi \vdash \chi$, because a proof is given by (appending proofs of the necessary statements to the beginning of) $\neg\phi, \neg\phi \to \psi, \psi, \psi \to \chi, \chi$.

Hence by part (1) of this theorem $\Gamma, \phi \vee \psi \vdash \chi$. Hence it is enough to show that if $\Gamma, \phi \vdash \chi$, and $\Gamma, \neg\phi \vdash \chi$, then $\Gamma \vdash \chi$. Equivalently it is enough to show there is a proof of:

$$(\phi \to \chi) \to (((\neg\phi) \to \chi) \to \chi).$$

Lemma (a): If $A \to B$, $B \to C$ we can prove $A \to C$.

Proof: By the deduction theorem it is enough to show that $A, A \to B, B \to C \vdash C$, which is clear.

Lemma (b): $(\phi \to \chi) \to (((\neg\chi) \to \phi) \to \chi)$.

By the deduction theorem it is enough to show that

$$\phi \to \chi, (\neg\chi) \to \phi \vdash \chi.$$

By lemma (a) and the deduction theory it is thus sufficient to show that $(\neg\chi \to \chi) \to \chi$.

An outline of a proof is given by

$\neg\chi \to (\neg\neg(\delta \to \delta)) \to \neg\chi)$ (axiom)

$(\neg\neg(\delta \to \delta)) \to \neg\chi) \to (\chi \to \neg(\delta \to \delta))$ $\mathrm{HW4} + \mathrm{DeductionTheorem}.$

$\neg\chi \to (\chi \to \neg(\delta \to \delta))$ Lemma(a)

$(\neg\chi \to (\chi \to \neg(\delta \to \delta))) \to ((\neg\chi \to \chi) \to (\neg\chi \to \neg(\delta \to \delta)))$ Axiom

$((\neg\chi \to \chi) \to (\neg\chi \to \neg(\delta \to \delta)))$ MP

$(\neg\chi \to \neg(\delta \to \delta)) \to ((\delta \to \delta) \to \chi)$ HW4

$(\neg\chi \to \chi) \to ((\delta \to \delta) \to \chi)$

We can now conclude the proof by using the deduction theorem and noting example 3.16.

Lemma (c) We have that

$$(((\neg\chi) \to \phi) \to \chi) \to (((\neg\phi \to \chi)) \to \chi).$$

From HW4 and the deduction theorem we have that$((\neg\chi) \to \phi) \leftrightarrow (\neg\phi) \to \chi)$. Hence we can use this and the deduction theorem to prove this.

By applying lemma (a) to Lemma (b) and lemma (c) we get the desired result.

(4) We set $\top = (\bot \to \bot)$. Specializing the proof of $B \to B$ in example 3.16 to $B = \bot$ shows this is true.
(5) Firstly we note that unpacking this sentence we get $(\phi \to \psi) \to \bot$. We first show that $\psi, ((\phi \to \psi) \to \bot) \vdash \bot$ showing by the deduction theorem that $((\phi \to \psi) \to \bot) \vdash (\psi \to \bot)$.

A proof is given by

$\psi \to (\phi \to \psi)$

ψ

$(\phi \to \psi)$ (Modus Ponens)

$(\phi \to \psi) \to \bot$

$\bot$ (Modus Ponens).

For ϕ we will first show that $((\phi \to \psi) \to \bot), \phi \to \bot \vdash \phi \to \psi$. By the deduction theorem it is equivalent to show that $((\phi \to \psi) \to \bot), \phi \to \bot, \phi \vdash \psi$. A "proof" is given by

$$\phi$$
$$\phi \to \bot$$
$$\bot$$
$$\bot \to \psi$$
$$\psi$$

where we are using $\bot \vdash \psi$ from example 3.20 (note: that we are abusing the notation of a proof slightly).

We hence have by modus ponens that $((\phi \to \psi) \to \bot), \phi \to \bot \vdash \bot$. By the deduction theorem we hence have that

$$((\phi \to \psi) \to \bot) \vdash ((\phi \to \bot) \to \bot).$$

Using modus ponens and the second axiom we hence have that

$$((\phi \to \psi) \to \bot) \vdash \phi.$$

(6) Unpacking $(\neg\phi) \wedge \phi$ we realize this is the sentence $(\phi \to \phi) \to \bot$. We use example 3.16 to show $(\neg\phi) \to (\neg\phi)$ and then apply modus ponens to derive $\bot$.
(7) A proof is given by the sequence $(\phi, \phi \to \psi, \psi)$.
(8) This was in fact proved under point (2).
(9) Because of the finite length of a proof, any proof must only contain finitely many sentences which are in $\cup_{i\in I}\Gamma_i$. Hence there must be some Γ_i which contains all these sentences. Hence any given proof of ϕ is in fact a proof $\Gamma_i \vdash \phi$ for some $i \in I$.

□

3.4. **Semantic Truth/Proof.** The idea behind propositional logic is that each symbol $P_i \in L$ can either be true or False.

Consider such a map of assignments

$$\mathcal{P} \xrightarrow{t} \{\bot, \top\} \cong \{0, 1\}.$$

Construction 3.22. There exists a unique extension of t as above to a as above extends to a function:

$$Form(L) \xrightarrow{t} \{0, 1\}.$$

via the recursive definition:

- $t(\neg\phi) = 1 - t(\phi)$.
- $t(\phi \vee \psi) = \max(t(\phi), t(\psi))$.
- $t(\phi \wedge \psi) = \min(t(\phi), t(\psi))$.
- $t(\bot) = 0$.
- $t(\top) = 1$.
- $t(\phi \to \psi) = \max(1 - t(\phi), t(\psi))$.

Note that we are using the unique reading lemma.

Definition 3.23 (Truth function)**.** We call a function $t : Form(L) \to \{1, 0\}$ a truth function if

- $t(\neg\phi) = 1 - t(\phi)$.
- $t(\phi \vee \psi) = \max(t(\phi), t(\psi))$.
- $t(\phi \wedge \psi) = \min(t(\phi), t(\psi))$.
- $t(\bot) = 0$.
- $t(\top) = 1$.

- $t(\phi \to \psi) = \max(1 - t(\phi), t(\psi))$.

Proposition 3.24. *Any truth function $t : Form(L) \to \{0,1\}$ is constructed from $t|_L : L \to \{0,1\}$ by construction 3.22.*

Proof. Clear. □

Definition 3.25 (Model)**.** Let Γ be a set of sentences of $Form(L)$. We call a truth function t a *model* of Γ, if $t(\phi) = 1$ for all $\phi \in \Gamma$.

Definition 3.26 (Semantic notion of Truth)**.** We write $\Gamma \vDash \phi$ if for all models t of Γ, $t(\phi) = 1$.

Warning 3.27. *People sometimes also write $t \vDash \phi$, if $t(\phi) = 1$.*

This is the *semantic* version of the notion that "The sentences of Γ imply the statement ϕ."

Theorem 3.28 (Soundness of propositional logic)**.** *If $\Gamma \vdash \phi$ then $\Gamma \vDash \phi$.*

Proof. Let $\Gamma \vdash \phi$, and let $t : Form(L) \to \{0,1\}$ be an arbitrary truth function.

Let $(S_k)_{k=1}^n$ be a proof of ϕ. We prove by induction that $t(S_k)$ is true for all $1 \le k \le n$. Recall that there were three possibilities for S_k, we could have any of:

- $S_k \in \Gamma$, in which case $t(S_k) = 1$ by definition of a model.
- S_k is an axiom, in which case $t(S_k) = 1$ (e.g. by checking truth tables [exercise]).
- $\exists\, i, j < k$ with the properties that S_j is $S_i \to S_k$. Hence by the induction hypothesis $t(S_i) = 1$, and
$$1 = t(S_i \to S_k) = \max\{1 - t(S_i), S_k\} = \max\{0, t(S_k)\},$$
hence $t(S_k) = 1$.

Hence $t(\phi) = t(S_n) = 1$.

As t was an arbitrary model, we hence have $\Gamma \vDash \phi$. □

Exercise 3.29. At first glance it looks like we did not include a base case in the induction proof inside Theorem 3.28. When did we cover the base case?

3.5. **Completeness Theorem.** An abstract interpretation of a proof system is as a map $\mathcal{P}(Form(L)) \xrightarrow{Pd} \mathcal{P}(Form(L))$. This map is given by
$$\mathcal{P}(Form(L)) \ni \Gamma \mapsto \{\phi | \phi \in Form(L), \Gamma \vdash \phi\}.$$

We will say that we can derive a contradiction from Γ if $\perp \in Pd(\Gamma)$.

The function Pd has the following properties, and these properties will be sufficient to provide a proof of the completeness Theorem for propositional logic.

Theorem 3.30. *The function Pd has the following properties:*

1. *For any Γ, and any $\phi \in Form(L)$, we have $(\neg\phi) \vee \phi \in Pd(\Gamma)$.*
2. *If $\Gamma \subset S \subset Pd(\Gamma)$, then $Pd(S) = Pd(\Gamma)$.*
3. *$Pd(\Gamma \cup \{\phi \vee \psi\}) \supset Pd(\Gamma \cup \phi) \cap Pd(\Gamma \cup \psi)$ (i.e. if $\Gamma \cup \{\phi\} \vdash \chi$, and $\Gamma \cup \{\psi\} \vdash \chi$, then $\Gamma \cup \{\phi \vee \psi\} \vdash \chi$).*
4. *$\top \in Pd(\Gamma)$ (i.e. $\Gamma \vdash \top$)*
5. *$Pd(\neg(\phi \to \psi)) \ni \neg\psi, \phi$. (i.e. $\neg(\phi \to \psi) \vdash \neg\psi, \phi$).*
6. *$\perp \in Pd(\phi, \neg\phi)$. (i.e. $(\neg\phi) \vee \phi \vdash \perp$).*
7. *$Pd(\phi, \phi \to \psi) \ni \psi$. (i.e. $\phi, \phi \to \psi \vdash \psi$.)*
8. *If $\Gamma_1 \subset \Gamma_2$ then $Pd(\Gamma_1) \subset Pd(\Gamma_2)$.*
9. *If $\{\Gamma_i\}_{i \in I}$ is well ordered by inclusion then we have that $Pd(\cup_{i \in I} \Gamma_i) = \cup_{i \in I} Pd(\Gamma_i)$. (Equivalently if $\cup_{i \in I} \Gamma_i \vdash \phi$, then there exists i such that $\Gamma_i \vdash \phi$).*

This is just a rewrite of Theorem 3.21.

Let us first consider the existence of models for propositional logic. In particular we will prove proposition 3.32. We first introduce the following definition:

Definition 3.31 (Consistency)**.** We say that a set $\Gamma \subset Form(L)$ is not consistent if $\Gamma \vdash \perp$.

Otherwise we call Γ consistent.

Proposition 3.32 (Lindenbaum's Lemma)**.** *A set Γ is consistent if and only if it has a model.*

Remark 3.33. In fact we will prove that for any function $Pd : \mathcal{P}(Form(L)) \to \mathcal{P}(Form(L))$ satisfying the properties of Theorem 3.30, if Γ is such that $Pd(\Gamma) \ni \perp$, then Γ has a model.

One direction is straightforward: If Γ is not consistent, then $\Gamma \vdash \perp$. By soundness this implies that $t(\perp) = 1$ for any model $t : Form(L) \to \{0, 1\}$. However by definition there is no such model.

We are now left to prove the harder direction: If Γ is consistent we need to construct/find/prove the existence of a model for Γ. Recall that specifying a model of t is equivalent to finding the set $\{\phi \in Form(L) | t(\phi) = 1\}$ or indeed to finding the $\{A_i \in L | t(A_i) = 1\}$. In the finite case we can try to go through the A_i (or the ϕ) and either add ϕ or $\neg\phi$ to Γ (having chosen an ordering on the ϕ).

We start by considering the case where L is finite.

If Γ does not give us a contradiction, then we need that at least of $\Gamma \cup \{\phi\}$ or $\Gamma \cup \{\neg\phi\}$ does not gives us a contradiction. This follows from properties (1-3) of Theorem 3.30. We choose one of the subset of $\{\phi, \neg\phi\}$ that does not give us a contradiction and add this to Γ.

Let $\Gamma' \supset \Gamma$ be the set we create after applying this process to all sentences of L.

Claim: Define $t : Form(L) \to \{0, 1\}$ by

$$t(\phi) = \begin{cases} 1, & \text{if } \phi \in \Gamma' \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

Then t is a model of Γ in the sense of Definition 3.25.

Proof: We need to check the conditions in the definition of a truth function:

- $t(\neg\phi) = 1 - t(\phi)$. This follows because exactly one (Exercise: why exactly one, use property 6 of Theorem 3.30) of ϕ and $\neg\phi$ is in Γ'.
- $t(\phi \vee \psi) = \max(t(\phi), t(\psi))$. Using subsection 3.2.1 this is equivalent to the last point in this list.
- $t(\phi \wedge \psi) = \min(t(\phi), t(\psi))$. Using subsection 3.2.1 this is equivalent to the last point in this list.
- $t(\perp) = 0$. By construction Γ' is consistent, so this follows by the definition of consistency.
- $t(\top) = 1$. This follows by property 4 of Theorem 3.30.
- $t(\phi \to \psi) = \max(1 - t(\phi), t(\psi))$. Note that exactly one of each of the sets $\{(\phi \to \psi), \neg(\phi \to \psi)\}$, $\{\phi, \neg\phi\}$ and $\{\psi, \neg\psi\}$ is in Γ'. If $\neg(\phi \to \psi) \in \Gamma'$ then by property 5 of Theorem 3.30 we have $\phi, \neg\psi \in \Gamma'$, and this relation is satisfied. If $\phi \to \psi, \phi \in \Gamma'$ then $\psi \in \Gamma'$ by property 7 of Theorem 3.30, and this relation is satisfied. If $\phi \to \psi, \neg\phi \in \Gamma'$ then this relation is satisfied regardless of which of $\{\psi, \neg\psi\} \in \Gamma'$.

This construction works in the case L (and hence $Form(L)$) is finite. In the case where L is infinite we have to be more clever. One way that we can describe t is by $t^{-1}(1)$, which we can describe as a maximal consistent subset of $Form(L)$. We can use this description to define a model t in the general case

Recall:

Let $(P, \leq_P)$ be a partially ordered set. We call a totally ordered subset of P a *chain*. We call an element $x \in P$ a *maximal element* if there is no $y \in P$ with $y \geq_P x$, and $y \neq x$.

Lemma 3.34 (Zorn's Lemma)**.** *Let P be a partially ordered set such that every chain has an upper bound in P (ie. for a chain $C \subset P$, $\exists x \in P$ such that $x \geq y$, $\forall y \in C$). Then the set of maximal elements of P is nonempty.*

Remark 3.35. Note that P is non-empty by considering the empty set as a chain.

Proof. Assume that there is no maximal element in P. We use the axiom of choice to define a function g from well ordered subsets of P to P, with the property that for $S \subset P$ well ordered, $g(S) \notin S$, and $g(S) > s$ for all $s \in S$ (we here use the existence of an upper bound, and the assumption that there is no maximal element).

We say a set S satisfies property g if for all $s \in S$ $s = g(\{s' \in S | s' \leq s\})$.

Let S_1 and S_2 satisfy property g then either[2] $S_1 \subset S_2$ or $S_2 \subset S_1$. (This should remind you of ordinals, and the proof will indeed proceed similarly). Consider the maximal subset[3] S_3 which is an initial sequence[4] of S_1 and S_2. Then as S_1 and S_2 satisfy property g, we have either that $g(S_3) \in S_i$ $(i = 1, 2)$ or $S_i = S_3$. The result follows by the assumption of maximality of S_3.

Let G be the union of all sets satisfying property g. Then G satisfies property g . Furthermore $G \cup \{g(G)\}$ satisfies property g. Contradiction. The result follows. □

Exercise 3.36. Use Zorn's lemma to show that every (not necessarily finite dimensional) vector space has a basis.

Proof of Proposition 3.32. Consider the set P of consistent sets of formula of our language, containing Γ.

We claim that this satisfies the hypotheses of Zorn's lemma if Γ is consistent (Exercise).

There is hence a[5] maximal such set S. We claim that t defined by

$$t(\phi) = \begin{cases} 1, & \text{if } \phi \in S \\ 0, & \text{otherwise.} \end{cases} \tag{3.5}$$

is a model of Γ.

We need to check the conditions of a truth function. This is essentially the same as the discussion in the finite case. If we do not have either ϕ or $\neg\phi$ in S this would contradict maximality. If it did not satisfy one of the other properties it would contradict the consistency of S. □

We can now approach Gödel's completeness Theorem for propositional logic.

Theorem 3.37 (Gödel's completeness Theorem for propositional logic)**.** *If* $\Gamma \vDash \phi$ *then* $\Gamma \vdash \phi$.

Remark 3.38. We hence have that $\Gamma \vdash \phi$ if and only if $\Gamma \vDash \phi$.

Proof. Suppose that $\Gamma \vDash \phi$. Then there is no model of $\Gamma \cup \{\neg\phi\}$. Hence by proposition 3.32 we must have that $\Gamma \cup \{\neg\phi\}$ is inconsistent.

That is to say $\Gamma \cup \{\neg\phi\} \vdash \perp$. By the Deduction Theorem 3.18 we hence have that $\Gamma \vdash (\neg\phi) \to \perp$, that is to say $\Gamma \vdash ((\phi \to \perp) \to \perp)$. Applying the second axiom and modus ponens shows that $\Gamma \vdash \phi$. □

3.6. **Compactness Theorem.** We here prove a corollary of Proposition 3.32. Note that because any proof by definition only uses finitely many axioms, we have that every

Theorem 3.39 (Compactness Theorem for Propositional Logic)**.** *Suppose that every finite subset of* Γ *admits a model. Then* Γ *admits a model.*

Proof. If every finite subset of Γ admits a model. Then by proposition 3.32 shows that every finite subset of Γ is consistent. However because any proof $\Gamma \vdash \perp$ must by definition of proof use only finitely many elements of Γ, we have that if every finite subset of Γ is consistent then Γ is consistent. Applying proposition 3.32 again shows that Γ has a model. □

Remark 3.40. This proof is rather disappointing in that given a model for every finite subset of Γ we have no idea how to construct a model for Γ. We will solve this problem later using ultrafilters.

[2]By $\subset$ we include the possibility of equality.

[3]This exists by taking the union of all such subsets.

[4]I.e. is equal to all elements less that a certain element.

[5]Not necessarily unique.

Remark 3.41. In some ways the Compactness Theorem (3.39) can be viewed as more fundamental that Gödel's theorem (or proposition 3.32). This is because it doesn't depend on which notion of proof you use – instead it is purely about semantics. Of course a proof theorist would probably not see it this way!

Remark 3.42. Suppose that we modify propositional logic to allow infinite disjunctions (that is to say we allow sentences of the form $\vee_{i=1}^{\infty}\phi_i$), then the compactness theorem does *not* hold in this new setting, as the below example shows:

Consider the case where

$$\Gamma := \{\vee_{i=1}^{\infty} A_i, \neg A_1, \neg A_2, \neg A_2, ...\},$$

for A_i (i a positive integer) a propositional variable.

Clearly any finite subset of Γ has a model, however Γ does not have a model.

3.7. **Boolean Algebras.** We want to give truth functions/models a more algebraic structure.

Definition 3.43 (Boolean Algebra)**.** A *Boolean Algebra* is a sextuple $(B, \wedge, \vee, \neg, 0, 1)$, where B is a set, $0, 1 \in B$, $\neg : B \to B$, and $\wedge, \vee : B \times B \to B$ are associative and commutative binary operations.

These must satisfy the boolean relations (for all $a, b, c \in B$):

$$\begin{aligned} a \vee (a \wedge b) &= a \\ a \wedge (a \vee b) &= a \\ a \vee 0 &= a \\ a \wedge 1 &= a \\ a \vee (b \wedge c) &= (a \vee b) \wedge (a \vee c) \\ a \wedge (b \vee c) &= (a \wedge b) \vee (a \wedge c) \\ a \vee \neg a &= 1 \\ a \wedge \neg a &= 0. \end{aligned}$$

Example 3.44. Take $B = \mathcal{P}(S)$ (the power set of S) for some set S. For $A, B, C \in \mathcal{P}(S)$ define:

$$\begin{aligned} A \wedge B &:= A \cap B, \\ A \vee B &:= A \cup B, \\ \neg A &:= \{s \in S | s \notin A\}, \\ 1 &= S, \\ 0 &= \varnothing. \end{aligned}$$

Exercise 3.45. Show that example 3.44 satisfies the axioms (boolean relations) of a boolean algebra.

Example 3.46. Consider the case of example 3.44 where $|S| = 1$. Then $B = \{\varnothing, S\}$. This gives a two element boolean algebra that we also denote by $\{0, 1\}$, where operations agree with logical operations.

Definition 3.47. A morphism of boolean algebras from $(B, \wedge_B, \vee_B, \neg_B, 0_B, 1_B)$ to $(C, \wedge_C, \vee_C, \neg_C, 0_C, 1_C)$ is a morphism $f : B \to C$ such that for all $b_1, b_2 \in B$:

- $f(b_1 \wedge_B b_2) = f(b_1) \wedge_C f(b_2)$.
- $f(b_1 \vee_B b_2) = f(b_1) \vee_C f(b_2)$.
- $f(0_B) = 0_C$.
- $f(1_B) = 1_C$.

Exercise 3.48. Show that if f is a morphism of boolean algebras as above we have $f(\neg_B b_1) = \neg_C f(b_1)$.

We now want to give another way to describe boolean algebras in terms of lattices.

Definition 3.49. A *lattice* is a partially ordered set S such that any two elements have a unique supremum and infimum.

Remark 3.50. This gives two binary operations $S \times S \to S$ which we denote by $a \vee b$ and $a \wedge b$ $(a, b \in S)$, taking two elements to their supremum and infimum respectively.

Example 3.51. Consider the lattice given by $\mathcal{P}(A)$ for some set A, with the partial order given by inclusion.

Definition 3.52. A *distributive* lattice S is a lattice in which the operations $\vee$ and $\wedge$ distribute over each other. That is to say:

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$$
$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

for all $a, b, c \in S$.

Definition 3.53. Let S be a lattice with a greatest element, which we will denote 1, and a least element that we will denote 0.

A *complement* of $x \in S$ is an element $y \in S$, such that $x \wedge y = 0$, $x \vee y = 1$.

Exercise 3.54. Show that for a distributive lattice a complement is unique if it exists. In this case we write the complement of x as $\neg x$.

Definition 3.55. A *complemented* lattice is a lattice, with greatest element, and least element, such that every element has a complement.

Proposition 3.56. *Boolean Algebras are equivalent to complemented, distributive lattices.*

Remark 3.57. This isn't really the best formulation of this statement. After developing category theory, we realize that we have already defined a category of boolean algebras. Similarly we have a category of complemented, distributive lattices, which is the subcategory of the category of partially ordered sets[6] determined by the class[7] of complemented, distributive lattices as a subclass of the class of partially ordered sets. We can then say that there is an *equivalence of categories* between the category of boolean algebras, and the category of complemented, distributive lattices. Furthermore this is given by two inverse functors, which have been specified on objects above.

If you know category theory a good exercise is to provide the background details for this remark (ie. define the functors on morphisms, and show that they are inverse to each other).

Proof. Let $(B, \wedge_B, \vee_B, \neg_B, 0_B, 1_B)$ be a Boolean algebra. We define a partial order $\leq$ on B as follows:

Define $a \leq_B b$ $(a, b \in B)$ if $b = a \vee c$ for some $c \in B$.

Exercise: Show that this is a complemented, distributive lattice (with the complement of $a \in B$ given by $\neg a$).

Exercise: Show the converse, that is given a complemented, distributive lattice S, we have previously defined operations $\wedge, \vee, \neg$, and elements $0, 1$. Show that these satisfy the axioms of a Boolean algebra.

Exercise: Show that if we start with a Boolean algebra, construct the partial order $\leq_B$, and then construct a Boolean algebra from the complemented, distributive lattice $(B, \leq_B)$ we obtain the Boolean algebra we started with. Conversely show that if we start with a distributive complemented lattice, form a Boolean algebra from it, and then form a complemented distributive lattice from that we end up with the lattice we started with. □

[6] Where morphisms are any morphisms of sets with the property that they preserve the partial order; that is to say that we consider maps of sets $f : C \to D$ such that if $a <_C b$ we have $f(a) <_D f(b)$.

[7] We of course need to deal with categories where the objects do not form a set. We mainly deal with this problem by ignoring it, though we note that we can either use a version of set theory in which appropriate classes are defined [e.g. Gödel–Bernays–Von-Neumann] or use all the sets of an appropriately small model of ZFC.

Construction 3.58 (Lindenbaum–Tarski Algebra for Propositional Logic). Let Γ be consistent. Consider the equivalent relation on $Form(L)$ given by $\phi \sim_\Gamma \psi$ if $\Gamma \vdash \phi \leftrightarrow \psi$.

Consider the partial order on $Form_L / \sim_\Gamma$ given by $\phi \preceq_\Gamma \psi$ if $\Gamma \vdash \phi \to \psi$.

Theorem 3.59. *The set with partial order defined in construction 3.58 is a complemented, distributive lattice, and hence a Boolean algebra.*

We denote this Boolean algebra by $Lind(L, \Gamma)$

Proof. It is clear that we have a partially ordered set. The least upper bound of two elements $[\phi]$, $[\psi]$ is $[\phi \vee \psi]$ (exercise: check this is well defined – that is to say we would have got the same equivalence class if we had chosen different representations of $[\phi]$ and $[\psi]$). The greatest lower bound is $[\phi \wedge \psi]$. hence $Lind(L, \Gamma)$ is a lattice.

Furthermore this lattice is distributive, and has complements $[\phi] = [\neg\phi]$.

The result follows. □

Proposition 3.60. *Models of Γ correspond to morphisms of boolean algebras*

$$Lind(L, \Gamma) \xrightarrow{t'} \{0, 1\}.$$

Proof. Firstly suppose that we have a morphism of boolean algebras t'. We get a truth function t by the composition

$$Form(L) \to Form(L)/ \sim_\Gamma \xrightarrow{t'} \{0, 1\}.$$

Exercise: Show that this is a model of Γ as in definition 3.25.

Secondly suppose that we have a model $t : Form(L) \to \{0, 1\}$.

By the definition of a model it descends to $t' : Form(L) \to \{0, 1\}$.

Exercise: Show that the definition of a truth function shows that this induces a morphism of boolean algebras. □

It is now worth considering the preimage of 1 under a morphism of Boolean algebras.

Definition 3.61. A *filter* of B is a subset $F \subset B$, with the property that for all $f_1, f_2 \in F$, $b \in B$ we have $f_1 \wedge f_2 \in F$, and $b \vee f_1 \in F$.

Proposition 3.62. *Let $B \xrightarrow{f} C$ be a morphism of Boolean algebras. We then have that $f^{-1}(1_C)$ is a filter.*

Exercise 3.63. Prove proposition 3.62.

Exercise 3.64. Conversely show that if $F \subset B$ is a filter, then there is a morphism of Boolean algebras $B \xrightarrow{f} C$ such that $f^{-1}(1) = F$.

Hint: Construct $C = B/ \sim_F$, for the equivalence relation $b_1 \sim_F b_2$ if
$((\neg b_1) \vee b_2) \wedge (b_1 \vee (\neg b_2)) \in F$

Note: This may appear unmotivated, however it should seem significantly more motivated when you see the definition of the Lindenbaum–Tarski algebra and see that this formula is precisely $b_1 \leftrightarrow b_2$.

Definition 3.65. We say a filter $F \subset B$ is *proper* if $B \neq F$.

Definition 3.66. An *ultrafilter* is a filter F with the property that for all $b \in B$ either $b \in F$ or $\neg b \in F$.

Exercise 3.67. If F is an ultrafilter, show that there is no proper filter F' with $F \subsetneq F'$.

Exercise 3.68. Show that if $B \xrightarrow{f} \{0, 1\}$ is a morphism of Boolean algebras, then $f^{-1}(1)$ is an ultrafilter.

Exercise 3.69. Conversely show that if $F \subset B$ is an ultrafilter then there is a morphism $B \xrightarrow{f} \{0, 1\}$, with the property that $f^{-1}(1) = F$.

Definition 3.70. For convenience we restate the definition of a filter in the case where the Boolean algebra is a special case of example 3.44.

An *ultrafilter on a set* S is a set $F \subset \mathcal{P}(S)$, such that for all $A \subset S$, either $A \in F$ or $S \backslash A \in S$, and such that F is closed under union, and closed under intersection with sets in $\mathcal{P}(S)$.

Definition 3.71. An *ultrafilter on a set* S is a set $F \subset \mathcal{P}(S)$, such that:

- $\emptyset \notin F$.
- If $A, B \in \mathcal{P}(S)$, $A \subset B$, $A \in F$, then $B \in F$.
- If $A, B \in F$ then $A \cap B \in F$.
- If $A \in \mathcal{P}(X)$ then either A or $S \backslash A$ is in F.

Exercise 3.72. Show that definitions 3.70 and 3.71 are equivalent.

Example 3.73. Let $s \in S$. An ultrafilter on S is given by $\mathcal{U}_s := \{A \in \mathcal{P}(S) | s \in A\} \subset \mathcal{P}(S)$ (Exercise: Why is this an ultrafilter?). We call ultrafilters of this form *principal ultrafilters*.

Proposition 3.74. *Any filter $F \subset B$, $F \neq B$ is contained inside an ultrafilter.*

Proof. Let $F \subset B$ be a filter $F \neq B$. Consider the set of all filters $F \subset F' \neq B$, ordered by inclusion. We will apply Zorn's lemma to this set. We claim that for any chain $\{F_i\}_{i \in I}$, $\cup_{i \in I} F_i$ is an upper bound for the chain. To justify this we need to check that $\cup_{i \in I} F_i$ is a filter, and $\cup_{i \in I} F_i \neq \mathcal{P}(S) \neq B$ (Exercise). Zorn's lemma shows that there exists a maximal element which we denote F_{max}.

We claim that F_{max} is an ultrafilter. Suppose not. Then there exists $b \in B$ such that $b, \neg b \notin B$.

Let

$$G := \{d \in B | \exists c \in F_{max}, e \in B \ \ (c \wedge b) \vee e = d\}.$$

Claim: G is a filter, and $G \neq B$.

Exercise: Show that G is a filter.

Suppose $0_B \in B$ Then we would have $0_B = (c \wedge b)$ for some $c \in F_{max}$ (note that we must have $e = 0_B$). Then $c \vee (\neg b) = \neg b \in F_{max}$. Contradiction. Hence $0_B \notin B$.

Then G contradicts the maximality of F_{max}, hence the assumption that $b, \neg b \notin F_{max}$ is false. Hence F_{max} is an ultrafilter. $\square$

We introduce the following definition for convenience:

Definition 3.75. We call a filter G *proper* if $G \neq B$.

Proposition 3.76. *Non-principal ultrafilters exist.*

Proof. Firstly take all cofinite subsets of an infinite set S (that is to say A such that $S \backslash A$ is finite). This is a filter (check) which we denote F on $\mathcal{P}(S)$.

This is contained inside some ultrafilter by proposition 3.74. It is clear the ultrafilter is not principal. $\square$

Example 3.77. We can see from the above proof that ultrafilters on $\mathcal{P}(S)$ can be "constructed" from filters F such that for any point $s \in S$, there is $V \in F$, with $s \notin V$.

One way to construct such sets is as follows. Let S be infinite. Embed $S \hookrightarrow i\mathbb{R}$, such that there exists a limit point $a \in \mathbb{R}$, $a \notin i(S)$.

Define F_a to be the filter consisting of sets $L \subset S$, such that there exists $\epsilon > 0$, such that $L \supset i(S) \cap B_\epsilon(a)$, where $B_\epsilon(a) := \{y \in \mathbb{R} || y - a| < \epsilon\}$.

Note that we could clearly have replaced $\mathbb{R}$ with any metric space, or with slightly more care by any topological space.

Proposition 3.78. *Suppose that a subset $X \subset \mathcal{P}(S)$, has the property that the intersection of any finite subset of X is non-empty. There exists an ultrafilter $\mathcal{U}$ with $X \subset \mathcal{U}$.*

Proof. We note that $X' := \{V \subset S | \exists U \in X, V \supset U\}$ is a filter, and is not equal to $\mathcal{P}(S)$. Hence by proposition 3.74 we have that $X \subset X' \subset \mathcal{U}$ for some ultrafilter $\mathcal{U}$. □

Lemma 3.79. *Any non-principal ultrafilter $\mathcal{U}$ on a set S contains the cofinite filter*

$$\mathcal{F} := \{S' \subset S | S \backslash S' \text{ finite}\}.$$

Proof. By the ultrafilter property for any $s \in S$ either $\{s\}$ or $S \backslash s$ is in $\mathcal{U}$. As by assumption $\mathcal{U}$ is non-principal we have that $(S \backslash s) \in \mathcal{U}$.

Hence for any finite set $\{s_1, ...s_n\} = F \subset S$, we have that

$$S \backslash F = \cap_{i=1}^{n} (S \backslash s_i) \in \mathcal{U}.$$

□

It is also worth asking what the precise relationship between theories in propositional logic (recall definition 3.7). More precisely we can ask:

Question 3.80. Which Boolean algebras are of the form $Lind(L, \Gamma)$ for some theory in propositional logic (L, Γ)?

Exercise 3.81. Let B be a boolean algebra. Show that there is a theory in propositional logic (L, Γ) such that $Lind(L, \Gamma) \cong B$.

Hint: Set $L = B \backslash \{0_B, 1_B\}$. Add propositions to Γ as needed.

Call this theory $Th(B)$.

Question 3.82. If there is a morphism $Lind(L_1, \Gamma_1) \to Lind(L_2, \Gamma_2)$ then what is (if any) the corresponding relationship between the theories (L_1, Γ_1) and (L_2, Γ_2)?

In particular if $Lind(L_1, \Gamma_1) = Lind(L_2, \Gamma_2)$, what is the relationship between (L_1, Γ_1) and $(L_2, \Gamma 2)$?

We will not fully answer this question, but will make a start at suggesting the maps Th and $Lind$ are *functorial*, in the sense that they are compatible with morphisms of boolean algebras, and interpretations of theories in propositional logic.

Definition 3.83. We say that an *interpretation* of (L_1, Γ_1) in (L_2, Γ_2) is a morphism $s : L_1 \to Form(L_2)$, with the below detailed properties.

Firstly note that s naturally extends to a function, which we also denote s, $s : Form(L_1) \to Form(L_2)$.

We require that if $\Gamma_1 \vdash \phi$, then $\Gamma_2 \vdash s(\phi)$.

We denote an interpretation by

$$(L_1, \Gamma_1) \xrightarrow{s} (L_2, \Gamma_2).$$

Exercise 3.84. Show that given an interpretation $(L_1, \Gamma_1) \xrightarrow{s} (L_2, \Gamma_2)$ we get a morphism of Boolean algebras

$$Lind(L_1, B_1) \to Lind(L_2, B_2).$$

Remark 3.85. We can go significantly further, alas probably beyond the scope of this course. The state of the art statement is that Th, and $Lind$ are an *adjoint functor pair*[8] between the categories of theories in propositional logic (morphisms being interpretations) and the category of Boolean algebras.

The perhaps important point, is that ultimately one can consider the category of Boolean algebras, and that of first order theories to be essentially the same.

There is a second part of this relation, which is *Stone duality*, which says there is an equivalence of categories between the category of Boolean algebras, and the category of Stone spaces. The key point is that the set of models of a theory in propositional logic, has some extra structure (that of a topological space) that makes it into a stone space. We hence have an equivalence between the syntactic

[8] Or in fact weak inverses.

"understanding" (the propositional theory, and/or the Boolean algebra), and the semantic "understanding" the Stone space of models of a theory. While we are going to try to get to Stone duality it is by no means certain that we can make it!

3.8. **Stone Duality.** (Todo: To be Written)

4. First Order Logic

Propositional Logic was clearly very limited in which mathematical statements we could say using it. First order logic allows us to consider families of statements $\phi(x)$ depending on a variable x, and also to quantify over variables. That is to say that we are allowed statements such as $\forall x(\phi(x))$ and $\exists x(\phi(x))$.

4.1. **First Order Theories.**

Definition 4.1. A *first order language with equality* corresponds to a set L which is the union of six disjoint subsets:

- Constants $c_1, ...$ (we may write as $\{c_i\}_{i\in I}$).
- Countably many variables $x, y, z, ...$
- Logical symbols: $\vee$, $\wedge$, $\to$, $\leftrightarrow$, $\neg$, $\top$, $\bot$, $\exists$, $=$, and $\forall$.
- Functions f, or a specified rank (or arity) n (n a non-negative integer [or can assume a positive integer]).
- Relations R, again of a specified rank, n a non-negative integer.
- Parentheses (brackets).

We define recursively, terms and formulas of the language.

The *terms* of the language are the smallest set of strings of characters in L that contain the constants, the variables, and for any function f of rank n contains $f(x_1, ..., x_n)$ where $x_1, ..., x_n$ are terms. Note: $A.$ $Priori.$ it may not be clear that there is a smallest such set, use an argument similar to that in the definition of formula in a first order language to show that this is well defined. We may denote this set $Term(L)$.

The *formula* of the language make up the smallest set $Form(L)$ of strings of characters in L that contains:

- $R(t_1, ..., t_r) \in Form(L)$ for R a relation of degree r, and $t_1, ..., t_r \in Term(L)$.
- If $A, B \in Form(L)$,, x a variable of L then the following are in $Form(L)$:

$$(A) \leftrightarrow (B), (A) \to (B), (A) \vee (B), (A) \wedge (B), \neg(A), \forall x(A), \exists x(A). \tag{4.1}$$

- For $t_1, t_2 \in Term(L)$ we have that $=(t_1, t_2) \in Form(L)$.
- $\top$, $\bot$.

Remark 4.2. Note that we didn't include commas in the set of symbols. Strictly speaking we should write e.g. $f(t_1t_2t_3)$ rather than $f(t_1, t_2, t_3)$.

Similarly we will often write $t_1 = t_2$ rather than $=(t_1, t_2)$.

Remark 4.3. Clearly we can deal with some subset of the logical symbols. We discuss dropping the equality symbol in remark 4.22, and treating it as an ordinary relation.

However dropping some sets of symbols will result in a reduction of the expressiveness of the language.

Remark 4.4. Note that a function of rank zero must just give us some constant. Similarly a relation of rank zero is essentially a propositional variable.

Lemma 4.5 (Unique reading lemma for first order logic)**.** *The set $Term(L)$ and $Form(L)$ are disjoint. If $t \in Term(L)$ then exactly one of the following is true:*

- *t is a constant.*
- *t is a variable.*

- *t is of the form $f(t_1...t_r)$ for uniquely specified terms $t_1, ..., t_r$ and a uniquely specified f or some rank r.*

If $P \in Form(L)$, then exactly one of the following is true:

- $P = \top$.
- $P = \bot$.
- *$P = R(t_1...t_r)$, for R a uniquely specified relationship of rank r, and $t_1, ..., t_r$ uniquely specified terms in $Term(L)$.*
- *P is $=(t_1, t_2)$ for uniquely specified terms $t_1, t_2 \in Term(L)$.*
- *P is of precisely one of the following forms, where x is a variable in L and A and B are uniquely specified formula of L:*

$$(A) \leftrightarrow (B), (A) \to (B), (A) \vee (B), (A) \wedge (B), \neg(A), \forall x(A), \exists x(A). \tag{4.2}$$

Proof. Essentially exactly the same as the case for propositional logic. Again the key step is the unique bracketing lemma 3.9. □

Definition 4.6. A *first order theory* is a pair of a first order language with equality L, together with a set $\Gamma \subset Form(L)$, with the property that for each $\psi \in \Gamma$, ψ has no free variables (see definition 4.9).

Remark 4.7. Sometimes the term sentence is used to refer to a formula with no free variables.

Remark 4.8. This definition is essentially the same as the definition in the case of propositional logic.

4.1.1. *Substituting Variables.* When we say $\forall y \exists x(x^2 = y)$ (when we are talking say about the complex numbers), it is clear that the variables x and y are not particularly relevant. We could equally write $\forall y \exists z(z^2 = y)$, and this would have the same meaning, and so should be considered the same.

We could however not substitute $x = y$: the sentence $\forall y \exists y(y^2 = y)$ has a different meaning (as can be seen from the fact that unlike the former sentence) this is *false* in the complex numbers.

Definition 4.9. We say that an appearance of a variable x in a formula φ is free, if it is not inside the brackets associated to a $\forall x(-)$ predicate or a $\exists x(-)$ predicate. Otherwise we say it is not free.

We sometimes write ϕ as $\phi(x_1, ..., x_n)$, where the x_i are the variables that appear freely in the formula ϕ. Then for any a_i in M, or $Term(L)$ or in $Variables(L)$ or in $Constants(L)$ we denote $\phi(a_1, ..., a_n)$ as the formula we get by replacing each free appearance of x_i with a_i. When replacing x_i with something in $Term(L)$, this term must not contain any variable, such that the occurrence of this variable would not be free if we replaced x_i with the given term. For a discussion of this in which more care has been taken see 1.5 and 1.6 of chapter II of Manin [3].

4.1.2. *Examples:*

Example 4.10 (Peano Arithmetic). Here we have a language consisting of a constant 0, together with a function of rank 1, S (Successor function), and an two functions of rank 2, $+$, $\cdot$. We then specify the following axioms:

- $\forall x \neg(0 = S(x))$.
- $\forall x, y((S(x) = S(y)) \to (x = y))$.
- $\forall x(x + 0 = x)$.
- $\forall x(x \cdot 0 = 0)$.
- $\forall x, y(x \cdot S(y) = x \cdot y + y)$.

We also add the *first order induction axioms*, which for each k a positive integer, and each $\phi(x, y_1, ..., y_k) \in Form(L)$ using variables $x, y_1, ...y_k$ freely, is an axiom of the form

$$\forall \vec{y}((\phi(0, \vec{y}) \wedge \forall x(\phi(x, \vec{y} \to \phi(S(x), \vec{y}))) \to \forall x \phi(x, \vec{y})),$$

where $\vec{y}$ denotes $y_1, ..., y_k$.

Remark 4.11. Usually Peano arithmetic is written as a theory of second order logic, in that rather than writing $\forall\vec{y}$, we should just write $\forall X$, where we let X range over all subsets of the constants of our model. Note that in practice we only need *finite* subsets, because any formula only involves a finite number of free variables.

Example 4.12 (Group Theory). Recall that a group G is a set G together with a single operation, satisfying various axioms. We can formalize this in first order logic as follows:

We use a language consisting of a constant e, and a formula of rank 2, $\cdot$. We then impose the axioms:

- $\forall x, y, z(x \cdot (y \cdot z) = (x \cdot y) \cdot z)$ (associativity).
- $\forall x((x \cdot e = e \cdot x) \wedge (x = e \cdot x))$ (Identity).
- $\forall x(\exists y((x \cdot y = e) \wedge (y \cdot x = e)))$ (inverses).

Exercise 4.13. Write down axioms for a first order theory of rings, and a first order theory of fields.

Example 4.14 (Abstract Projective Plane). This is a formalization of the projective geometry of lines in the (projective) plane. We will call it an *abstract projective plane.*

In our language we have two relationships R_{pt} and R_{line} of rank 1. These will specify which constants are points and which are lines.

We have a relation of order 2 denoted $\in$ (we will informally denote $a \in b$ rather than $\in(a, b)$. Note that will refer to the relation of a point lying on a given line).

We then introduce the axioms:

- $\forall x((R_{pt}(x) \vee R_{line}(x)) \wedge (\neg(R_{pt}(x) \wedge (R_{line}(x)))))$. (Everything is a point or a line).
- $\forall x, y((x \in y) \to (R_{pt}(x) \wedge R_{line}(y)))$ (Incidence refers to a point lying on a line).
- $\forall x, y((R_{pt}(x) \wedge R_{pt}(y)) \to (\exists z(x \in z \wedge y \in z)))$ (There is a line passing through two points).
- $\forall x, y, z, w((x \in z \wedge y \in z \wedge x \in w \wedge y \in w) \to z = w)$(The line passing through two points is unique)
- $\forall x, y((R_{line}(x) \wedge R_{line}(y)) \to (\exists z(z \in x \wedge z \in y)))$ (There is a point lying on any two lines).
- $\forall x, y, z, w((x \in z \wedge y \in z \wedge x \in w \wedge y \in w) \to x = y)$ (The point where two lines intersect is unique)

Remark 4.15. We can also describe projective geometry starting with a field, and then using homogeneous coordinates.

If we start with an abstract projective plane that came from homogeneous coordinates we can reconstruct the coordinates. In general if we have an abstract projective plane satisfying Pappus' hexagon theorem we can recover coordinates.

This tells us that the notion of equivalence we use for first order theories is going to have to be much more involved than propositional theories in that we would like to view the first order theory of the projective plane, with Pappus theorem added as an axiom, should be equivalent to a formalization we gained by formalizing homogeneous coordinates valued in some appropriate field.

One approach to this is given by syntactic categories; these two ways to formalize the projective plane will have the same *syntactic category* considered as a category together with some additional structure. Sadly we probably won't really get to this in the course.

Example 4.16 (Algebraically Closed Fields). We need first the axioms of a field from exercise 4.13. We then add the axioms of algebraic closedness, For each $n \in \mathbb{N}$ we add an axiom:

$$\forall y_1, ..., y_n \exists x(x^n + y_1 x^{n-1} + ... + y_n = 0).$$

If we only want to consider algebraically closed fields of characteristic p we add an axiom $1+1+...+1 = 0$, where we have used $p-1$ addition signs.

If we want to consider only fields of characteristic zero, then for each prime p we add an axiom $\neg(1 + 1 + ... + 1 = 0)$, where we have again used $p - 1$ addition signs.

Example 4.17 (Ordered Field). We use the axioms for a field from exercise 4.13.

We add the a relation symbol $<$ to the alphabet. We add the axioms:

- $\forall x, y(x = y \vee x < y \vee x < x)$.
- $\forall x, y, z((x < y) \to (x + z < y + z))$
- $\forall x, y((0 < x \wedge 0 < y) \to 0 < x \cdot y)$
- $\forall x, y, z((x < y \wedge y < z) \to x < z)$.
- $\forall x, y(\neg(x < y) \wedge (y < x))$
- $\forall x(\neg(x < x))$.

Example 4.18 (Real Closed Field). We start by taking the axioms from example 4.17. We then add the axiom

$$\forall x((x > 0) \to (\exists y(y \cdot y = x))),$$

and for every odd integer $n > 0$ we add the axiom

$$\forall y_1, ..., y_n \exists x(x^n + y_1 x^{n-1} + ... + y_n = 0).$$

Remark 4.19. You may not have come across the notion of a real closed field before. The idea behind the definition is that any real closed field has the same first order properties as the real numbers. That is to say if ϕ is a first order sentence that is true for $\mathbb{R}$, then it is true for any other real closed field.

Other examples of real closed fields are the real algebraic numbers, and the hyperreal numbers.

4.2. **Models.**

Definition 4.20. A *model* of a first order theory (L, Γ) is a set[9] M, together with maps:

$$\phi : Constants(L) \to M,$$

$$\phi : Functions^n(L) \to Hom(M^n, M)$$

$$\phi : Relations^n(L) \to \mathcal{P}(M^n).$$

To a model M (or (M, ϕ)) we construct the additional functions:

Let $\overline{M} = Hom(Variables(L), M)$.

To each formula ψ we assign

$$\phi(\psi) : \overline{M} \to \{1, 0\}.$$

Informally this is the function that is 1 when the formula is true for the given values of the variables, and is 0 when the formula is false for the given values of the variables. If the formula ψ only contains n free[10] variables $x_1, ..., x_n$, then this corresponds to a map $M^n \to \{1, 0\}$, and $\phi(\psi)$ is obtained as the composition

$$\overline{M} \to M^n \to \{0, 1\},$$

where the map $\overline{M} \to M^n$ records the image of the variables $x_1, ...x_m$ under a given morphism of sets.

We will sometimes also denote the map $M^n \to \{0, 1\}$ by $\phi(\psi)$.

Note that we can also see this as a map from a formula ψ to the set $\phi(\psi)^{-1}(1) \subset M^n$. This is the viewpoint utilised by the description of first order theories and their models via the *syntactic category.*

Construction 4.21. We now construct $\phi(\psi)$ rigorously.

Firstly note that for any term t we have a map $\phi(t) : \overline{M} \to M$. This is because, this is defined as follows:

- For constants $\phi(c_1)$ is the constant map $\overline{M} \ni \xi \mapsto \phi(c_1)$ for all $\xi \in \overline{M}$.

[9]In fact can we don't need this to be a set. It can e.g. be a class in the sense of e.g. Von Neumann–Bernays–Gödel set theory. For this course we will ignore this subtlety and assume it is a set.

[10]To be explained

- For a variable x the map $\phi(x)$ takes $\overline{M} \ni \xi : Variables(L) \to M$ to
$$(\phi(x))(\xi) := \xi(x).$$
- If we have already defined $\phi(t_1), ..., \phi(t_r) \in Hom(\overline{M}, M)$ we define $\phi(f(t_1, ..., t_r))$ (for f a function of rank n) as the composition
$$\overline{M} \xrightarrow{(\phi(t_1),...,\phi(t_r))} M^r \xrightarrow{\phi(f)} M$$

and if we have $\phi(t_1), ..., \phi(t_r) \in M$, then we define
$$\phi(f(t_1, ..., t_r)) := \phi(f)(\phi(t_1), ..., \phi(t_r)).$$

We now show that for any formula P we have a function $\phi(P) : \overline{M} \to M$. We define this as follows:

- Suppose that R is a relation of rank r. We define
$$\phi(R(t_1, ..., t_r))(\xi) = \begin{cases} 1, & \text{if } ((\phi(t_1))(\xi), ..., (\phi(t_r))(\xi)) \in \phi(R) \\ 0 & \text{otherwise.} \end{cases}$$
- We define
$$\phi(= (t_1, t_2)) = \begin{cases} 1, & \text{if } ((\phi(t_1))(\xi) = (\phi(t_r))(\xi)) \\ 0 & \text{otherwise.} \end{cases}$$
- We define
  - $\phi(\neg A) = 1 - \phi(A)$.
  - $\phi(A \vee B) = \max(\phi(A), \phi(B))$.
  - $\phi(A \wedge B) = \min(\phi(A), \phi(B))$.
  - $\phi(\bot) = 0$ (The constant function).
  - $\phi(\top) = 1$.
  - $\phi(A \to B) = \max(1 - \phi(A), \phi(B))$.
  - $\phi(\forall x A)(\xi) = \min_{\upsilon}(\phi(A)(\upsilon))$ where we take the minimum over all $\upsilon \in \overline{M}$, such that for all variables $y \neq x$, we have $\xi(y) = \upsilon(y)$.
  - $\phi(\exists x A)(\xi) = \max_{\upsilon}(\phi(A)(\upsilon))$ where we take the minimum over all $\upsilon \in \overline{M}$, such that for all variables $y \neq x$, we have $\xi(y) = \upsilon(y)$.

Remark 4.22. Rather than including equality as a logical symbol we could just treat it as a relation of order 2. It is then the case that some models would have the relation $\phi(=)$ as a relation other than the equality symbol in some models would be mapped to a relation other than equality on M.

We call a model *normal* if $\phi(=)$ is equality on M.

If you work with first order theories without equality, the statements of some theorems need to be changed so they refer only to normal models.

Definition 4.23. We say that $M \vDash \psi(a_1, ...a_r)$ if $\phi(\psi(a_1, ..., a_r)) = 1$ (on the right hand side we should interpret 1 as the corresponding constant function on M).

We say that $\Gamma \vDash \psi(a_1, ..., a_r)$ if for every model M of Γ we have $M \vDash \psi(a_1, ..., a_r)$.

Note that we must follow the rules for when and how substitutions are allowed developed in section 4.1.1.

Remark 4.24. We are however going to be significantly more sloppy. Sometimes we are going to write $M \vDash \psi$ for a formula ψ which includes variable(s) x (we'll just give the version with one variable). In this case we should either see this as $M \vDash \psi(c)$, where we replace the variable with a new constant (ie. we expand L to add a new constant), or equivalently, as $M \vDash \forall x \psi(x)$.

Definition 4.25 (Morphisms and Elementary embedding)**.** There are various notions of morphisms of models of a first order theory.

For this course we will provisionally define a morphism between two models M_1 and M_2 of the same first order theory to be a map of sets $M_1 \xrightarrow{f} M_2$ with the property that this commutes with the maps ϕ_1, ϕ_2, in the following senses:

This means e.g. that the diagram

$$\begin{tikzcd}
 & M_1 \arrow[dd, "f"] \\
Constants(L) \arrow[ur, "\phi_1"] \arrow[dr, "\phi_2"'] & \\
 & M_2
\end{tikzcd}$$

commutes (that is to say $\phi_2 = f \circ \phi_1$).

For a function $g \in Functions^n(L)$ one imposes that $f \circ \phi_1(g) = \phi_2(g) \circ \Delta(f)$, where $\Delta(f) : M_1^n \to M_2^n$ is the map given by applying f to each factor.

For a relation $R \in Relations^n(L)$. We denote $\mathcal{P}(f))$ to be the map $\mathcal{P}(M_1^n) \to \mathcal{P}(M_2^n)$ induced by $\Delta(f)$. We require that $\mathcal{P}(f)(\phi_1(R)) \subset (\phi_2(R))$.

We define an *elementary morphism* to be an morphism of models with the property that for every formula $\phi(x_1, ..., x_n) \in Form(L)$, in variables $x_1, ...x_n$ we have that

$$M_1 \vDash \phi(a_1, ..., a_n) \Leftrightarrow M_2 \vDash \phi(f(a_1), ..., f(a_n)).$$

We define an *elementary embedding* to be an elementary morphism which is an injection of sets $M_1 \xrightarrow{f} M_2$.

Definition 4.26 (Definable Set)**.** Let M be a model. We say that a set $S \subset M^n$ is *definable*, if there is $\psi(x_1, ..., x_{m+n}) \in Form(L)$ a formula with $m + n$ free variables, and there is $\vec{b} \in M^m$, with the property that

$$\left(M \vdash \psi(\vec{a}, \vec{b})\right) \Leftrightarrow \vec{a} \in S.$$

4.3. **Syntactic Category.** (Todo: To be Done: This is (an) analogue of the Boolean Algebra story for propositional logic)

4.4. **Models as Functors.** (Todo: This)

4.5. **Ultraproducts and Compactness.** We recall from definitions 3.70 and 3.71 the notion of ultrafilters on a set S.

We will use this to define an element called the ultraproduct.

Definition 4.27. Let $\{A_s\}_{s \in S}$ be a set of sets parameterized by the set S. Furthermore let $\mathcal{U} \subset \mathcal{P}(S)$ be an ultrafilter on S. We then define the *(set theoretic) ultraproduct* of $\{A_s\}_{s \in S}$ along the ultrafilter $\mathcal{U}$ as the quotient of

$$\prod_{s \in S} A_s,$$

by the equivalence relation $\sim_{\mathcal{U}}$ which we define as follows. Let $a := (a_s)_{s \in S}, b := (b_s)_{s \in S} \in \prod_{s \in S} A_s$. We write $a \sim_{\mathcal{U}} b$ if the set $S \supset D := \{s \in S | a_s \neq b_s\} \notin \mathcal{U}$.

We write the ultraproduct as:

$$\prod_{\mathcal{U}} A_s := \left(\prod_{s \in S} A_s\right) / \sim_{\mathcal{U}} .$$

Remark 4.28. Note that $\prod_{s \in S} A_s$ can be defined as functions f from $S \to \coprod_{s \in S} A_s$, such that $f(s) \in A_s$.

Example 4.29. Take $\mathcal{U} = \mathcal{U}_j$ a principal ultrafilter as defined in example 3.73 for a fixed element $j \in S$. We then have that

$$\prod_{\mathcal{U}} A_s = A_j.$$

Exercise 4.30. Let $S \subset \mathbb{N}$ be the set of prime numbers. Pick a non-principal ultrafilter $\mathcal{U}$ on S. Show that

$$\prod_{\mathcal{U}} \overline{\mathbb{F}}_p$$

is a field (look up axioms if needed) of characteristic zero (with the field operations inherited from $\prod_{p\in S} \overline{\mathbb{F}}_p$). Here $\overline{\mathbb{F}}_p$ is the algebraic closure of $\mathbb{F}_p$ (i.e. we add every number that is a root of a polynomial over $\mathbb{F}_p$).

What is the cardinality of this ultraproduct?

Is there a map $\overline{\mathbb{Q}} \to \prod_{\mathcal{U}} \overline{\mathbb{F}}_p$? Is it injective?, surjective?

4.5.1. *Application of Ultrafilters to Ramsey Theory.* Let G be a graph [11], where the set of edges is coloured red or blue. That is to say we partition the set of edges into $E = E_R \coprod E_B$. We require that G has no loops (ie. no edges from a vertex to itself, and between every pair of vertices there is at most one edge).

Proposition 4.31. *Let G as above be infinite. There is an infinite subset of the vertices $V' \subset V$, such that all edges between vertices of V' are the same colour.*

Proof. Firstly we note that we can assume without loss of generality that every pair of distinct vertices contains an edge.

We let $\mathcal{U}$ be a non-principal ultrafilter on the set of vertices V.

For an edge between vertices (u, v) we write $c(u, v) \in \{B, R\}$ as the colour of the edge between u and v.

We define $c(u)$ as the colour such that the set of vertices $\{v\}$, such that $c(u.v)$ is of the given colour, is in $\mathcal{U}$.

We note that there must be at least one colour such that the set of u with $c(u)$ being the given colour is in $\mathcal{U}$. Without loss of generality let this be blue. We let the given set be denoted by $A \in \mathcal{U}$

We pick $u_1 \in A$, and define $A_1 := \{u \in A | c(u, u_1) = B\}$. We iterate this process (pick $u_i \in A_{i-1}$, let $A_i := \{\in A_{i-1} | c(u, u_i) = B\}$).

Take $\{u_n\}_{n\in\mathbb{N}}$ as the desired infinite graph. (Exercise: Why does the process not terminate?) □

Exercise 4.32. Convince yourself that we can generalize proposition 4.31 to the case where we use finitely many colours.

Exercise 4.33. Where does the proof of proposition 4.31 go wrong if you used a principal ultrafilter?

Exercise 4.34. Compare the above proof of proposition 4.31 to a proof that does not use ultrafilters.

Proposition 4.35. *Show that there exists $N = N(M, r)$, such that if we colour all edges of a graph G with N vertices from a set of r colours, then there exists a subset $V' \subset V$ of the vertices V, such that $|V'| = M$, and all edges between vertices in V' are the same colour.*

Exercise 4.36. Use the compactness theorem for first order logic (not yet covered in notes − Theorem 4.44) to prove proposition 4.35.

Hint: Develop a first order theory of coloured graphs.

4.5.2. *Other Applications.*

Remark 4.37. There are many other interesting applications of ultrafilters, and ultraproducts. One is of course the compactness theorem, and other can be found in the applications of the compactness theorem section here.

I really wanted to include the applications to topological dynamics, and ergodic theory, however these use the topological structure of the space of ultrafilters on a set S in a serious way, and as such I didn't feel able to do this.

[11] that is to say a set of vertices V, and a set of edges E, each edge connects two vertices [that is to say there is a map $E \to V^2/S_2$, where the quotient by S_2 "forgets" the order of the pair of vertices. If we were working with directed graphs we would not perform this quotient].

4.5.3. *Łoś Ultraproduct Theorem.* We now show that we can define ultraproducts in the setting (category) of models of theories in first order logic (with equality).

Let $\{M_s\}_{s\in S}$ be a set of models of a first order theory (L,Γ). Pick an ultrafilter $\mathcal{U}$ on S. We can define the set theoretic ultraproduct

$$\prod_{\mathcal{U}} M_s := \left(\prod_{s\in S} M_s\right) / \sim_{\mathcal{U}} .$$

To give this the structure of a model of L,Γ, we need to provide maps

$$\phi_{\mathcal{U}} : Constants(L) \to \prod_{\mathcal{U}} M_s,$$
$$\phi_{\mathcal{U}} : Functions^n(L) \to Hom(\prod_{\mathcal{U}} M_s^n, \prod_{\mathcal{U}} M_s)$$
$$\phi_{\mathcal{U}} : Relations^n(L) \to \mathcal{P}(\prod_{\mathcal{U}} M_s^n).$$

We first define the map on constants as the composition:

$$Constants(L) \xrightarrow{\prod_{s\in S}\phi_s} \prod_{s\in S} M_s \to \prod_{\mathcal{U}} M_s.$$

Secondly for the map on functions we note that for a function f the map

$$\prod_{s\in S} \phi_s(f) \in Functions^n(L) \to Hom((\prod_{s\in S} M_s)^n, \prod_{s\in S} M_s)$$

fits into a unique (why?) commutative diagram, which defines $\phi_{\mathcal{U}}(f)$:

$$\begin{tikzcd}
(\prod_{s\in S} M_s)^n \arrow[r, "\prod_{s\in S}\phi_s(f)"] \arrow[d] & \prod_{s\in S} M_s \arrow[d] \\
(\prod_{\mathcal{U}} M_s)^n \arrow[r, "\phi_{\mathcal{U}}(f)"] & \prod_{\mathcal{U}} M_s.
\end{tikzcd}$$

We now need to define the relations. For a relation R or rank r we define the relation $\phi_{\mathcal{U}}(R) \in \mathcal{P}((\prod_{\mathcal{U}} M_s)^n)$ as the set

$$\{([(a_{1,s})_{s\in S}], ..., [(a_{r,s})_{s\in S}]) | \{s \in S | \phi_s(R) \ni (a_{1,s}, ..., a_{r,s})\} \in \mathcal{U}\}.$$

Exercise 4.38. Check that $\phi_{\mathcal{U}}(R)$ is well defined for a relation R of the language L. Note that this is necessary because we made a choice of representatives of the equivalence classes of $\sim_{\mathcal{U}}$, a. priori. the above definition could depend on this choice.

Furthermore we have that these maps make $\prod_{\mathcal{U}} M_s$ into a model for (L,Γ).

Now that we have defined the ultraproduct of a set of models of a first order theory (L,Γ) we will prove the following theorem

Theorem 4.39 (Łoś Ultraproduct Theorem)**.** *Let $\{M_s\}_{s\in S}$ be models of a first order theory (L,Γ). Let $\mathcal{U}$ be an ultrafilter on S and consider the model $\prod_{\mathcal{U}} M_s$.*

For every $\psi \in Form(L)$ we have that $\prod_{\mathcal{U}} M_s \models \psi([\{a_{1,s}\}], ..., [\{a_{n,s}\}])$ if and only if $\{s \in S | M_s \models \psi(a_{1,s}, ..., a_{n,s})\} \in \mathcal{U}$.

Proof. Firstly we note that substitution/variable assignment is compatible with the ultrafilter structure. That is to say for any term $t = [(t_j)_{j\in S}]$, we have that

$$t[[\{a_{1,s}\}], ..., [\{a_{n,s}\}]] = [(t_j\{a_{1,j}\}], ..., [\{a_{n,j}\}])_{j\in S}].$$

If this is not clear one way to rigorously prove it is to use the recursive definition of terms.

We use the recursive definition of $Form(L)$. We first show this claim for the formula $= (t_1, t_2)$, $R(t_1, ..., t_r)$, $\top$ and $\perp$.

We will do this for $=(t_1, t_2)$, and leave the remaining cases as exercises.

$$\begin{aligned}
&\prod_{\mathcal{U}} M_s \vDash (=(t_1,t_2))[[\{a_{1,s}\}],...,[\{a_{n,s}\}]] \\
&\Leftrightarrow t_1[[\{a_{1,s}\}],...,[\{a_{n,s}\}]] = t_2[[\{a_{1,s}\}],...,[\{a_{n,s}\}]] \\
&\Leftrightarrow \{j \in J | t_1[\{a_{1,j}\},...,\{a_{n,j}\}] = t_2[\{a_{1,j}\},...,\{a_{n,j}\}]\} \in \mathcal{U} \\
&\Leftrightarrow \{j \in J | (=(t_1,t_2))[\{a_{1,j}\},...,\{a_{n,j}\}]\} \in \mathcal{U}
\end{aligned}$$

which completes this case.

We then need to show that if the claim holds for the formula A. B, then for any variable x it holds for

$$(A) \leftrightarrow (B), (A) \rightarrow (B), (A) \vee (B), (A) \wedge (B), \neg(A), \forall x(A), \exists x(A).$$

The first five hold immediately by the definition of an ultraproduct (exercise: check).

We now show it for the formula $\exists x(P)$. Note that the for all case can either be seen as similar, or via the equivalence of $\forall x Q$, and $\neg(\exists x(\neg Q))$.

$$\begin{aligned}
&\prod_{\mathcal{U}} M_s \vDash \exists x P[[\{a_{1,s}\}],...,[\{a_{n,s}\}]] \\
&\Leftrightarrow \exists [b_s] \in \prod_{\mathcal{U}} M_s \text{ such that } \prod_{\mathcal{U}} M_s \vDash \phi[[b_s],[\{a_{1,s}\}],...,[\{a_{n,s}\}]] \\
&\Leftrightarrow \{s \in S | M_s \vDash \phi[[b_s],[\{a_{1,s}\}],...,[\{a_{n,s}\}]]\} \in \mathcal{U} \\
&\Leftrightarrow \{s \in S | M_s \vDash \exists x \phi[[\{a_{1,s}\}],...,[\{a_{n,s}\}]]\} \in \mathcal{U}
\end{aligned}$$

□

4.5.4. *A Categorical Interpretation of Ultraproducts.* Let $\mathcal{U}$ be an ultraproduct on a set S. We can consider the ordered set $\mathcal{U}$ as a category (which we will denote $\mathcal{J}_{\mathcal{U}}$), where objects are elements $S \in \mathcal{U}$, and we have a single morphisms $S_1 \rightarrow S_2$ if $S_1 \subset S_2$ (including the case of equality). As there is at most one morphism between any two objects composition is uniquely defined.

Exercise 4.40. The category $\mathcal{J}_{\mathcal{U}}^{op}$ is a filtered category in the sense of definition 6.14.

Remark 4.41. We can interpret the Łoś ultraproduct theorem as saying that $\prod_{\mathcal{U}} M$ is a cocone over the diagram indexed by $\mathcal{J}_{\mathcal{U}}^{op}$ given by the models $\prod_{s \in S} M_s$, to any $S \in \mathcal{U}$. We map (the opposite of) each inclusion $S_1 \subset S_2$ to the projection

$$\prod_{s \in S_2} M_s \rightarrow \prod_{s \in S_1} M_s.$$

Note that this is an elementary morphism.

The Łoś ultraproduct theorem tells us that for any $S_1 \in \mathcal{U}$ there is an elementary morphism

$$\prod_{s \in S_1} M_s \rightarrow \prod_{\mathcal{U}} \mathcal{M}_s$$

this is because the surjective map $\prod_S M_s \rightarrow \prod_{\mathcal{U}} M_s$ factors through $\prod_{S_1} M_s$, providing the required map

$$\begin{tikzcd}
\prod_S M_s \arrow[rr] \arrow[dr] & & \prod_{\mathcal{U}} M_s \\
& \prod_{S_1} M_s \arrow[ur] &
\end{tikzcd}$$

We claim that in fact $\prod_{\mathcal{U}} M_s$ is universal among cocones, and hence is the filtered colimit of the diagram $\mathcal{J}_{\mathcal{U}}^{op} \rightarrow Mod$, where Mod is the category of models for a given first order theory, with morphisms given by elementary morphisms.

Proposition 4.42. *The ultraproduct $\prod_{\mathcal{U}} M_s$ is the colimit of the diagram $\mathcal{J}_{\mathcal{U}}^{op} \rightarrow Mod$ described above.*

Proof. We have already shown that $\prod_{\mathcal{U}} M_s$ is a cocone.

Let C' be another cocone. We claim that the map $\prod_S M_s \to C'$ factors through $\prod_{\mathcal{U}} M_s$.

Note that the map $\prod_S M_s \to \prod_{\mathcal{U}} M_s$ is a surjection. We then have that by the defining property of cocones, the map $\prod_S M_s \to C'$ factors through $\prod_{\mathcal{U}} M_s$. By definition of this factoring this commutes with all maps from elements in the image of $\mathcal{J}_{\mathcal{U}}^{op}$. $\square$

Exercise 4.43. Consider the category $\mathcal{J}$ corresponding to the ordered set of the natural numbers (standard ordering). Show that given a diagram $\mathcal{J} \to Mod$, where Mod is the category of elementary models of some theory, such that each morphism in the image is an injection on sets, this diagram has a colimit.

Hint: Construct this colimit explicitly.

4.5.5. *The Compactness Theorem in First Order Logic.* We now have a constructive proof of the below theorem:

Theorem 4.44 (Compactness Theorem in first order logic)**.** *Let (L, Γ) be a first order theory. Every finite subset Γ' of Γ has a model $M_{\Gamma'}$, if and only if Γ has a model.*

Proof. Firstly note that the only if direction is clear. If M is a model of Γ it is a model of Γ' for every finite subset $\Gamma' \subset \Gamma$.

For each $g \in G$ there is a model of (L, g) which we denote M_g. Define $S_g \subset G$ as $S_g := \{h \in G | g \subset h\}$. Consider $X := \{S_g | g \in G\} \subset \mathcal{P}(G)$. The set X satisfies the finite intersection property of proposition 3.78 and hence is contained inside an ultrafilter $\mathcal{U}$ on G.

Consider the model of $(L, \varnothing)$ given by

$$M_\Gamma := \prod_{\mathcal{U}} M_g.$$

We claim that this is a model of (L, Γ). This is because for any $\psi \in \Gamma$, we have that $S_{\{\psi\}} \subset \{g \in G | M_g \vDash \psi\}$. Hence $\{g \in G | M_g \vDash \psi\} \in \mathcal{U}$. Hence $M_\Gamma \vDash \psi$ by Łoś ultraproduct theorem (4.39). $\square$

Remark 4.45. If you are familiar with topology it is natural to ask why this is called the compactness theorem. Consider the set $\mathcal{M}$ of all models of L, modulo elementary equivalence[12]. We generate a topology on $\mathcal{M}$ by the open sets $U_\psi := \{M \in \mathcal{M} | M \vDash \psi\}$ for all $\psi \in Form(L)$. The compactness theorem is then equivalent to saying that $\tilde{M}$ is compact with respect to this topology (Exercise).

(Todo: Break up into some subexercises?)

4.6. **Applications of Compactness.**

4.6.1. *Non-standard Arithmetic.* Recall the example of Peano Arithmetic (as a first order theory) in example 4.10.

Add a constant c.

Consider the set of axioms $\{\neg(c = m) | m \in \mathbb{N}\}$ (you can write m by iterated application of the successor function).

Clearly any finite subset of these axioms is satisfiable. Hence by the compactness theorem we have that all of them are satisfiable.

Hence there is a model of the first order[13] Peano axioms which includes a constant c that is not one of the “normal" natural numbers.

[12]Ie. there is an equivalence relation on models, defined by saying that $M_1 \sim M_2$ if there exist inverse maps $f : M_1 \to M_2$ and $g : M_2 \to M_1$ that are elementary maps of models. We take the equivalence classes of this relation.

[13]This condition is very important. There is not a version of the compactness theorem for second order logic, and there are some versions of second order arithmetic, which can specify $\mathbb{N}$.

4.6.2. *Non-standard Analysis.* Recall the first order theory of a real field from example 4.18, removing (optional) the axiom about the existence of roots of a polynomial.

We again add a constant c to the language. We consider the set of axioms $\{c > 0\} \cup \{c \leq 1/n\}_{n=1}^{\infty}$ (we should be slightly careful about defining the second, as this is equivalent to $c \times n < 1$ this is not difficult). By applying the compactness theorem we see that there is a model of the axioms of real closed fields which has an infinitesimal c.

We can also use the structure of the proof of the compactness theorem to give a more explicit version of this. Let $\mathcal{U}$ be a non-principal ultrafilter on the natural numbers.

Consider $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$.

Exercise 4.46. Show that the diagonal embedding

$$\mathbb{R} \xrightarrow{i} \mathbb{R}^{\mathbb{N}}/\mathcal{U}$$

is only an isomorphism if $\mathcal{U}$ is a principal ultrafilter.

Definition 4.47. In light of this we pick a non-principal ultrafilter $\mathcal{U}$ which we fix for the rest of this discussion. We denote $\mathbb{R}^* := \mathbb{R}^{\mathbb{N}}/\mathcal{U}$, and call this the *hyperreals*.

Exercise 4.48. Show that $[(1, 2, 3, ...)] \in \mathbb{R}^{\mathbb{N}}/\mathcal{U}$ is greater than $i(r)$ for any $r \in \mathbb{R}$.

Provide an example of a positive element in $\mathbb{R}*$, which is less than $i(r)$ any positive element $r \in \mathbb{R}$.

One of the advantages of this more explicit version, is that for any function $f : \mathbb{R} \to \mathbb{R}$ we automatically have a function $\mathbb{R}^* \to \mathbb{R}^*$.

Similarly for a subset $A \subset \mathbb{R}$ we get a natural subset $A^* \subset \mathbb{R}^*$. Also for $\mathbb{R}$ a model of the first order theory of real closed fields, we have that $\mathbb{R}^*$ is also a model the first order theory of real closed fields.

Exercise 4.49. Show that for $A \subset \mathbb{R}$, $A = A^*$ if and only if A is finite.

Definition 4.50. The *infinitesimal hyperreals* are the elements of the set $\mu := \{x \in \mathbb{R}^* | |x| \leq 1/n \ \ \forall n \in \mathbb{N}\}$.

Definition 4.51. We define the equivalence relation $\approx$ by $x \approx y$ if and only if $|x - y| \in \mu$.

Exercise 4.52. Confirm that this is an equivalence relation.

Proposition 4.53. *Let $\{s_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers. This is a morphism $\mathbb{N} \to \mathbb{R}$, and hence extends to a morphism $\mathbb{N}^* := \mathbb{N}^{\mathbb{N}}/\mathcal{U} \to \mathbb{R}^*$.*

We have that $s_n \to L$ if and only if $s_N \cong L$ for all infinite[14] *$N \in \mathbb{N}^*$.*

Exercise 4.54. Prove the above proposition.

Similarly we can formulate non-standard criteria for many other notion in analysis.

Proposition 4.55. *Let $f : A \to \mathbb{R}$, $a \in A$. Then f is continuous at a if and only if for all $x \approx a$ we have $f(x) \approx f(a)$.*

Exercise 4.56. Prove proposition 4.55.

Finally we can give a non-standard definition of the derivative of a function f.

Proposition 4.57. *Let $f : A \to \mathbb{R}$, and $a \in A$ (in fact in the interior of A), then we have that f is differentiable at a with derivative $d \in \mathbb{R}$ if for all hyperreals, $\xi \approx a$, $\xi \neq a$ we have*

$$\frac{f(\xi) - f(a)}{\xi - a} \approx d.$$

Proof. (Todo: This should be written.) □

(Todo: write an exercise on Riemann integration in the non-standard paradigm)

[14]That is to say, $N \in \mathbb{N}^*$ such that for all $n \in \mathbb{N}$ we have $N > n$.

4.6.3. *Upward Löwenheim–Skolem Theorem.*

Definition 4.58. Let L be a first order language. We call the cardinal $|\sigma(L)| := |Function(L)| \oplus |Relations(L)|$ the *signature* of L.

Theorem 4.59 (Löwenheim–Skolem Theorems)**.** *Let M be a model of (L, Γ). Let $\kappa \geq |\sigma(L)|$. There exists a model N of (L, Γ) such that either:*

- *$\kappa > |M|$, and there is an elementary embedding $M \to N$.*
- *$\kappa \leq |M|$, and there is an elementary embedding $N \to M$.*

In this section we only deal with the second case (sometimes called the downward Löwenheim–Skolem theorem). See corollary 4.88 and proposition 4.90 for a proof of the first part.

Proof of Theorem 4.59 for $\kappa > |M|$,assuming the $\kappa \leq |M|$ part. For every $m \in M$ we add a constant c_m to our theory. We also add κ new symbols c_a for $(a \in \kappa)$, and we add axioms $c_1 \neq c_2$ for these new symbols.

We now need to also add axioms that for any function of rank n, $c_{f(m_1,...,f(m_n))} = f(c_{m_1}, ..., c_{m_n})$, and for any relation $R(c_{m_1}, ...c_{m_n}) = R(m_1, ..., m_n)$.

By the compactness theorem it is clear that this has a model.

By applying the downward part of this theorem we have that there is a model of cardinality κ.

It is clear that we have an elementary embedding by construction. $\square$

4.6.4. *Ax-Gröthendieck Theorem via Exercises.*

Exercise 4.60. Write down an countable set of axioms, that together state a field is of characteristic zero. (Recall that in an earlier exercise you were asked to write down a first order theory of fields).

Theorem 4.61 (Steinitz Theorem)**.** *Given two algebraically closed fields of the same characteristic, and the same cardinality are isomorphic.*

Exercise 4.62. Prove Theorem 4.61; Hint show the transcendence degree over some field k of the relevant characteristic is the same for both, and use and the axiom of choice to give an isomorphism.

Definition 4.63. We call a theory (L, Γ) complete, if for every sentence (formula with no free variables) ψ, either $\Gamma \vDash \psi$ or $\Gamma \vDash \neg\psi$.

Theorem 4.64 (Vaught's Test)**.** *If for each cardinal infinite cardinal $\kappa \geq max(\sigma(L), \aleph_0)$ (see 4.58) there is a unique model of a theory (L, Γ), and (L, Γ) has no finite models, then (L, Γ) is complete.*

Exercise 4.65. Prove Theorem 4.64 (Vaught's test) using the Löwenheim–Skolem theorems (4.59).

Exercise 4.66. Show that the theory of algebraically closed fields of a given characteristic has no finite models.

Exercise 4.67. Use the Steinitz Theorem (4.61) to show that the theory of algebraically closed fields (of a given characteristic) is complete.

Remark 4.68. Those of you who are reading ahead may be concerned that this at first glance looks like it might contradict Gödel incompleteness theorem. The resolution is as follows; we can not encode the integers inside an algebraically closed field of a fixed characteristic.

You may still be concerned; why can we not encode the integers in an algebraically closed field of characteristic zero?

The answer is that the integers are *not definable* in this theory (see definition 4.26).

Exercise 4.69. Use the fact that the theory of algebraically closed fields of a given characteristic (0 or p) is complete to prove that the following are equivalent:

- The theory of algebraically closed fields of characteristic zero models a sentence ψ.
- The theory of algebraically closed fields of characteristic p models the sentence ψ for all but finitely many primes p.

Proposition 4.70 (Simple case of Ax–Gröthendieck Theorem)**.** *Let k be an algebraically closed field Let $P : k^n \to k^n$ be a polynomial map. Then if P is injective it is surjective.*

Exercise 4.71. Specify a set of first order sentences/formulas (in the theory of fields) that expresses proposition 4.70.

Definition 4.72. An *algebraic set* over a field k is the zero locus of some set of polynomials $Q_1, ..., Q_n$ in $V = k^n$ (for k a field, and n a positive integer).

Theorem 4.73 (Ax–Gröthendieck Theorem)**.** *Let V be an algebraic set over an algebraically closed field. Then for any polynomial map*

$$P : V \to V$$

if P is injective, then P is surjective.

Exercise 4.74. Specify a set of first order sentences/formulas (in the theory of fields) that expresses proposition 4.73. Note that this is a generalization of exercise 4.71.

Exercise 4.75. Show that the sentences you wrote in exercise 4.74 are satisfied in the finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ of characteristic p.

Argue that this means they are satisfied in the algebraic closure $\overline{\mathbb{F}_p}$ of $\mathbb{F}_p$. Hint: Either use Hilbert Nullstellensatz, or directly reduce to a finite subfield of $\overline{\mathbb{F}_p}$ and use the same argument as used for $\mathbb{F}_p$.

Conclude that they are satisfied in a field of characteristic zero, and hence in all fields of characteristic zero.

Remark 4.76. The version of Ax–Gröthendieck Theorem due to Grothendieck (found in EGA IV) is in fact more general that the statement we proved by model theory above.

I am now aware of whether or not it is possible to recover the more general statement via model theory.

Remark 4.77 (Decidability). We call a theory (L, Γ) decidable, if for any sentence χ there is an algorithm that will decide whether or not $\Gamma \vDash \phi$. We are going to have to be very sloppy in this remark, because we have not formally formalized the concept of an algorithm – so we are going to have to work intuitively.

Note that the theory of algebraically closed fields of a fixed characteristic is recursively axiomatizable, that is to say there is an algorithm that lists the axioms of the theory.

Hence the following algorithm sketch shows that the theory of algebraically closed fields of a fixed characteristic. We can use the recursive axiomatization to recursively list all ψ such that ψ is implied by the axioms. Our program terminates when we reach either χ or $\neg\chi$. Because the theory is complete we have that the algorithm will terminate.

4.7. **Syntactic Proof in First Order Logic.** Recall our definition of a proof in propositional logic (definition 3.14). Here we had an axiom scheme given by definition 3.12 (where we again play the sleight of hand where we only used the logic symbols $\to$, $\perp$).

Firstly we should note that $\exists x(P)$ should be considered as equivalent to $\neg\forall x(\neg P)$. Hence we can use only the logical symbols $\forall$, $\to$, and $\perp$.

We take the axioms (axiom scheme) for propositional logic [see definition 3.12] (where we now allow the formula to be formula of our first order language). We then add the additional axioms (if we are doing first order logic without equality then we can remove the final three axioms):

Definition 4.78 (Quantifier and Equality Axioms)**.**

- If x is a variable not free in φ, we then have

$$(\forall x(\varphi \to \psi)) \to (\varphi \to (\forall x(\psi))).$$

- If $\psi = \varphi[t]$, in the sense that we have replaced all free occurences of a variable x in ψ by a term t, then we have:
$$((\forall x)\varphi) \to \psi.$$
- $$x = x$$
.
- $$(x = y) \to (t(z_0, .., z_{i-1}, x, z_{i+1} ..., z_n) = t(z_0, .., z_{i-1}, y, z_{i+1} ..., z_n))$$
- If φ is an atomic formula[15]
$$(x = y) \to (\varphi(z_0, .., z_{i-1}, x, z_{i+1} ..., z_n) = \varphi(z_0, .., z_{i-1}, y, z_{i+1} ..., z_n))$$

For propositional logic we had a single rule of inference given by Modus Ponens. For first order logic we add the rule of inference:

- Generalization: From φ infer $(\forall x)\varphi$.

We can now define

Definition 4.79 (Syntactic Proof in First Order Logic)**.** In first order logic a syntactic proof $\Gamma \vdash \varphi$ is a finite sequence of sentences $S_1, ..., S_n$ $(n \in \mathbb{N})$ of our language, with the properties that $S_n = \varphi$, and that for each $1 \leq k \leq n$ either:

- $S_k \in \Gamma$.
- S_k is one of the axioms listed in definition 3.12, or in definition 4.78.
- $\exists\ i, j < k$ with the properties that S_j is $S_i \to S_k$. (Modus Ponens)
- S_k is $\forall x(\varphi)$, where S_i is φ for some $i < k$.

Remark 4.80. Note that the Deduction Theorem (3.18) still holds for the same reason as in the propositional logic case.

Similarly the properties of Theorem 3.21 also hold true.

Theorem 4.81 (Soundness Theorem for First Order Logic)**.** *If* $\Gamma \vdash \varphi$ *then* $\Gamma \vDash \varphi$.

Proof. Essentially the same as the propositional case. There are some additional details that need to be checked. ☐

Definition 4.82 (Consistent, Inconsistent)**.** We define a set $\Gamma \subset Form(L)$ to be consistent or inconsistent in the same way as in propositional logic.

That is to say if $\Gamma \vdash \perp$ we call it inconsistent, and otherwise we call it consistent.

Lemma 4.83. *Every set of formulas* Γ *is contained in a*[16] *maximal consistent set of formula* Γ'.

Proof. The proof in the propositional case works (via Zorn's lemma in proposition 3.32) essentially unchanged. ☐

4.8. **Henkin Models and Completeness.** In addition to what we did for the Gödel completeness proof propositional case, we need to actually construct a set M. We will follow Henkin's proof of Gödel completeness.

Definition 4.84 (Witnesses)**.** We say $\Gamma \subset Form(L)$ has *witnesses* if for all $\varphi(x)$ (ie. a formula with variable x) we have that there exists a constant c_φ such that
$$\Gamma \vdash (\exists x(\varphi)) \to \varphi(c_\varphi)$$

[15] Recall that by an atomic formula we mean the equality formula, $\top$, $\perp$, and those corresponding to relationships of the language.

[16] Not necessarily unique.

Proposition 4.85. *Let L be a language and Γ be a consistent set of formula. Let $|C| = |Form(L)|$. Let $L' = L \cup C$ (where we consider C as constants in L'). We can form a set $\Gamma' \supset \Gamma$ of formula which is consistent, and has witnesses.*

Proof. Informally what we are going to try to do is add variables to act as witnesses together with axioms saying that these variables act as witnesses. The reason that this is not entirely straightforward is that we need not only to add witnesses for $\chi \in Form(L)$, but also for $\chi \in Form(L')$ – by adding witnesses we have increased the number of formula we must add witnesses for. Even worse the witnesses and axioms that they are witnesses could be inconsistent with Γ. We hence must give the following more careful procedure:

Pick a set C with $|C| = |L| \oplus \aleph_0$. Let β_C be the corresponding ordinal[17].

Using the axiom of choice we provide a well ordering/enumeration all formula in $Form(L')$ by non-zero ordinals $\alpha < \beta_C$. Similarly we impose a well ordering/enumeration of C by $\alpha < \beta_C$

For each ordinal $\zeta < \alpha_C$ we define T_ζ inductively below. For each ζ we define d_ζ as the minimal element of C such that d_ζ does not appear in any formula of $T_{\zeta-1}$.

- $T_0 = \Gamma$.
- $T_{S(\zeta)} := T_\zeta \cup \{((\exists x \chi_{S(\zeta)}) \to \chi(d_{S(\zeta)}))\}$.
- For ζ a limit ordinal we define
$$T_\zeta = \cup_{\xi < \zeta} T_\xi.$$

To complete the proof we need to check that (a) this enumeration works (our definition of d_ζ does not cause us to "run out" of d_ζ's), and that T_ζ is consistent. Noting that $|C|$ is a limit ordinal, we realize that for every ζ, in T_ζ we have used (cardinality) less that $|C|$ elements of C (as we have added ζ formulas, each which only uses finitely many elements of C), and hence there does exist some d_ζ remaining.

We start by inductively proving that T_ζ is consistent.

By assumption T_0 is consistent.

As any syntactic proof uses only finitely many axioms, we have that for a limit ordinal ζ if $T_\zeta \vdash \perp$, then $T_{xi} \vdash \perp$ for some $\xi < \zeta$.

Now assume that T_ζ is consistent, we will show that $T_{S(\zeta)}$ is consistent. Assume the converse, then we must have that $T_\zeta \vdash \neg(((\exists x \chi_{S(\zeta)}) \to \chi(d_{S(\zeta)})))$. By the definition of $\to$ we then have that $T_\zeta \vdash \exists x \chi_{S(\zeta)}$, and $\neg\chi(d_{S(\zeta)})$. As d_ζ does not appear in T_ζ we then have that $\forall x(\neg\chi(x))$, that is to say $\neg\exists x\chi$. But this contradicts the consistency of T_ζ. This completes the induction. □

Proposition 4.86 (Lindenbaum's Lemma for First order Logic)**.** *Let Γ be a set of formula of L. Then Γ has a model if and only if it is consistent.*

Proof. Firstly we deal with the easy direction. If Γ is not consistent, then by the soundness theorem 4.81 Γ not have a model.

Let us now suppose that Γ is consistent. Firstly we note that by proposition 4.85, there is an extension of the language L', and a consistent set $\Gamma' \supset \Gamma$. Note that a model of (L', Γ') is also a model of (L, Γ).

We now note that there is a maximal consistent set $\tilde{\Gamma}$ with the property that $\Gamma' \subset \tilde{\Gamma} \subset Form(L')$.

Consider the set C of witnesses in (L', Γ'). We introduce the equivalence relation $\sim$ on C defined by $c_1 \sim c_2$ if and only if $\tilde{\Gamma} \vdash (c_1 = c_2)$.

Exercise: Use the maximality of $\tilde{\Gamma}$ to show that $\sim$ is an equivalence relation.

We now wish to give $C/\sim$ the structure of a model of $(L', \tilde{\Gamma})$.

Let $a \in Constants(L)$, then let c_a be a witness for the statement $\exists x(x = a)$. We define ϕ to map a to $[c_a]$. We now need to show this is well defined. If c_a, and c'_a are two witnesses we note that we then have $\tilde{\Gamma} \vdash c_a = a$ and $\tilde{\Gamma} \vdash c'_a = a$. We hence have by maximality and consistency of $\tilde{\Gamma}$ that $\tilde{\Gamma} \vdash c_a = c'_a$, and hence $[c_a] = [c'_a]$, and so ϕ is well defined.

[17]That is to say, we will treat this as an ordinal rather than as a cardinal.

Let $f \in Functions^n(L)$. For $c_1, ..., c_n \in C$ we need to define an element $\phi(f)([c_1], ..., [c_n])$. We define $\phi(f)([c_1], ..., [c_n])$ as the equivalence class of a witness to the statement $\exists x(x = f(c_1, ..., c_n))$. Exercise: Show that this equivalence class is well defined, and did not depend on a choice of the representative $c_1, ..., c_n$.

Let $R \in Relations^n(L)$. We define $\{[c_1], ..., [c_n]\} \in \phi(R)$ if and only if $R(c_1, ..., c_n) \in \tilde{\Gamma}$.

Exercise: Check that R did not depend on the choice of representatives of the conjugacy class.

Now that we have defined the maps ϕ, showing that $C/\sim$ is a model of L' (and hence of L), we now must show that $C/\sim \vDash \tilde{\Gamma}$. We do this by induction on formula.

Firstly note that for any terms t_1 and t_2 without free variables, we have that

$$(C/\sim) \vDash = (t_1, t_2)$$

if and only if $\tilde{\Gamma} \ni (t_1 = t_2)$ (as it contains statements specifying both of them are equal to some $[c]$).

Secondly for an atomic formula χ_{at} (a relationship, equation, $\top$, $\bot$) we have that

$$(C/\sim) \vDash \chi_{at}(t_1, ..., t_n)$$

if and only if $\tilde{\Gamma} \ni \chi_{at}(t_1, ..., t_n)$ by consistency, definition of $\sim$, and the relationship symbols.

We have that

$$(C/\sim) \vDash (\psi \to \chi)$$

if and only if $(C/\sim) \vDash \neg\psi$, and $(C/\sim) \vDash \chi$.

We now need to deal with the quantifiers. We work with the quantifier $\exists$ only.

If $(C/\sim) \vDash \exists x\chi$, then there exists $[c]$ such that $\chi([c])$, that is to say $\tilde{\Gamma} \vdash \chi(c)$ (by induction hypothesis). We have

$$\vdash \chi(c) \to (\exists x(\chi)),$$

and hence $\tilde{\Gamma} \vdash \exists x(\chi)$.

On the other hand if $\tilde{\Gamma} \vdash \exists x(\chi)$ then the existence of witnesses means that there is $c \in C$ such that $\tilde{\Gamma} \vdash \chi(c)$ (using Modus Ponens), and hence we have that $(C/\sim) \vDash \chi([c])$, that is to say $(C/\sim) \vDash \exists x\chi$.

This completes the induction, it shows that $(C/\sim) \vDash \tilde{\Gamma}$, and hence that $(C/\sim)$ is a model of (L, Γ). □

Remark 4.87. Models constructed out of constants of a theory with witnesses in a way essentially similar to the above are called *Henkin Models*.

We can now prove a weak version of the downward part of the Löwenheim–Skolem Theorem (4.59).

Corollary 4.88 (Weak version of Downward Löwenheim–Skolem Theorem). Let (L, Γ) be a consistent first order theory. Let $\kappa = |L| \oplus \aleph_0$, There exists a model N of (L, Γ) with $|N| \leq \kappa$.

Remark 4.89 (Skolem's Paradox). Originally corollary 4.88 was thought to be paradoxical for the following reason. Apply this to the ZFC. The standard construction of the real numbers (Dedekind cuts of rational numbers, which can in term be constructed by two group completions from $\mathbb{N}$, which is definable in set), gives a set of real numbers. However this is "countable" – but we can prove that the real numbers are countable within ZFC (Cantor's diagonal argument).

The resolution to this seeming paradox is as follows: We recall that the definition of uncountable is that there is no bijection $f : \mathbb{R} \to \mathbb{N}$. In the countable model of ZFC there is still no such bijection. It is only in the meta language that we can argue that this set is "countable" however this is a notion of countability in the meta language that does not agree with the notion of countability within ZFC.

Proof. Upon examination of the construction in the proof of 4.86 we note that this result follows. The key point is to note that $|Form(L)|$ (and hence the C of the construction in proposition 4.85) is of size equal to $|L| \oplus \aleph_0$. □

Proposition 4.90 (Downward Löwenheim–Skolem Theorems). *Let M be a model of (L, Γ). Let $\kappa \geq |\sigma|$, and $\kappa \leq |M|$. There exists a model N of (L, Γ) such that there is an elementary embedding $N \to M$.*

Proof. We introduce the following closure operator. Let $A \subset M$, $F(A) = A \cup \{b \in M\}$, where we add precisely one $b \in M$ such that $M \models \phi(b, a_1, ..., a_n)$for each $\phi(x, a_1, ..., a_n)$ such that $a_1, ..., a_n \in M$, x is a variable, and $M \models \exists x \phi(x, a_1, ..., a_n)$.

We pick $A \subset M$ with $|A| = \kappa$. We then define $N = \cup_{n \in \mathbb{N}} F^n(A)$.

Clearly this is an elementary embedding (this test for being an elementary morphism is called the Tarski–Vaught test). ☐

Theorem 4.91 (Gödel Completeness Theorem)**.** *Let* (L, Γ) *be a first order theory. Then*

$$\Gamma \models \varphi \Rightarrow \Gamma \vdash \varphi.$$

Proof. We proceed from the first order version of Lindenbaum's lemma (proposition 4.86) in exactly the same manner as in the propositional case:

Suppose that $\Gamma \models \phi$. Then there is no model of $\Gamma \cup \{\neg\phi\}$. Hence by proposition 3.32 we must have that $\Gamma \cup \{\neg\phi\}$ is inconsistent.

That is to say $\Gamma \cup \{\neg\phi\} \vdash \bot$. By the Deduction Theorem 3.18 we hence have that $\Gamma \vdash (\neg\phi) \to \bot$, that is to say $\Gamma \vdash ((\phi \to \bot) \to \bot)$. Applying the second axiom and modus ponens shows that $\Gamma \vdash \phi$. ☐

Remark 4.92. Again we have that soundness and completeness in first order logic give us

$$\Gamma \models \varphi \Leftrightarrow \Gamma \vdash \varphi.$$

4.9. **Types.**

Warning 4.93. *This use of the word types in this section, is* not *the same as the use in "type theory."*

Definition 4.94 (Type)**.** We work with a model M for a language L. We define an extension of the language L_A be adding a constant c_a for each $a \in A \subset M$ (for some subset $A \subset M$).

We define an *n-type of M over A* as a set of formulas $p(\vec{x})$ with free variables in the set $\vec{x} = (x_1, ..., x_n)$, with the property that for any finite subset $p_{fin}(\vec{x}) \subset p(\vec{x})$ there is a $\vec{b} \in M^n$ with the property that $M \models p_{fin}(\vec{b})$.

Definition 4.95. We say that a type $p(\vec{x})$ is *complete* if it is maximal in the poset (partially ordered set) of types ordered by inclusion.

Definition 4.96. We say that a type $\vec{p}(\vec{x})$ is *realized* in M if there is $\vec{b} \in M^n$ such that $M \models p(\vec{b})$.

Definition 4.97 (Local Definability)**.** We say that a type $p(\vec{x})$ is locally definable if there exists a formula ψ, such that $M \models \psi \to \chi$ for all $\chi \in p(\vec{x})$.

Exercise 4.98. Explain why if a type p of M is locally definable, then it is realized in M.

There is a natural *topological* interpretation of this condition.

Consider the set of all formulas in n variables $\vec{x}$. Consider the equivalence relation $\sim_{\mathcal{M}}$ defined by $\psi \sim_{\mathcal{M}} \chi$ if and only if

$$\mathcal{M} \models \forall \vec{x}(\psi(\vec{x}) \leftrightarrow \chi(\vec{x})).$$

Exercise 4.99. Show that the set of all formulas in n variables $\vec{x}$ quotiented by $\sim_{\mathcal{M}}$ is a boolean algebra.

Hint: This is similar to how we showed the Lindstöm–Tarski algebra is a Boolean algebra.

We can now apply Stone duality. We take the space of ultrafilters. Points of this space correspond to complete n-types of M. It is endowed with the topology that for any formula χ we have an open set U_χ corresponding to all types containing the formula χ. This makes it into a Stone space (compact, totally disconnected, Hausdorff).

A type is locally definable if and only if it is an isolated point (ie. both closed and open) in this topology. (Exercise: unravel the definitions to see why this is true).

Definition 4.100. We say that a model N of L_A omits a type $p(\vec{x})$, if there does not exist $\vec{b} \in N^n$ such that $N \vDash p(\vec{b})$.

Theorem 4.101 (Omitting Types Theorem)**.** *Let (L, Γ) be a first order theory, in which the language L is countable.*

If an n-type p (over $\varnothing$) is not locally definable, then there is a model N of (L, Γ) that omits $p(\vec{x})$

Proof. We will use a modification of the construction of the Henkin model in lemma 4.86. We base this off the proof in [4].

Firstly noting that every type is contained inside a complete type, we can assume without loss of generality that $p(\vec{x})$ is complete.

We let $C = \{c_n | n \in \mathbb{N}\}$ be a new countable set of constants. Let $L' := L \cup C$. Enumerate by $\{\varphi_i\}_{i \in \mathbb{N}}$ the set of sentences of L'.

We let $\{\vec{d_i}\}_{i \in \mathbb{N}}$ be an enumeration of all subsets of C of cardinality n.

We will construct a sequence $\Gamma = \Gamma_0 \subset \Gamma_1 \subset \Gamma_2 \subset ...$ such that $\cup_{s \in \mathbb{N}} \Gamma_s$ has witnesses, is complete, is consistent, and implies that the Henkin model omits $p(\vec{x})$.

We define $\Gamma_0 = \Gamma$, and $\Gamma_s := \Gamma \cup \{\theta_k\}_{k=1}^{s} = \Gamma_{s-1} \cup \{\theta_s\}$.

Assume we have constructed Γ_s, with the property that Γ_s is consistent.

Then

- In this case we deal with the completeness: If $s + 1 \cong 1 \pmod 3$ then:
  - If $\Gamma_s \cup \{\phi_{s/3}\}$ is consistent, we set $\theta_{s+1} := \theta_s \wedge \phi_{s/3}$.
  - Otherwise we set $\theta_{s+1} := \neg\phi_{s/3}$.
- In this case we ensure the resulting theory has witnesses. If $s + 1 \cong 2 \pmod 3$ then:
  - If $\phi_{(s-1)/3} = \exists x(\psi(x))$ and $\Gamma_s \vDash \phi_{(s-1)/3}$, then we set $\theta_{s+1} = \psi(c_\zeta)$, where ζ is the smallest number such that c_ζ has not yet appeared in any formula θ_k, ϕ_j ($k \leq s$, $j \leq (s-1)/3$).
  - Otherwise we set $\theta_{s+1} := \top$.
- In this step we deal with ensuring the Henkin model omits $p(\vec{x})$. In the case where $s - 2 \cong 0$ (mod 3) we do the following: From θ_s construct the formula ψ where for each constant $c \in C \backslash \{c_1, ..., c_n\}$ we replace each occurence of it in θ_s with a variable x, and then add $\exists x$ as a prefix to the formula.

  By assumption there exists a formula $\nu(\vec{x}) \in p(\vec{x})$ with the property that
  $$\Gamma \nvDash \forall \vec{x}(\psi(x) \to \nu(\vec{x})).$$
  We set $\theta_{s+1} := \neg\nu(\vec{d_{(s-2)/3}})$.

Note that in each state Γ_{s+1} is consistent.

Construct the Henkin model for the complete, consistent (Exercise) theory (with witnesses) $\cup_{s \in \mathbb{N}} \Gamma_s$. Clearly this omits the type $p(\vec{x})$. □

Exercise 4.102. Where was finiteness of L crucial in the above proof of the omitting types theorem 4.101. What happens if you try to generalize this to a transfinite induction?

Example 4.103. Consider the theory of algebraically closed fields of characteristic zero.

For each algebraic number we have the type consisting of all formula that hold true for the given algebraic number. These are locally definable (essentially by the equation defining the given algebraic number).

We also have a type corresponding to the formula that hold for a transcendental number (over the base field $\overline{\mathbb{Q}}$). Given that $\overline{\mathbb{Q}}$ is a model which omits this type, we know by the Omitting Types Theorem that this type is not locally definable[18].

[18] Can we see this directly? I haven't really thought about this.

4.10. **Definability.**

Remark 4.104. Versions of these results in propositional logic appeared as exercises in homework 4.

Definition 4.105. Let L_1, L_2 be two languages. Let $\Gamma_1 \subset Form(L_1)$ and $\Gamma_2 \subset Form(L_2)$. We call Γ_1 and Γ_2 *seperable* if there exists $\theta \in Form(L_1 \cap L_2)$, such that $\Gamma_1 \models \theta$, and $\Gamma_2 \models \neg\theta$. If Γ_1 and Γ_2 are not seperable we call them *inseperable*.

Theorem 4.106 (Craig Interpolation Theorem). *Let L_1, L_2 be two languages. Let $\chi \in Form(L_1)$, $\psi \in Form(L_2)$ such that $\chi \models \psi$ (in $L_1 \cup L_2$).*

There is then a formula $\theta \in L_1 \cap L_2$ (when we write $L_1 \cap L_2$ we assume there are sufficiently many variables in the intersection, add more if necessary), such that

$$\chi \models \theta \text{ and } \theta \models \psi$$

Proof. Firstly by considering only the symbols that appear in χ and ψ we can assume without loss of generality that $L_1 \backslash Constants(L_1)$ and $L_2 \backslash Constants(L_2)$ are finite.

We assume the contrary.

Let $C = \{c_n | n \in \mathbb{N}\}$ be a countable set of constants labeled by natural numbers.

Let $L_1' = L_1 \cup C$, $L_2' = L_2 \cup C$ and $L_{12}' = (L_1 \cap L_2) \cup C$.

We inductively define two sets

Set $T_0 = \{\chi\}$, $U_0 = \{\neg\psi\}$.

We apply the following simultaneous induction. We first enumerate all formulas in L_1' as χ_i for $i < \omega$, and we enumerate all formulas in L_2' as ψ_i for $i < \omega$. We do not use the $\forall$ symbol.

- If $T_m \cup \{\chi_m\}$ and U_m are inseparable, and $\chi_m = \exists x \sigma(x)$ we define $T_{m+1} := T_m \cup \{\chi_m\} \cup \{\sigma(c_\zeta)\}$ were ζ is the minimum ordinal such that c_ζ does not appear in any formula in $U_m \cup T_m$.
- If $T_m \cup \{\phi_m\}$ and U_m are inseparable, and χ_m if not of the form $\exists x \sigma(x)$ then we define $T_{m+1} := T_m \cup \{\chi_m\}$.
- Otherwise we define $T_{m+1} := T_m$
- If $U_m \cup \{\psi_m\}$ and T_{m+1} are inseparable and $\psi_m = \exists x \sigma(x)$ we define $U_{m+1} := U_m \cup \{\psi_m\} \cup \{\sigma(c_\mu)\}$ where μ is the minimum ordinal such that c_ζ does not appear in any formula in $U_m \cup T_{m+1}$.
- If If $U_m \cup \{\psi_m\}$ and T_{m+1} are inseparable and ψ_m is not of the form $\exists x \sigma(x)$ we define $U_{m+1} := U_m \cup \{\psi_m\}$.
- Otherwise we define $U_{m+1} := U_m$.

We define $U_\omega := \cup_{n \in \mathbb{N}} U_n$, and $T_\omega := \cup_{n \in \mathbb{N}} T_n$.

Exercise: Show that U_ω and T_ω are maximal consistent sets in L_2' and L_1' respectively.

Exercise: Show that U_ω and T_ω are inseperable.

Exercise: Show that $U_\omega \cap T_\omega$ is a maximal consistent set in L_{12}', and as such we have $U_\omega \cap Form(L_{12}') = T_\omega \cap Form(L_{12}')$.

Show that if we take the Henkin models M and N of (L_1', T_ω) and (L_2', U_ω) respectively as constructed in the course of the proof of proposition 4.86, there is an isomorphism of these models $M \to N$ considered as models of $(L_{12}', U_\omega \cap Form(L_{12}'))$.

Exercise: Show that this allows us to construct the structure of a $(L_1' \cup L_2', U_\omega \cup T_\omega)$-model on the set N.

This means that $\{\chi, \neg\psi\}$ is consistent by soundness. Contradiction, the result follows. $\square$

Theorem 4.107 (Robinson Consistency Theorem). *Let L_1, L_2 be two languages (such that $L_1 \cap L_2$ has enough variables). Suppose that $(L_1 \cap L_2, \Gamma_{12})$ is a complete theory, and suppose that (L_1, Γ_1) and (L_2, Γ_2), and $(L_1 \cap L_2, \Gamma_{12})$ are consistent theories, where $\Gamma_1 \supset \Gamma_{12} \subset \Gamma_2$.*

Then $(L_1 \cup L_2, \Gamma_1 \cup \Gamma_2)$ is consistent.

Proof. Suppose the contrary. Then by compactness there is some finite subset $\Gamma' \subset \Gamma_1 \cup \Gamma_2$ that is inconsistent. Let $\Gamma_1' := \Gamma' \cap \Gamma_1$, and let $\Gamma_2' = \Gamma' \cap \Gamma_2$.

Let $\chi = \wedge_{\xi \in \Gamma_1'} \xi$, and $\psi = \wedge_{\zeta \in \Gamma_2'} \zeta$.

We then have that as $\{\chi, \psi\}$ is inconsistent, by the deduction theorem (and completeness) we have that $\chi \models \neg\psi$.

By the Craig interpolation theorem (4.106) we have that there is $\nu \in Form(L_1 \cap L_2)$, with the property that $\chi \vdash \nu$, and $\nu \vdash \neg\psi$. We hence have that $\psi \vdash \neg\nu$. Now kby completeness of Γ_{12}, we have either that $\Gamma_{12} \models \nu$ or $\Gamma_{12} \vdash \neg\nu$. Either of these contradicts the consistency of both Γ_1 and Γ_2. □

Theorem 4.108 (Beth Definability Theorem)**.** *Let L be a language, and let $L' \supset L$ be an extension of the language L. Let P be a predicate of L'.*

The following are equivalent:

- *There is a set $\Sigma(P) \subset Form(L \cup \{P\})$, with the property that*
$$\Sigma(P) \cup \Sigma(P') \models \forall \vec{x}(P(\vec{x}) \leftrightarrow P'(\vec{x})),$$
where by $\Sigma(P')$ we mean the modification of the set $\Sigma(P)$ where for each $\psi \in \Sigma(P)$ we have replaced all occurrences of the symbol P with the symbol P'.
- *We have*
$$\Sigma(P) \models \forall \vec{x}(P(\vec{x}) \leftrightarrow \chi(\vec{x})),$$
where $\vec{x}$ ranges over the free variables, of $P(\vec{x})$ and χ, and $\chi \in Form(L)$.

Remark 4.109. Informally we can state the Beth definability theorem as (L, Γ) defines P explicitly if and only if it defines P implicitly.

Exercise 4.110. Show the easy direction of Theorem 4.108 (that is if $\Gamma \models \forall \vec{x}(\psi \leftrightarrow \chi)$, then we have the specified behaviour of models).

Proof of Beth Definability Theorem (4.108). The difficult direction remains.

Assume that there is a set $\Sigma(P) \subset Form(L \cup \{P\})$, with the property that
$$\Sigma(P) \cup \Sigma(P') \models \forall \vec{x}(P(\vec{x}) \leftrightarrow P'(\vec{x})),$$
where by $\Sigma(P')$ we mean the modification of the set $\Sigma(P)$ where for each $\psi \in \Sigma(P)$ we have replaced all occurrences of the symbol P with the symbol P'.

Firstly introduce new constants $c_1, ..., c_n$. We then have $\Sigma(P) \cup \Sigma(P') \models \forall P(\vec{c}) \leftrightarrow P'(\vec{c})$. We now use compactness to pick a finite subset $S \subset \Sigma(P) \cup \Sigma(P')$ such that $S \models P(\vec{c}) \leftrightarrow P'(\vec{c})$.

Let $\chi(P)$ (and $\chi(P')$) be obtained by taking the and of all formula in S where we replace all instances of the symbol P' with P (replace all instances of the formula P by P' respectively).

We have $\chi(P) \wedge \chi(P') \models P(\vec{c}) \to P'(\vec{c})$, by completeness and the deduction theorem we thus have
$$\chi(P) \wedge P(\vec{c}) \models \chi(P') \to P'(\vec{c}).$$

Applying the Craig interpolation theorem we have $\theta \in Form(L \cup \{c_1, ..., c_n\}$, such that
$$\chi(P) \wedge P(\vec{c}) \models \theta, \text{ and } \theta \models \chi(P') \to P'(\vec{c}).$$

Applying the deduction theorem and completeness again we thus have that
$$\chi(P) \vdash \theta \leftrightarrow P(\vec{c}).$$

Substitution and generalization gives the result (noting that $\Sigma(P) \models \chi(P)$):
$$\Sigma(P) \models \forall \vec{x}(P(\vec{x}) \leftrightarrow \chi(\vec{x})).$$
□

We state the following without proof:

Theorem 4.111 (Makkai Definability Theorem)**.** *Suppose that for each model M of a first order theory, we have a set $\tilde{M} \subset M^k$ such that*

- *For a elementary embedding $f : M_1 \to M_2$, we have $\tilde{M}_1 = f^{-1}(\tilde{M}_2)$.*
- *For models $\{M_j\}_{j\in J}$, and an ultrafilter $\mathcal{U}$ on J we have*

$$\widetilde{\prod_{\mathcal{U}} M_j} = \prod_{\mathcal{U}} \tilde{M}_j.$$

For a proof see [2] (I have also found the course notes [1] extremely helpful).

4.11. **Saturated Models.** (Todo: I should really have a section on this, but I probably won't)

4.12. **Quantifier Elimination.**

Definition 4.112. We say that a formula χ is quantifier free if and only if the symbols $\forall$ and $\exists$ do not appear in χ.

Definition 4.113 (Quantifier Elimination)**.** We say that a theory (L, Γ) has quantifier elimination if and only if it has the property that for every formula ψ, we have that $\Gamma \vDash \forall_{\vec{c}}(\exists x(\psi(x, \vec{v}) \to \chi(\vec{v})))$ for some quantifier free formula χ.

Example 4.114. An example of quantifier elimination, is that in the real numbers we can replace the sentence $\exists b(b^2 = c)$ with the sentence $c \geq 0$.

Example 4.115. Another example of quantifier elimination is $\exists x((x \in \mathbb{R}) \wedge (x^2 - bx - c = 0))$ is equivalent to the sentence $b^2 - 4c \geq 0$.

Exercise 4.116. What is the difference between quantifier elimination and witnesses?

Definition 4.117 (Model Complete)**.** We say that a theory is *model complete* if every injection[19] of models has the property that it is an elementary embedding.

Exercise 4.118. Show that if a theory has quantifier elimination then it is model complete.

Proposition 4.119. *Show that if a theory is model complete then it has quantifier elimination.*

We show this as a series of exercises.

Exercise 4.120. Firstly show that if either $\Gamma \vDash \psi$ or $\Gamma \vDash \neg\psi$ then there is a quantifier free χ such that $\Gamma \vDash \phi \leftrightarrow \psi$.

Exercise 4.121. We now consider the case where $\Gamma \nvDash \psi$, and $\Gamma \nvDash \neg\psi$.
Define $C(\vec{v}) = \{\chi(\vec{v}) | \chi \text{ quantifier free}, \Gamma \vDash \forall\vec{v}(\psi(v) \to \chi(\vec{v}))\}$.
Add $\vec{d}$ a set of new constants to L.
Show that if $\Gamma \cup C(\vec{d}) \vDash \psi(\vec{d})$, then proposition 4.119 follows.

We now show that $\Gamma \cup C(\vec{d}) \vDash \psi(\vec{d})$.
(Todo: Provide a proof that $\Gamma \cup C(\vec{d}) \vDash \psi(\vec{d})$ – or break it into several exercises. Include an interpretation of quantifier elimination in terms of definable sets. Prove that algebraically closed fields have quantifier elimination. However I don't think I will actually ever finish this section.)

5. Appendix A: Some Alternatives

We have talked about *classical* propositional logic, and first order logic. There are many variants on this idea. These modifications can either be based on philosophical differences about what truth/provability should mean, or on more practical concerns.

This section only exists to raise awareness of some alternatives. The statements we make about them are not only woefully incomplete, but also vague and imprecise.

A very incomplete list of examples:

[19]Strictly we mean an injection of the underlying sets.

- Intuitionistic/Constructive logic. This removes the law of the excluded middle. This was based on philosophical ideas about what it meant for something to exist, leading to Brouwer rejecting the idea that mathematical objects existed without a construction.

  However on a more practical note the Curry–Howard isomorphism says [extremely informally] that a proof of existence in constructive logic is the same as an algorithmic construction. This is a reason to consider this ever if you do not (philosophically) reject the idea of the excluded middle.

  A second reason to consider this within standard mathematics, is that there are mathematical structures (topoi) with an *interal logic*, which can be constructive.

  Kripke semantics are a semantic interpretation of truth that matches the constructive syntactic interpretation.
- Higher Order Logic. In first order logic we quantify over variables. In second order logic we also quantify over sets. In third order we quantify over sets of sets. This is often needed within standard mathematics[20].

  Exercise 5.1. Consider the axioms for a topology on a set X. In what order of logic does this most naturally fit?

- There are various forms of logic (modal logic, fuzzy logic, many valued logics,...) that don't have 0 and 1 as the only possible values of a truth function. There are various reasons and approaches to this.
- Relevance Logics. These want to only interpret $a \to b$ as true if "the truth of a is relevant to the truth of b." In particular this means that the principal of explosion no longer holds. This in particular allows paraconsistent logics, which allow things to both be true or false, however by the principal of explosion not holding this doesn't mean that everything is both true and false.

6. Appendix B: Category Theory

Definition 6.1. A *category* $\mathcal{C}$ consists of a collection[21] of objects $Ob(\mathcal{C})$, together with for every pair of objects A, B a collection of morphisms (or arrows) which we denote $Hom(A.B)$. We require:

- For all objects A, B, C there is an operation (composition) $Hom(A, B) \times Hom(B, C) \to Hom(A, C)$, denoted $(f, g) \mapsto f \circ g$.
- Associativity of composition For all objects A, B, C, D and $f \in Hom(A, B)$, $g \in Hom(B, C)$, $h \in Hom(C, D)$ we have that:

$$f \circ (g \circ h) = (f \circ g) \circ h.$$

- Identity: For any object A there exists $1_A \in Hom(A, A)$, with the property that for an object B, and any morphisms $f \in Hom(A, B)$, $g \in Hom(B, A)$ we have that $f = f \circ 1_A$ and $g = 1_A \circ g$.

Example 6.2.
- The category of sets. Objects are sets[22], morphisms are functions between sets, composition is composition of functions, and for any set S, 1_S is the identity function $S \to S$.
- The category $Vect_k$ of vector spaces over some field k. Objects are vector spaces over k. Morphisms are k-linear maps. Composition and identity are as above. We can also restrict to the subcategory $Vect_k^{f,d}$ where we only allow finite dimensional vector spaces as objects.

[20] One might ask – can't we just interpret being a set, or a set of sets, or... as a property of variables, introduce relations for being an element of a set, and use first order logic. The answer is that there are different semantic interpretations of a formula in second order logic. Roughly speaking this restricts the actual subsets that one quantifies over, from those one is quantifying over in the standard semantics. We refer elsewhere for details.

[21] I say a collection rather than a set because some categories (such as the category of sets) the collection of objects is too large to be a set. For other certain collections of morphisms are also too large to form a set. There are various approaches to dealing with this problem – here we take the approach of ignoring it. Other approaches would be to use a version of set theory that allows the relevant classes. Another is to work with a small model of your favourite logical theory of sets.

[22] Say the class of sets.

- The category of smooth (e.g. C^ω) manifolds. Objects are manifolds, morphisms are smooth maps.
- The category of groups, morphisms are group homomorphisms. Similarly the category of Rings, Fields.
- The category $Bool$ of Boolean algebras, morphisms are morphisms of Boolean algebras.
- The category of partially ordered sets. Objects are sets equipped with a partial order. Morphisms are morphisms of sets that preserve this order, that is to say morphisms f such that $a \leq b$ implies that $f(a) \leq f(b)$.
- The category $Lat^{c,d}$ of complemented, distributive lattices. This is a subcategory of the category of partially ordered sets, where we only take objects that are complemented, distributive lattices, and morphisms between these objects that preserve the notion of the supremum and infimum of pairs of points in the lattice.
- The category $Prop$ where objects are theories in propositional logic, and morphisms are interpretations (see definition 3.83).

Definition 6.3. Let $\mathcal{C}$ be a category. We define the *opposite category* $\mathcal{C}^{op}$ by setting $Ob(\mathcal{C}^{op}) := Ob(\mathcal{C})$, and setting $Hom_{\mathcal{C}^{op}}(A, B) := Hom_{\mathcal{C}}(B, A)$.

Intuitively we can see forming the opposite category as "reversing the direction of all morphisms/arrows of $\mathcal{C}$."

We now need to define the notion of a map between categories, which we call a functor.

Definition 6.4. Let $\mathcal{C}, \mathcal{D}$ be categories. A *functor* $F : \mathcal{C} \to \mathcal{D}$ is the data of a map $F : Ob(\mathcal{C}) \to Ob(\mathcal{D})$, together with for each pair of objects $A, B \in Ob(\mathcal{C})$ a map $Hom_{\mathcal{C}}(A, B) \xrightarrow{F} Hom_{\mathcal{D}}(f(A), f(B))$. These maps must be compatible with the identities, and composition as follows:

- For all $A \in ob(\mathcal{C})$ we must have that $F(1_A) = 1_{F(A)}$.
- For all $A, B, C \in Ob(\mathcal{C})$ the following diagram must commute:

$$\begin{tikzcd}
Hom_{\mathcal{C}}(A,B) \times Hom_{\mathcal{D}}(B,C) \arrow[r, "\circ_{\mathcal{C}}"] \arrow[d, "F \times F"] & Hom_{\mathcal{C}}(A,C) \arrow[d, "F"] \\
Hom_{\mathcal{D}}(F(A),F(B)) \times Hom_{\mathcal{D}}(F(B),F(C)) \arrow[r, "\circ_{\mathcal{D}}"] & Hom_{\mathcal{D}}(F(A),F(C)),
\end{tikzcd}$$

that is to say we must have $F \circ (\circ_{\mathcal{C}}) = (\circ_{\mathcal{D}}) \circ (F \times F)$ as functions $Hom_{\mathcal{C}}(A, B) \times Hom_{\mathcal{D}}(B, C) \to Hom_{\mathcal{D}}(F(A), F(C))$.

It is clear that we can compose functors. Note that for every category $\mathcal{C}$ there is an identity functor $Id_{\mathcal{C}} : \mathcal{C} \to \mathcal{C}$.

Exercise 6.5. Return to proposition 3.56. Show that there are in fact functors $F : Bool \to Lat^{c,d}$, and $G : Lat^{c,d} \to Bool$, enriching the map on objects given in the proposition, such that $F \circ G = Id_{Bool}$, and $G \circ F = Id_{Lat^{c,d}}$.

Unlike the case with sets, there is also a notion of a morphism between two functors $F, G : \mathcal{C} \to \mathcal{D}$ which we call a *natural transformation.*

Definition 6.6. Let $F, G : \mathcal{C} \to \mathcal{D}$ be two functors. We define a natural transformation $\eta : F \Rightarrow G$ is the data of a morphism $F(A) \xrightarrow{\eta_A} G(A)$ for every $A \in Ob(\mathcal{C})$. We require that for any $A, B \in Ob(\mathcal{C})$, and any $f \in Hom_{\mathcal{C}}(A, B)$ the following diagram commutes in $\mathcal{D}$;

$$\begin{tikzcd}
F(A) \arrow[r, "\eta_A"] \arrow[d, "F(f)"] & G(A) \arrow[d, "G(f)"] \\
F(B) \arrow[r, "\eta_B"] & G(B),
\end{tikzcd}$$

that is to say we require $G(f) \circ_{\mathcal{D}} \eta_A = \eta_B \circ_{\mathcal{D}} F(f)$ in $Hom_{\mathcal{D}}(F(A), G(B))$.

Definition 6.7. An *equivalence of categories* is given by a pair of functors $F : \mathcal{C} \to \mathcal{D}$, $G : \mathcal{D} \to \mathcal{C}$ together with natural transformations $F \circ G \Rightarrow Id_{\mathcal{D}}$, and $Id_{\mathbb{C}} \Rightarrow G \circ F$

Example 6.8. Consider the three functors $Vect_k \to Vect_k$, which on objects are the identity, take V to its dual V^*, and take V to it's double dual V^{**}. Define what these functors are on morphisms.

Show that there is a natural transformation $Id \Rightarrow (-)^{**}$. Show that this gives an equivalence of categories.

Is there a natural transformation $Id \Rightarrow (-)^*$?

Definition 6.9. Let $\mathcal{J}$ be a category, and $\mathcal{J} \xrightarrow{i} \mathcal{C}$ a functor (which we will call a diagram). A *cone* of the diagram is an element $C \in \mathcal{C}$, together with for all $J \in \mathcal{J}$ a morphism $C \xrightarrow{f_J} i(J)$, and for any morphism $g \in Hom_{\mathcal{J}}(J_1, J_2)$, the diagram

$$\begin{tikzcd} & C \arrow[dl, "f_{J_1}"'] \arrow[dr, "f_{J_2}"] & \\ i(J_1) \arrow[rr, "i(g)"] & & i(J_2) \end{tikzcd}$$

commutes.

Definition 6.10. We call an element $C \in \mathcal{C}$ a *limit* of this diagram $\mathcal{J} \xrightarrow{i} \mathcal{C}$, if

- C is a cone of the diagram J.
- For any other cone C', there exists a unique morphism $C' \to C$, that commutes with the maps from C and C' to each element of J.

Definition 6.11. Let $\mathcal{J}$ be a category, and $\mathcal{J} \xrightarrow{i} \mathcal{C}$ a functor (which we will call a diagram). A *cocone* of the diagram is an element $C \in \mathcal{C}$, together with for all $J \in \mathcal{J}$ a morphism $i(J) \xrightarrow{f_J} C$, and for any morphism $g \in Hom_{\mathcal{J}}(J_1, J_2)$, the diagram

$$\begin{tikzcd} & C & \\ i(J_1) \arrow[ur, "f_{J_1}"] \arrow[rr, "i(g)"] & & i(J_2) \arrow[ul, "f_{J_2}"'] \end{tikzcd}$$

commutes.

Definition 6.12. We call an element $C \in \mathcal{C}$ a *colimit* of this diagram $\mathcal{J} \xrightarrow{i} \mathcal{C}$, if

- C is a cocone of the diagram J.
- For any other cocone C', there exists a unique morphism $C' \leftarrow C$, that commutes with the maps from each element of J to C and C'.

Example 6.13. Consider the category of sets, and the diagram where $\mathcal{J}$ is the category with two objects and no non-identity morphisms. Let $i : \mathcal{J} \to Set$ maps these to sets A, B.

The colimit is $A \coprod B$ – the disjoint union of these sets.

The limit is $A \times B$ – the product of these sets.

Definition 6.14. A *filtered category* is a non-empty category $\mathcal{C}$ with the property that for any $A, B \in Ob(\mathcal{C})$, there exists $C \in \mathcal{C}$, such that $Hom_{\mathcal{C}}(A, C) \neq \varnothing$, and $Hom_{\mathcal{C}}(B, C) \neq \varnothing$, and such that for any pair $f, g \in Hom_{\mathcal{C}}(A, B)$ there exists $D \in \mathcal{C}$, and $h \in Hom_{\mathcal{C}}(B, D)$, such that $h \circ f = h \circ g$.

Definition 6.15. A *filtered colimit* is a colimit over a diagram $\mathcal{J} \to \mathcal{C}$, where the category $\mathcal{C}$ is filtered in the sense of definition 6.14.

References

[1] J Lurie. Course notes for categorical logic. `https://www.math.ias.edu/~lurie/278x.html`, 2018.

[2] Michael Makkai. *Duality and definability in first order logic*, volume 503. American Mathematical Soc., 1993.

[3] Yu I Manin. *A course in mathematical logic for mathematicians*, volume 53. Springer Science & Business Media, 2 edition, 2009.

[4] David Marker. *Model theory: an introduction*, volume 217. Springer Science & Business Media, 2006.

[5] Reed Solomon. Ordinals and cardinals. `https://www2.math.uconn.edu/~solomon/math5026f18/OrdCard2.pdf`. Course Notes.