Logic

Luka Ilić

December 10, 2018

Contents

Preface

The following notes, are to be regarded as such – notes. They should contain most of what is written down in the logic lecture at Imperial College London (2018) by professor Evans. More likely then not, there will be a considerable amount of spelling errors (–please report everything to email down below or in Github comments–) that hopefully do not alter any important meaning. These notes will be constantly reread (by you the readers as well as myself) so I hope that at the end of the term most errors will be corrected so that anybody reading this will find good lecture notes for the exam.

At the moment I still have problems with my labels so if you need to jump to a reference, just click on it

Anybody willing to help me, can write me an email at

luka.ilic18@imperial.ac.uk

Help will only consist of being able to edit errors yourself. (So no need for any texing, except if you really want to.)

This is a project for my fellow students, so I hope it will be appreciated and used. I wish everybody reading this a lot of fun with the following content.

1 Propositional logic

1.1 Truth functions

Convention: In this course we write T for true and F for false.

Definition. 1.1. The alphabet of propositional logic consists of following symbols: *propositional variables* denoted (mostly) by $p, q, \dots$ or $p_1, p_2, \dots, q_1, q_2, \dots$ and the *connectives* $\wedge, \vee, \neg, \to, \leftrightarrow$.

Definition. 1.2. A *propositional formula* is a string of symbols obtained in the following way:

1. Any variable is a formula.
2. If ϕ and ψ are formulas then so are $(\phi \wedge \psi), (\phi \vee \psi), (\neg\phi), (\phi \to \psi), (\phi \leftrightarrow \psi)$.
3. Any formula is obtained in this way.

Definition. 1.3. A *truth function* of n variables is a function

$$f : \{T, F\}^n \to \{T, F\} \quad .$$

Exercise. How many functions are there for n variables?

Definition. 1.4. Suppose ϕ is a formula with variables $p_1, \dots, p_n$ then we obtain a truth function $F_\phi : \{T, F\}^n \to \{T, F\}$ whose value at $(x_1, \dots, x_n) \quad x_i \in \{T, F\}$ is the truth value of ϕ when p_i has value x_i. The function F_ϕ is the *truth function of* ϕ.

Remark. The truth tables for the connectives are the following:

p	$\neg p$
T	F
F	T

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

p	q	$p \leftrightarrow q$
T	T	T
T	F	F
F	T	F
F	F	T

.

Example. What is the truth function of

$$(((p \to q) \wedge (q \to (\neg p))) \to (\neg p)) \quad ?$$

Definition. 1.5. A propositional formula ϕ whose truth function F_ϕ is always true is called *tautology*. Say that formulas ϕ, ψ are *logically equivalent* (l.e.) if they have the same truth function.

Remark. ϕ, ψ are l.e. iff $(\phi \leftrightarrow \psi)$ is a tautology. Also, suppose that we got some formula ϕ with variables $p_1, \dots, p_n$ and $\phi_1, \dots, \phi_n$ are formulas with variables $q_1, \dots, q_r$. For each $i \leq n$ substitute ϕ in place of p_i in ϕ. Then the result is a formula ψ and if ϕ is a tautology, then so is ψ.

Proof. The first statement is easy. For the second remark that

$$F_\psi(q_1, \dots, q_r) = F_\phi(F_{\phi_1}(q_1, \dots, q_r), \dots, F_{\phi_n}(q_1, \dots, q_r))$$

by induction on the number of connectives in ϕ. □

Example.

1. $(p_1 \wedge (p_2 \wedge p_3))$ is l.e. to $((p_1 \wedge p_2) \wedge p_3)$,
2. same with $\vee$,
3. $(p_1 \vee (p_2 \wedge p_3))$ is l.e. to $((p_1 \vee p_2) \wedge (p_1 \vee p_3))$
4. similar the other way around.
5. etc.

Remark. Note that by the remark above, we can boost these equivalences by substituting formulas for the variables.

Definition. 1.6. Say that a set of connectives is *adequate* if for every $n \geq 1$, every truth function of n variables is the truth function of some formula which involves only connectives from the set and variables $p_1, \dots, p_n$.

Theorem. 1.7. *The set* $\{\neg, \wedge, \vee\}$ *is adequate.*

Proof. Let $G : \{T, F\}^n \to \{T, F\}$

1. $G(v) = F$ for all $v \in \{T, F\}$. Take ϕ to be $(p_1 \wedge (\neg p_1))$ then $G = F_\phi$
2. (*Disjunctive Normal Form* List the $v \in \{T, F\}^n$ with $G(v) = T$ as $v_1, \dots, v_r$. Write $v_i = (v_{i1}, \dots, v_{in})$ Define

   $$q_{ij} = \begin{cases} p_j & \text{if } v_{ij} = T \\ (\neg p_j) & \text{if } v_{ij} = F \end{cases}$$

   So q_{ij} has value T iff p_j has value v_{ij}. Let ψ_i be

   $$(q_{i1}, \dots, q_{in}),$$

   then $F_{\psi_i}(v) = T$ iff each q_{ij} has value T iff $v = v_i$.

   Let θ be $(\phi_1 \vee, \dots, \vee \phi_r)$. Then $F_\theta(v) = T$ iff $F_{\psi_i}(v) = T$ for some i which is equivalent to $v = v_i$ for some $i \leq r$. Thus $F_\theta(v) = T$ iff $G(v) = T$ i.e. $F_\theta = G$. As θ was constructed using only $\neg, \vee, \wedge$ the statement follows.

□

Definition. 1.8. A formula in the form as θ in the proof above (1.7) is said to be in *disjunctive normal form (dnf).*

Remark. Apart from the very intuitive and useful dnf, having a small adequate set at our disposal is useful for the following reason. It shortens induction proofs over the structure of formulas by a considerable amount, as the reader will surely experience in due time.

Corollary. 1.9. *Suppose* χ *is a formula which truth function is not always false. Then* χ *is l.e. to a formula in dnf.*

Proof. Take $G = F_\chi$ and apply the second case from the proof above. □

Example. For

$$\chi : \quad ((p_1 \to p_2) \to (\neg p_2))$$

the truth function $F_\chi(v)$ is true precisely when $v = \{T, F\}$ or $v = \{F, F\}$. Hence the dnf is:

$$((p_1 \wedge (\neg p_2)) \vee ((\neg p_1) \wedge (\neg p_2))).$$

Corollary. 1.10. *The following sets of connectives are adequate:*

1. $\neg, \vee$
2. $\neg, \wedge$
3. $\neg, \to$.

Proof.
1. By 1.7 we just need to show, that $\wedge$ can be expressed using $\neg, \vee$. $(p \wedge q)$ is l.e. to $(\neg((\neg p) \vee (\neg q))$.
2. similar to the approach above. $(p \vee q)$ is l.e. to $(\neg((\neg p) \wedge (\neg q))$.
3. Due to the cases above, it suffices to express either $\wedge$ or $\vee$ using $\neg, \to$. $(p \vee q)$ is l.e. to $((\neg p) \to q)$.

□

Example. Some sets of connectives that are not adequate are:

1. $\wedge, \vee$
2. $\neg, \leftrightarrow$

Proof.
1. If ϕ is build using $\wedge, \vee$ then $F_\phi(T, \ldots, T) = T$ as proven by induction over number of connectives.
2. exercise.

□

Example. The *NOR* connective $\downarrow$ has truth table:

p	q	$(p \downarrow q)$
T	T	F
T	F	F
F	T	F
F	F	T

It is adequate on its own.

(exercise - express $\neg, \wedge$)

1.2 A formal system for propositional logic

Idea: Try to generate all tautologies from certain basic assumptions (axioms) using appropriate deduction rules.

Definition. 1.11. This is important!

A *formal deduction system* Σ has the following ingredients:

1. An *alphabet* A of symbols ($A \neq \emptyset$).
2. A non empty set $\mathcal{J}$ of the set of all finite sequences ('strings') of the elements of A: the *formulas* of Σ.
3. A subset $\mathcal{A} \subseteq \mathcal{J}$ called the *axioms* of Σ.
4. A collection of *deduction rules.*

Definition. 1.12. A *proof* in Σ us a finite sequence of formulas in $\mathcal{J}$

$$\phi_1, \ldots, \phi_n$$

such that each ϕ_i is either an axiom *or* is obtained from $\phi_1, \ldots, \phi_{i-1}$ using one of the deduction rules. The last (or any) formula in a proof is a *theorem* of Σ. Write $\vdash_\Sigma \phi$ for 'ϕ is a theorem of Σ'.

Remark. 1. If $\phi \in \mathcal{A}$ then $\vdash_\Sigma \phi$.

2. We should have an algorithm to test whether a string of symbols really is a formula and whether it is an axiom. Then someone who is willing to follow an algorithm precisely (computer) should be able to generate all possible proofs in σ and check whether something is a proof. (We say Σ is *recursive* in this case.)

Definition. 1.13. The formal system L for propositional logic consists of:

- **Alphabet**: variables $p_1, p_2, p_3 \ldots$ connectives $\neg, \rightarrow$ punctuation),(.
- **Formulas**: as defined in 1.2 and will be called L*-formulas.*
- **Axioms**: Suppose ϕ, ψ, χ are L-formulas. The following are axioms of L:

  A1 $(\phi \rightarrow (\psi \rightarrow \phi))$

  A2 $((\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi)))$

  A3 : $(((\neg\psi) \rightarrow (\neg\phi)) \rightarrow (\phi \rightarrow \psi))$
- **deduction rule**: *Modus Ponens (MP)* from $\phi, (\phi \rightarrow \psi)$ deduce ψ.

Example. Suppose ϕ is an L-formula. Then $\vdash_L (\phi \rightarrow \phi)$. A proof in L could be as follows:

1. $(\phi \rightarrow ((\phi \rightarrow \phi) \rightarrow \phi))$ use A1
2. $(\chi \rightarrow (\phi \rightarrow ((\phi \rightarrow \phi) \rightarrow \phi)))$ use A1 and MP
3. $((\phi \rightarrow (\phi \rightarrow \phi)) \rightarrow (\phi \rightarrow \phi))$
4. $(\phi \rightarrow (\phi \rightarrow \phi))$

5. $(\phi \to \phi)$.

Definition. 1.14. Suppose Γ is a set of L- formulas. A deduction from Γ is a finite sequence of formulas of L-formulas $\phi_1, \ldots, \phi_n$ such that each one of these ϕ_i is either

- an axiom
- a formula in Γ or
- is obtained from previous formulas $\phi_1, \ldots, \phi_{i-1}$ using the deduction rule (MP).

Write $\Gamma \vdash_L \phi$ if there is a deduction from Γ ending in ϕ.

Say Γ is a *consequence* of Γ.

Remark. Being a consequence of the empty set ($\Gamma = \emptyset$) is the same as being a theorem of L. ($\emptyset \vdash_L \phi \Leftrightarrow \vdash_L \phi$)

Theorem. 1.15 (Deduction Theorem.)**.** *Suppose Γ is a set of L-formulas and ϕ, ψ are L-formulas. Suppose*

$$\Gamma \cup \{\phi\} \vdash \psi$$

then

$$\Gamma \vdash_L (\phi \to \psi)$$

.

Corollary. 1.16. *Suppose ϕ, ψ, χ are L-formulas such that $\vdash_L (\phi \to \psi)$ and $\vdash_L (\psi \to \chi)$. Then $\vdash_L (\phi \to \chi)$*

Proof. Use 1.15 with $\Gamma = \emptyset$: Show $\{\phi\} \vdash_L \chi$. Here is a deduction of *chi* from *phi*:

1. $(\phi \to \psi)$ (theorem of L)
2. $(\psi \to \chi)$ (theorem of L)
3. ϕ (assumption)
4. ψ (MP)
5. χ (MP).

Thus $\{\phi\} \vdash_L \chi$. By 1.15: $\emptyset \vdash_L (\phi \to \chi)$ i.e. $\vdash_L (\phi \to \chi)$. □

Lemma. 1.17. *Suppose ϕ, ψ are L-formulas. Then*

1. $\vdash_L ((\neg\psi) \to (\psi \to \phi))$.
2. $\{(\neg\psi), \psi\} \vdash_L \phi$.
3. $\vdash_L (((\neg\phi) \to \phi) \to \phi)$

Proof.

1. problem sheet 1.
2. by 1. and MP (twice)

3. Suppose χ is any formula. Then $\{(\neg\phi), ((\neg\phi) \to \phi)\} \vdash_L \chi$ (by MP and 2.) Let α be any axiom and let χ be $(\neg\alpha)$. Apply 1.15 to 3 to get:

$$\{((\neg\phi) \to \phi)\} \vdash_L ((\neg\phi) \to (\neg\alpha))$$

A3: $(((\neg\psi) \to (\neg\phi)) \to (\phi \to \psi))$ and MP generate:

$$\{((\neg\phi) \to \phi)\} \vdash_L (\alpha \to \phi) \quad .$$

Since α is an axiom, by MP

$$\{((\neg\phi) \to \phi)\} \vdash_L \phi$$

and the application of 1.15 gives us:

$$\vdash_L (((\neg\phi) \to \phi) \to \phi) \quad .$$

□

Proof of 1.15: Suppose $\Gamma \cup \{\phi\} \vdash_L \psi$ using a deduction of length n. Show by induction on n that $\Gamma \vdash_L (\phi \to \psi)$.

Base step: $n = 1$. In this case ϕ is either an axiom or in Γ or is *phi*. In the first two cases $\Gamma \vdash_L \phi$ (one line deduction!) Using the A1 axiom $(\psi \to (\phi \to \psi))$ and MP we obtain $\Gamma \vdash_L (\phi \to \psi)$. In the last case – that $\phi = \psi$ – we already know

$$\Gamma \vdash (\phi \to \phi) \text{ by (1.13.)}$$

induction step: In our deduction of ψ from $\Gamma \cup \{\phi\}$ either ψ is an axiom or ψ is obtained from earlier steps using MP. In the last case these are formulas $\chi, (\chi \to \psi)$ earlier in the deduction. in the first case we argue as in the base case to get $\Gamma \vdash_L (\phi \to \psi)$. Otherwise we use the inductive hypothesis to get

$$\Gamma \vdash_L (\phi \to \chi)$$

and

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

We have the A2 axiom

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

Using the two formulas we obtained, this axiom and MP twice we obtain $\Gamma \vdash_L (\phi \to \psi)$ as required, completing the induction step.

□

1.3 soundness and completeness of L

Theorem. 1.18 (soundness of L)**.** *Suppose ϕ is a theorem of L. Then ϕ is a tautology.*

Remark (notation). A (propositional) *valuation* v is an assignment of truth values to the propositional variables $p_1, p_2, \dots$. So $v(p_i) \in \{T, F\}$ (for $i \in \mathbb{N}$). Note that, using the truth table rules, this assigns a truth value $v(\phi) \in \{T, F\}$ to every L-formula ϕ.

Proof of 1.18: By the induction on the length of a proof of ϕ it is enough to show:

1. every axiom is a tautology;
2. MP preserves tautologies. I.e. if $\psi, (\psi \to \chi)$ are tautologies, then so is χ.

For 1. use truth tables or argue as follows:

A2 Suppose for 1. if there is a valuation v with

$$v(((\phi \to (\psi \to \chi)) \to ((\phi \to \psi) \to (\phi \to \chi)))) = F$$

Then

$$v((\phi \to (\psi \to \chi))) = T$$

and

$$v(((\phi \to \psi) \to (\psi \to \chi))) = F$$

$v((\phi \to \psi) = T$ and $v((\phi \to \chi)) = F$ So by the last equality $v(\phi) = T, v(\chi) = F$ and also $(\psi) = T$ which contradicts the first equation.

A1 exercise

A3 exercise

For 2. if v is a valuation and $v(\psi) = T + v(\psi \to \chi) = T$ then $v(\chi) = T$.

□

Theorem. 1.19 (Generalization of soundness)**.** *Suppose Γ is a set of formulas and ϕ is a formula with $\Gamma \vdash_L \phi$. Suppose v is a valuation with $v(\phi) = T$ for all $\phi \in \Gamma$. Then $v(\phi) = T$.*

Proof. exercise □

Theorem. 1.20 (Completeness Theorem for L.)**.** *Suppose ϕ is a tautology. Then $\vdash_L \phi$.*

Remark (steps in the proof).

1. If $v(\phi) = T$ for all valuations v – we want to show $\vdash_L \phi$.
2. Try to prove a generalization: Suppose that for every v with $v(\Gamma) = T$ (i.e. $v(\phi) = T \forall \phi \in \Gamma$) we have $v(\phi) = T$. Then $\Gamma \vdash_L \phi$.
3. Equivalently, if $\Gamma \nvdash_L \phi$ show there is a valuation v with $v(\Gamma) = T$ and $v(\phi) = F$.

Definition. 1.21. A set Γ of L- formulas is *consistent* if there is no L-formula ϕ with

$$\Gamma \vdash_L \phi \text{ and } \Gamma \vdash_L (\neg\phi) \quad .$$

Proposition. 1.22. *Suppose Γ is a consistent set of L-formulas and $\Gamma \nvdash_L \phi$. Then $\Gamma \cup \{(\neg\phi)\}$ is consistent.*

Proof. Suppose not. So there is some formula ψ with

$$\Gamma \cup \{(\neg\phi)\} \vdash_L \psi$$

and

$$\Gamma \cup \{(\neg\phi)\} \vdash_L (\neg\psi)$$

Apply 1.15 to the second line above, then

$$\Gamma \vdash_L ((\neg\phi) \to (\neg\psi)).$$

By A3 and MP we obtain

$$\Gamma \vdash_L (\psi \to \psi).$$

Then $\Gamma \cup \{(\neg\phi)\} \vdash_L \phi$. By 1.15:

$$\Gamma \vdash_L ((\neg\phi) \to (\neg\phi)$$

and by a result from above

$$\vdash_L (((\neg\phi) \to \phi) \to \phi)$$

So by this, $\Gamma \vdash_L ((\neg\phi) \to (\neg\phi)$ and Modus Ponens we obtain

$$\Gamma \vdash_L \phi.$$

This contradicts $\Gamma \nvdash_L \phi$. ☐

Proposition. 1.23 (Lindenbaum Lemma)**.** *Suppose Γ is a consistent set of L-formulas. Then there is a consistent set of formulas $\Gamma^* \supseteq \Gamma$ such that for every ϕ either*

$$\Gamma^* \vdash_L \phi$$

or

$$\Gamma^* \vdash_L (\neg\phi) \quad .$$

(sometimes say Γ^* is complete*.)*

Proof. The set of all L-formulas is *countable*, so we can list the L-formulas as $\phi_0, \phi_1, \ldots$. (Why countable? Alphabet is countable: $\neg, \to,), (, p_1, p_2, \ldots$. Formulas are finite sequences from this alphabet, hence only countably many.) Define inductively sets of L-formulas $\Gamma_0 \subseteq \Gamma_1 \subseteq \Gamma_2 \subseteq \ldots$, where

$$\Gamma_0 = \Gamma$$

and

$$\Gamma^* = \bigcup_{i \in \mathbb{N}} \Gamma_i \quad .$$

Suppose Γ_n has been defined. If $\Gamma_n \vdash_L \phi_n$ then let $\Gamma_{n+1} := \Gamma_n$. Otherwise, if $\Gamma_n \nvdash_L \phi_n$ then let $\Gamma_{n+1} := \Gamma_n \cup \{(\neg\phi_n)\}$. An easy induction using 1.22 shows that each Γ_i is consistent.

Claim:: Γ^* is consistent. If $\Gamma^* \vdash_L \phi$ and $\Gamma^* \vdash_L (\neg\phi)$ then as deductions are finite sequences of L -formulas

$$\Gamma_n \vdash_L \phi \text{ and } \Gamma_n \vdash_L (\neg\phi)$$

for some $n \in \mathbb{N}$: contradiction. Now to show that Γ^* is complete, let ϕ be any formula. So $\phi = \phi_n$ for some $n \in \mathbb{N}$. If $\Gamma^* \nvdash_L \phi$ then $\Gamma_n \nvdash_L \phi$. So by construction $\Gamma_{n+1} \vdash_L (\neg\phi)$. Therefore $\Gamma^* \vdash_L (\neg\phi)$ and since ϕ was arbitrary this proves the statement. □

Lemma. 1.24. *Let Γ^* be as above, then there is a valuation v such that for every L-formula ϕ*

$$v(\phi) = T \Leftrightarrow \Gamma^* \vdash_L \phi$$

Corollary. 1.25. *Suppose Δ is a set of L-formulas which is consistent and*

$$\Delta \nvdash_L \phi \quad .$$

Then there is a valuation v with

$$v(\Delta) = T \text{ and } v(\phi) = F \quad .$$

Proof. Let $\Gamma := \Delta \cup \{(\neg\phi)\}$. By 1.22, Γ is consistent. By the Lindenbaum lemma there is $\Gamma^* \supseteq \Gamma$ which is consistent and such that for every χ either $\Gamma^* \vdash_L \chi$ or $\Gamma^* \vdash_L \neg\chi$. By 1.24 there exists a valuation with $v(\Gamma^*) = T$. In particular $v(\Delta) = T$ and $v((\neg\phi)) = T$ thus $v(\phi) = F$. □

Theorem. 1.26 (completeness/Adequacy theorem for L)**.** *If $v(\phi) = T$ for every valuation v, then $\vdash_L \phi$.*

Proof. Suppose $\nvdash_L \phi$. then apply the corollary above with $\Delta = \emptyset$. (Why is this consistent? Soundness Theorem.) There is a valuation with $v(\phi) = F$. □

Proof of 1.24. Γ^* consistent set of L-formulas such that for every L-formula ϕ either $\Gamma^* \vdash_L \phi$ or $\Gamma^* \vdash_L \neg\phi$. Want valuation v with $v(\phi) = T$ for all $\phi \in \Gamma^*$. (i.e. $v(\phi) = T \Leftrightarrow \Gamma^* \vdash_L \phi$.) Note that for each variable p_i either $\Gamma^* \vdash_l p_i$ or $\Gamma^* \vdash_L (\neg p_i)$. So let v be the valuation with $v(p_i) = T \Leftrightarrow \Gamma^* \vdash_L p_i$. Now, prove by induction on the length of ϕ that $v(\phi) = T \Leftrightarrow \Gamma^* \vdash_L \phi$.

Base case: ϕ is just a propositional variable – this case holds by the definition of v.

inductive step: **Case 1:** ϕ is $(\neg\psi)$. "$\Rightarrow$": $v(\phi) = T$ then $v(\psi) = F$ and by the induction hypothesis $\Gamma^* \nvdash_L \psi$. Then by the completeness of Γ^*, $\Gamma^* \vdash_L (\neg\psi)$ i.e. $\Gamma^* \vdash_L \phi$. "$\Leftarrow$": Suppose $\Gamma^* \vdash_L \phi$. By consistency $\Gamma^* \nvdash_L \psi$. By the induction hypothesis $v(\psi) = F$. As v is a valuation we obtain $v(\phi) = T$ which concludes the first case.

Case 2: ϕ is $(\psi \to \chi)$. "$\Rightarrow$": Suppose $v(\phi) = F$. Then $v(psi) = T$ and $v(\chi) = F$. By the induction hypothesis $\Gamma^* \vdash_L \psi$ and $\Gamma^* \nvdash_L \chi$. If

$$\Gamma^* \vdash_L \phi$$

then using modus ponens and$\Gamma^* \vdash \psi$ we obtain

$$\Gamma^* \vdash_L \chi$$

which is a contradiction. So $\Gamma^* \nvdash_L \phi$. "$\Leftarrow$": Suppose $\Gamma^* \nvdash_L (\psi \to \chi)$. Then $\Gamma^* \nvdash \chi$ (due to A1). Also, due to a result from above, $\Gamma^* \nvdash_L (\neg\psi)$. By combining these results we obtain $v(\chi) = F$ and $v(\psi) = T$ therefore $v(\phi) = F$ which concludes the induction step.

□

Corollary. 1.27. *Suppose Δ is a set of L-formulas and ϕ is an L-formula. Then*

1. *Δ is consistent if and only if there is a valuation v with $v(\Delta) = T$.*
2. *$\Delta \vdash_L \phi$ if and only if for every valuation v with $v(\Delta) = T$*

we have $v(\phi) = T$.

Proof. Exercise – deduce these from the preliminaries to 1.26. Warning: in the second statement Δ is not necessarily consistent. □

Theorem. 1.28 (Compactness Theorem for L)**.** *Suppose Δ is a set of L-formulas. The following are equivalent:*

1. *There is a valuation v with $v(\Delta) = T$.*
2. *For every finite subset $\Delta_0 \subseteq \Delta$ there us a valuation w with $w(\Delta_0) = T$.*

Proof. By the above corollary, the first is true iff Δ is consistent. Similarly the second holds iff every finite subset is consistent. But if $\Delta \vdash_L \phi$ and $\Delta \vdash_L (\neg\phi)$ then as deductions are finite (+ therefore only involves finitely many L-formulas in Δ), for some finite $\Delta_0 \subseteq \Delta$, $\Delta_0 \vdash_L \phi$ and $\Delta_0 \vdash_L (\neg\phi)$. □

Exercise. Let P be the set of sequences of

2 Predicate Logic

also called first-order logic Plan :

1. introduce mathematical objects that this logic can reason about. I.e. *First-order structures*
2. introduce the formulas. I.e. *First order languages*
3. describe formal system
4. show that the theorems of the formal system are exactly the formulas true in all structures. (Goedel Completeness Theorem

2.1 Structures

Definition. 2.1. Suppose A is a set and $n \in \mathbb{N}$, An *n-ary relation* of A is a subset $\overline{R} \subseteq A^n$. An *$n$-ary function* on A is a function $\overline{f} : A^n \to A$.

Example.

1. ordering $\leq$ on $\mathbb{N}$ is 2-ary relation on $\mathbb{N}$.
2. addition, multiplication are 2-ary functions on their respective sets.
3. a 1-ary relation is just a subset. (even numbers in $\mathbb{N}$, etc.)

Notation: If $R \subseteq A^n$ is an n-ary relation, then we write $R(a_1, \dots, a_n)$ for $(a_1, \dots, a_n) \in R$.

Definition. 2.2. A *first-order structure* $\mathcal{A}$ consists of:

1. A nonempty set A – the *domain* of $\mathcal{A}$
2. A set $\{\overline{R}_i : i \in I\}$ of *relations* on A.
3. A set $\{\overline{f}_j : j \in J\}$ of *functions* on A.
4. A set $\{\overline{c}_k : k \in K\}$ of *constants* in A. $(c_k \in A)$

Remark. The sets I, J, K can be empty and are indexing sets (usually subsets of N). The information

$$(n_i : i \in I), (m_j : j \in J), K$$

is called the *signature* of $\mathcal{A}$. Might denote the structure by

$$\mathcal{A} = \langle A; (\overline{R}_i : i \in I), (\overline{f}_j : j \in J), (\overline{c}_k : k \in K) \rangle$$

Example.

1. *Orderings* $A = \mathbb{N}, \mathbb{Q}, \mathbb{R}$ and $I = \{1\}, J = K = \emptyset$, $n_i = 2$. and $\overline{R}_1(a, b)$ means $a \leq b$.
2. *Groups* The domain is just the underlying set of the group and we take the signature: $\overline{R}, \overline{m}, \overline{i}, \overline{e}$, where the first is the 2-ary relation of equality, then multiplication, inversion and the neutral element.
3. *Rings* The domain is just the underlying set of the group and we take the signature: $\overline{R}, \overline{m}, \overline{a}, \overline{i}, \overline{0}, \overline{1}$, where the first is the 2-ary relation of equality, then multiplication, addition, subtraction and the neutral elements for addition and multiplication.
4. *Graphs* $A = V$ a set of vertices, a binary relation $\overline{E} \subseteq A^2$ the set of vertices (connected elements) and again the binary relation $\overline{R}$ for equality.

2.2 First-order languages

Definition. 2.3. A *first-order-language* $\mathcal{L}$ has an alphabet of symbols of the following types:

1. *variables*: $x_0, x_1, x_2, \dots$
2. *punctuation*:), (,   , (the comma is a symbol as well)
3. *connectives*: $\neg, \to$

4. *quantifier*: $\forall$
5. *relation symbols*: R_i $i \in I$
6. *function symbols*: f_j $j \in J$
7. *constant symbols*: c_k $k \in K$

Here I, J, K are indexing sets (can once again be empty). Each R_i comes equipped with an *arity* n_i; each f_j comes equipped with an arity m_j. The information

$$(n_i : i \in I), (m_j : j \in J), K$$

is called the *signature* of $\mathcal{L}$. A first order structure $\mathcal{A}$ with the same signature as $\mathcal{L}$ is referred to as an *$\mathcal{L}$-structure*.

Definition. 2.4. A *term* of $\mathcal{L}$ is defined as follows:

1. any variable is a term
2. any constant symbol is a term
3. if f is an n-ary function symbol of $\mathcal{L}$ and $t_1, \dots, t_n$ are terms, then

   $$f(t_1, \dots, t_n)$$

   is a term.
4. any term arises in this way.

Definition. 2.5.

1. An *atomic formula* of $\mathcal{L}$ is of the form $R(t_1, \dots, t_n)$, where R is a n-ary relation symbol of $\mathcal{L}$ and $t_1, \dots, t_n$ are terms.
2. the *formulas* of $\mathcal{L}$ are defined as follows:
   a) any atomic formula is a formula
   b) if ϕ, ψ are $\mathcal{L}$ formulas, then

      $$(\neg\phi), (\phi \rightarrow \psi), (\forall x)\phi$$

      are $\mathcal{L}$-formulas, where x is any variable.
   c) every formula arises this way.

Example. Suppose $\mathcal{L}$ has

- 2-ary function symbol f,
- 1-ary relation symbol P,
- 2-ary relation symbol R,
- constants $c_1, c_2, \dots$

Some terms:

$$x, c, f(x_1, c_1), f(f(x_1, c_1), x_2), \dots$$

Some atomic formulas:

$$P(x_1), R(f(x_1, c_1), c_2), etc.$$

Some formulas: ...

Definition. 2.6. Suppose ϕ, ψ are $\mathcal{L}$-formulas, then

$$(\exists x)\phi \text{ means } (\neg(\forall x)(\neg\phi)$$

and the other shorthands as in the propositional logic.

Definition. 2.7. Suppose $\mathcal{L}$ is a first-order language with relation,function and constant symbols

- R_i (of arity n_i) for $i \in I$
- R_i (of arity m_j) for $j \in J$
- c_k for $k \in K$.

An *$\mathcal{L}$-structure* is a structure

$$\mathcal{A} = \langle A, (\overline{R}_i : i \in I), (\overline{f}_j : j \in J), (\overline{c}_k : k \in K) \rangle$$

of the same signature as $\mathcal{L}$. There is a correspondence between the relation, function and constant symbols and the actual relations, functions and constants in $\mathcal{A}$. THis correspondence, or $\mathcal{A}$ itself, is called an *interpretation* of $\mathcal{L}$.

Definition. 2.8. With the same notation as above, suppose $\mathcal{A}$ is an $\mathcal{L}$-structure. A *valuation in*

$$\mathcal{A}$$

is a function v from the set of terms on $\mathcal{L}$ to A satisfying:

a) $v(c_k) = \overline{c}$

b) if $t_1, \dots, t_m$ are terms of $\mathcal{L}$ and f is an m-ary function symbol, then $v(f(t_1, \dots, t_m)) = \overline{f}(v(t_1), \dots, v(t_m))$ where $\overline{f}$ is the interpretation of f in $\mathcal{A}$.

Lemma. 2.9. *Suppose $\mathcal{A}$ is an $\mathcal{L}$-structure and $a_0, a_1, \dots \in A$. Then there is a unique valuation v in $\mathcal{A}$ with $v(x_l) = a_l$ for all $l \in \mathbb{N}$. ($x_0, x_1, \dots$ are the variables of $\mathcal{L}$)*

Proof. By induction on the length of terms: show that if we let

1. $v(x_l) = a_l$ for all $l \in \mathbb{N}$.
2. $v(c_k) = \overline{c}_k$ for all $k \in K$.
3. $v(f(t_1, \dots, t_m) = \overline{f}(v(t_1), \dots, v(t_m))$

then v is a well-defined valuation. (rest exercise). □

Example. Groups: The domain is just the underlying set of the group and we take the signature: $\overline{R}, \overline{m}, \overline{i}, \overline{e}$, where the first is the 2-ary relation of equality, then multiplication, inversion and the neutral element. Let $\mathcal{G}$ be a group and $g, h \in \mathcal{G}$. Let v be a valuation with $v(x_0) = g, v(x_1) = h$. Then

$$v(m(m(x_0, x_1), i(x_0))) = \overline{m}(v(m(x_0, x_1)), v(i(x_0)) = \cdots = ghg^{-1}$$

Definition. 2.10. Suppose $\mathcal{A}$ is an $\mathcal{L}$ structure and x_l is any variable. Suppose v, w are valuations in $\mathcal{A}$. We say that v, w are *x_l-equivalent* if $v(x_m) = w(x_m)$ whenever $m \neq l$.

Definition. 2.11. Suppose $\mathcal{A}$ is an $\mathcal{L}$-structure and v is a valuation in $\mathcal{A}$. Define, for an $\mathcal{L}$-formula ϕ, what is meant by
v satisfies ϕ in $\mathcal{A}$,

1. atomic formulas Suppose R is an n-ary relation symbol and $t_1, \ldots, t_n$ are terms of $\mathcal{L}$. Then v satisfies the atomic formula $R(t_1, \ldots, t_n)$ if and only if $\overline{R}(v(t_1), \ldots, v(t_n))$ holds in $\mathcal{A}$.

2. $\mathcal{L}$-formulas Suppose that ϕ, ψ are $\mathcal{L}$ formulas and we already know about valuations satisfying ϕ, ψ. Then

   - v satisfies $(\neg\phi)$ iff v does not satisfy ϕ in $\mathcal{A}$.
   - v satisfies $(\phi \to \psi)$ in $\mathcal{A}$ iff it is not the case that v satisfies ϕ and v does not satisfy ψ.
   - v satisfies $(\forall x_l)\phi$ iff every valuation w, that is x_l-equivalent to v satisfies ϕ.

Remark. The definition above does not work, if the structure is empty. (So we will not allow it).

If v satisfies ϕ, write $v[\phi] = T$. Otherwise write $v[\phi] = F$.

If every valuation in $\mathcal{A}$ satisfies ϕ then we say that ϕ is *true* in $\mathcal{A}$, or $\mathcal{A}$ is *a model of* ϕ ($\mathcal{A} \vDash \phi$).

If $\mathcal{A} \vDash \phi$ for every $\mathcal{L}$-structure $\mathcal{A}$, we say that ϕ is *logically valid* and write $\vDash \phi$. – These are analogues of tautologies in the propositional logic. Difference: In propositional logic there is an algorithm to decide whether a given formula is a tautology. There is *no* such algorithm to decide whether a given $\mathcal{L}$-formula is logically valid or not. – Consequence of the Goedel Incompleteness Theorem.

Example.

1. Suppose $\mathcal{L}$ has a binary relation symbol R. The $\mathcal{L}$-formula

   $$R(x_1, x_2) \to (R(x_2, x_3) \to R(x_1, x_3)))$$

   is true in $\mathcal{A} = \langle \mathbb{N}; < \rangle$, where $<$ is the interpretation of R. If the formula is not true, then there is a valuation v such that v satisfies $R(x_1, x_2)$ and v does not satisfy $R(x_2, x_3) \to R(x_1, x_3)$. So $v[R(x_2, x_3) = T$ and $v[R(x_1, x_3)] = F$. Let $v(x_i) = a_i \quad \in \mathbb{N}$. So

   $$a_1 < a_2 \quad a_2\ a_3 \quad a_1 \not< a_3$$

   which is impossible since $<$ is transitive on $\mathbb{N}$.

2. The same formula is not true in the structure $\mathcal{B}$ with domain $\mathbb{N}$ where we interpret $R(x_i, x_j)$ as $x_i \neq x_j$. Take a valuation in $\mathcal{B}$ with $v(x_1) = 1 = v(x_3)$ and $v(x_2) = 2$.

3. Lemma: Suppose $\mathcal{A}$ is an $\mathcal{L}$-structure and ϕ an $\mathcal{L}$-formula. Let v be a valuation in $\mathcal{A}$, then: v satisfies $(\exists x_1)\phi$ (in $\mathcal{A}$) if and only if there is a valuation w which is x_1 equivalent to v such that w satisfies ϕ.

   Proof. "$\Rightarrow$" Suppose v satisfies $(\neg(\forall x_1)(\neg\phi))$. Hence v does not satisfy $(\forall x_1)(\neg\phi)$. So there is a valuation w x_1-equivalent to v such that w does not satisfy $(\neg\phi)$. Such a w satisfies w. "$\Leftarrow$" exercise. ☐

Example. $(\forall x_1)(\exists x_2)R(x_1, x_2)$ is true in $\langle\mathbb{Z}, <\rangle$ and $\langle\mathbb{N}, <\rangle$ but not in $\langle\mathbb{N}, >\rangle$.
Exercise. Suppose ϕ is any $\mathcal{L}$-formula. Then

- $(\exists x_1)(\forall x_2)\phi \to (\forall x_2)(\exists x_1)\phi)$ is logically valid.
- $((\forall x_2)(\exists x_1)\phi \to (\exists x_1)(\forall x_2)\phi)$ is not necessarily logically valid.

the first can be shown with valuation arguments. The second can be shown by giving an example.
Example (Some logically valid formulas). Consider the propositional formula

$$\chi \quad (p_1 \to (p_2 \to p_1))$$

Suppose $\mathcal{L}$ is a first order language and ϕ_1, ϕ_2 are $\mathcal{L}$-formulas. Substitute ϕ_1 in place of p_1 and ϕ_2 in place of p_2 in χ. We obtain an $\mathcal{L}$-formula

$$\theta \quad (\phi_1 \to (\phi_2 \to \phi_1))$$

One can check that θ is logically valid.

Definition. 2.12. Suppose χ is an $\mathcal{L}$-formula involving propositional variables $p_1, \ldots, p_n$. Suppose $\mathcal{L}$ is a first order language and $\phi_1, \ldots, \phi_n$ are $\mathcal{L}$-formulas. A *substitution instance* of χ is obtained by replacing each p_i by ϕ_i. Call the result θ.

Theorem. 2.13. *1. θ is an $\mathcal{L}$-formula.*

2. If χ is a tautology, then θ is logically valid.

Proof. 1. is clear by the definition of a formula.

2. Take an $\mathcal{L}$-structure $\mathcal{A}$ and a valuation v in $\mathcal{A}$. Use this to define a propositional valuation w with
$$w(p_i) = v[\phi_i] \quad i \leq n.$$
   Then prove by induction on the number of connectives in χ that the value
$$w(\chi) = v[\theta]$$
   In particular if χ is a tautology then $v[\theta] = T$. (E.g. in the inductive step you consider the case where χ is $(\alpha \to \beta)$. SO θ is $(\theta_1 \to \theta_2)$ where θ_1 is obtained from α etc. By the induction hypothesis
$$w(\alpha) = v[\theta_1] \text{ and } w(\beta) = v[\theta_2]$$
   Thus $w(\alpha \to \beta) = v[\theta_1 \to \theta_2]$. The other connectives and details are left as exercise.

☐

Remark. Not all logically formulas arise in this way. E.g. $(\exists x_1)(\forall x_2)\phi \to (\forall x_2)(\exists x_1)\phi)$

2.3 Bound and free variables in formulas

Example.

$$(R_1(x_1, x_2) \to (\forall x_3)R_2(x_1, x_3))$$

Definition. 2.14. Suppose ϕ, ψ are $\mathcal{L}$- formulas and $(\forall x_i)\phi$ occurs as a subformula of ψ, i.e. ψ is $\ldots(\forall x_i)\phi\ldots$. We say that ϕ is the *scope* of the quantifier $(\forall x_i)$ in ψ. An occurrence of a variable x_j in ψ is *bound* if it is in the scope of a quantifier $(\forall x_j)$ in ψ (or if it is simply written after a quantifier). Otherwise it is a *free* occurrence (of x_j). Variables having a free occurrence in ψ are called the *free variables* of ψ. A formula with no free variables is called a *closed formula* or a *sentence* (of $\mathcal{L}$).

Example.

1. $\psi_2 : \quad ((\forall x_1)R_1(x_1, x_2) \to R_2(x_1, x_2))$ x_1 is free since $R_1(x_1, x_2)$ is the scope for $(\forall x_1)$ but it is occurring freely in $R_2(x_1, x_2))$. Compare with $((\forall x_1)(R_1(x_1, x_2) \to R_2(x_1, x_2)))$. Now x_1 is bound.
2. $\psi_3 : \quad ((\exists x_1)R_1(x_1, x_2) \to (\forall(x_2)R_2(x_2, x_3))$. Here, x_1 is bound but x_2, x_3 are free.

Remark (notation). If ψ is an $\mathcal{L}$-formula with free variables amongst $x_1, \ldots, x_n$, we might write $\psi(x_1, \ldots, x_n)$. If $t_1, \ldots, t_n$ are terms, by $\psi(t_1, \ldots, t_n)$ we mean the $\mathcal{L}$-formula obtained by replacing each *free* occurrence of x_i by t_i.

E.g.

$$\psi(x_1, x_2) : \quad ((\forall x_1)R_1(x_1, x_2) \to (\forall x_3)R_2(x_1, x_2, x_3))$$

$$t_1 : \quad f_1(x_1) \text{ and } t_2 : \quad f_2(x_1, x_2)$$

So

$$\psi(t_1, t_2) : \quad ((\forall x_1)R_1(x_1, t_2) \to (\forall x_3)R_2(t_1, t_2, x_3))$$

Theorem. 2.15. *Suppose ϕ is a closed $\mathcal{L}$-formula and $\mathcal{A}$ is an $\mathcal{L}$-structure, then either*

$$\mathcal{A} \models \phi \text{ or } \mathcal{A} \models (\neg\phi).$$

More generally, if ϕ has free variables amongst $x_1, \ldots, x_n$ and v, w are valuations in $\mathcal{A}$ with $v(x_i) = w(x_i)$ for $i = 1, \ldots, n$, then

$$v[\phi] = T \Leftrightarrow w[\phi] = T.$$

(Allow $n = 0$ for the case with no free variables.)

Proof. Note that the first statement follows from the generalization. If ϕ has no free variables, then, for any valuations v, w in $\mathcal{A}$, they agree on the free variables of ϕ.

Proving the generalization by induction on the number of connectives and quantifiers in ϕ.

Base case: ϕ is atomic $R(t_1, \ldots, t_m)$ for terms t_j. The t_j only involve variables amongst $x_1, \ldots, x_n$. As v, w agree on these variables, they agree on t_j, i.e. $v(t_j) = w(t_j)$. So $v[R(t_1, \ldots, t_m)] = T$ iff

$$\overline{R}(v(t_1), \ldots, v(t_m))$$

which is equivalent to

$$w[R(t_1, \ldots, t_m)] = T.$$

induction step: ϕ is $(\neg\psi), (\chi \to \psi)$ or $(\forall x_i)\psi$. The first two cases are left as exercise. Suppose ϕ is $(\forall x_i)\psi$. Suppose $v[\phi] = F$. Then there is a valuation v‘ that is x_l-equivalent to v with $v‘[\phi] = F$. The free variables on ψ are amongst $x_1, \dots, x_n, x_i$. Let w‘ be the valuation x_i-equivalent to w with $w‘(x_i) = v‘(x_i)$. Then $v‘, w‘$ agree on the free variables on ψ. By the induction hypothesis

$$v‘[\psi] = w‘[\psi] \text{ so } w‘[\psi] = F.$$

As w‘ is x_i-equivalent to w we obtain $w[(\forall x_i)\psi] = F$.

□

Definition. 2.16. Let ϕ be an $\mathcal{L}$-formula, x_i a variable, t an $\mathcal{L}$-term, then we say t *is free for* x_i *in* ϕ, if there is no variable x_j in t such that x_i has a free occurrence within the scope of a quantifier $(\forall x_j)$ in ϕ.

Theorem. 2.17. *Suppose $\phi(x_1)$ is an $\mathcal{L}$-formula (possibly with other free variables). Let t be a term free for x_1 in ϕ. Then*

$$\models ((\forall x_1)\phi(x_1) \to \phi(t)).$$

In particular, if $\mathcal{A}$ is an $\mathcal{L}$-structure with $\mathcal{A} \models (\forall x_1)\phi(x_1)$ then $\mathcal{A} \models \phi(t)$.

Lemma. 2.18. *With this notation, suppose v is a valuation in $\mathcal{A}$. Let w be a valuation in $\mathcal{A}$ with is x_1 equivalent to v, with $w(x_1) = v(t)$. Then*

$$w[\phi(x_1)] = T \Leftrightarrow v[\phi(t)] = T$$

Proof. The proof is posted on blackboard (omitted here) and works via induction on the connectives and quantifiers. □

Proof of theorem. Suppose v is a valuation with $v[\phi(t)] = F$. Claim: $v[(\forall x_1)\phi(x_1)] = F$. To show this claim, take w to be a x_1-equivalent valuation to v. such that $w(x_1) = v(t)$. Then by the lemma above

$$w[\phi(x_1)] = F \text{ implies } v[(\forall x_1)\phi(x_1)] = F$$

Thus the implication is shown. □

2.4 The formal system $K_{\mathcal{L}}$

Definition. 2.19. Suppose $\mathcal{L}$ is a first order language. The *formal system* $K_{\mathcal{L}}$ has as

1. **formulas**: the $\mathcal{L}$-formulas
2. **axioms**: For ϕ, ψ, χ $\mathcal{L}$-formulas we have

   A1 $(\phi \to (\psi \to \phi))$

   A2 $((\phi \to (\psi \to \chi)) \to ((\phi \to \psi) \to (\phi \to \chi)))$

   A3 : $(((\neg\psi) \to (\neg\phi)) \to (\phi \to \psi))$

   K1 : $(\forall x_i)\phi(x_i) \to \phi(t)$, where t is a term free for x_i in ϕ.

   K2 : $((\forall x_i)(\phi \to \psi) \to (\phi \to \forall x_i\psi))$ if x_i is not free in ϕ.

3. **deduction rules**

    a) *Modus Ponens (MP)*: from $\phi, (\phi \to \psi)$ deduce ψ.

    b) Gen(generalization): From ϕ deduce $(\forall x_i)\phi)$.

A *proof* in $K_{\mathcal{L}}$ arises in the same way as a proof in L. A *theorem* of $K_{\mathcal{L}}$ is the last formula in some proof. Write $\vdash_{\mathcal{L}} \phi$ for "ϕ is a theorem in $k_{\mathcal{L}}$".

Remark. Sometimes we omit the $K_{\mathcal{L}}$ and write $\vdash \phi$.

Definition. 2.20. Suppose Σ is a set of $\mathcal{L}$-formulas and ψ an $\mathcal{L}$-formula. A *deduction from* Σ is a finite sequence of formulas, ending with ψ, each of which is one of:

1. an axiom
2. a formula in Σ
3. obtained fro earlier formulas in the deduction using the deduction rules, with the only restriction, that when gen is applied it does not include a variable occurring freely in Σ. Write $\Sigma \vdash \phi$ if there is a deduction of ϕ fro Σ.

    Remark. If Σ consists of closed formulas we do not need to worry about the restriction on Generalization.

Theorem. 2.21. *Suppose ϕ is an $\mathcal{L}$-formula which is a substitution instance of a tautology in propositional logic. Then $\vdash_{K_{\mathcal{L}}} \phi$.*

Example. $\vdash_{K_{\mathcal{L}}} (\neg\neg\phi \to \phi)$ is such a case, etc.

Proof. There is a tautology χ with propositional variables $p_1, \dots, p_n$ and $\mathcal{L}$-formulas $\psi_1, \dots, \psi_n$ such that ϕ is obtained from χ by substituting ψ_i for p_i $(i \leq n)$. By completeness of propositional logic (1.26) there is a proof in L of χ: $\chi_1, \dots, \chi_r$ (each one a propositional formula and $\chi_r = \chi$. If we substitute ψ_1, ψ_n for $p_1, \dots, p_n$ in all χ_j we obtain a sequence of $\mathcal{L}$-formulas $\phi_1, \dots, \phi_r$ which is a proof of $\phi = \phi_r$. □

Theorem. 2.22 (Soundness)**.** *If $\vdash_{K_{\mathcal{L}}} \phi$ then $\vDash \phi$ (i.e. it is logically valid).*

Proof. Like in the proof for L (1.18) we need to show

1. Axioms are logically valid
2. Deduction rules preserve logical validity

1. The axioms A1, A2, A3 are logically valid since they are substitution instances of the propositional tautologies, which are logically valid themselves.
    K1 is logically valid by an earlier result.
    K2: Suppose we have a valuation v such that

$$v[(\phi \to (\forall x_i)\psi)] = F$$

implying $v[\phi] = T$ and $v[(\forall x_i)\psi)] = F$. So there is a valuation v‘ x_i equivalent to v with v‘$[\psi] = F$. Since v and v‘ agree on all variables free in ϕ an earlier result tells us

$$v[\phi] = v‘[\phi] = T$$

$$v‘[(\phi \to \psi)] = F.$$

Thus $v[(\forall x_i)(\phi \to \psi)] = F$ and $v[K2] = T$.

2. deduction rules are left as exercise.

□

Exercise. Suppose $\Sigma \vdash \psi$, then for every valuation v with $v[\sigma] = T$ for all $\sigma \in \Sigma$ we have $v[\psi] = T$.

Corollary. 2.23. *There is no $\mathcal{L}$-formula ϕ with $\vdash_{K_\mathcal{L}} \phi$ and $\vdash_{K_\mathcal{L}} (\neg\phi)$.*

Theorem. 2.24 (Deduction Theorem)**.** *Suppose $\mathcal{L}$ is a first order language. Σ a set of $\mathcal{L}$-formulas and ϕ, ψ $\mathcal{L}$-formulas. Then $\Sigma \cup \{\phi\} \vdash \psi$ implies $\Sigma \vdash (\phi \to \psi$.*

Proof. Follows from proof of the deduction Theorem for L (1.15) by induction on the length of the deduction.

Base case: One line deduction Argue exactly as in (1.15)

inductive step: Suppose ψ follows from earlier formulas in the deduction by MP or Gen.

MP Exactly as in (1.15)

Gen Suppose ψ is obtained using Gen, then $\psi = (\forall x_i)\theta$ and $\Sigma \cup \{\phi\} \vdash \theta$ and x_i is not free in any formula in $\Sigma \cup \{\phi\}$. By induction we have:

$$\Sigma \vdash (\phi \to \theta)$$

and by K2 $\Sigma \vdash (\forall x_i(\phi \to \theta) \to (\phi \to \forall x_i\theta))$. By Gen $\Sigma \vdash \forall x_i(\phi \to \theta)$ and x_i is not free in any formula in Σ. So by MP we get $\Sigma \vdash (\phi \to (\forall x_i)\theta)$.

□

2.5 Goedel completeness Theorem

Definition. 2.25. A set Σ of $\mathcal{L}$-formulas is *consistent* if there is no $\mathcal{L}$-formula ϕ with

$$\Sigma \vdash_{K_\mathcal{L}} \phi \text{ and } \Sigma \vdash_{K_\mathcal{L}} (\neg\phi).$$

By Soundness (or 2.22) $\emptyset$ is consistent.

Remark. If Σ is inconsistent then $\Sigma \vdash \chi$ for any $\mathcal{L}$-formula χ.

We want to show, that if Σ is a consistent set of closed $\mathcal{L}$-formulas then there is an $\mathcal{L}$-structure $\mathcal{A}$ with $\mathcal{A} \vDash \sigma$ for all $\sigma \in \Sigma$. From there the completeness Theorem will follow.

Remark. Simplification: We are going to assume, that $\mathcal{L}$ is countable. I.e. The variables, constants, relations and functions are all just countably many. So we can enumerate all the $\mathcal{L}$-formulas by the natural numbers. In particular, we can enumerate the closed $\mathcal{L}$-formulas as

$$\psi_0, \psi_1, \psi_2, \dots \quad .$$

Proposition. 2.26. *Suppose Σ is a (consistent) set of closed $\mathcal{L}$-formulas and ϕ is a closed $\mathcal{L}$-formula*

1. *If $\Sigma \nvdash \phi$ then $\Sigma \cup \{\neg\phi\}$ is consistent.*
2. *There is a consistent set $\Sigma^* \supseteq \Sigma$ of closed $\mathcal{L}$-formulas such that, for every closed $\mathcal{L}$-formula ψ either*

$$\Sigma^* \vdash \psi \text{ or } \Sigma^* \vdash (\neg\psi)$$

Proof. Similar to the proofs before the completeness Theorem in L (1.26. □

Theorem. 2.27. *Suppose that Σ is a consistent set of closed $\mathcal{L}$-formulas. Then there is a countable $\mathcal{L}$-structure $\mathcal{A}$ with $\mathcal{A} \vDash \Sigma$ (meaning that $\mathcal{A} \vDash \sigma$ for every $\sigma \in \Sigma$.*

Proof. Hard part, will come later. □

Theorem. 2.28. *Let Σ be a set of closed $\mathcal{L}$-formulas and ϕ a closed $\mathcal{L}$-formula. If every model of Σ is a model of ϕ, then $\Sigma \vdash \phi$. [Where* being a model of Σ implies being a model of ϕ *means that if $\mathcal{A} \vDash \sigma$ for all $\sigma \in \Sigma$ then $\mathcal{A} \vDash \phi$.]*

Remark. This is the opposite direction to Soundness (2.22).

Proof. We may assume, that Σ is consistent, since otherwise everything is a consequence. of Σ. By assumption there is no model of $\Sigma \cup \{(\neg\phi)\}$. So by 2.27, $\Sigma \cup \{(\neg\phi)\}$ is inconsistent. So by the first part of the above proposition, $\Sigma \vdash \phi$. □

Theorem. 2.29 (Godel Completeness Theorem for $K_{\mathcal{L}}$)**.** *If ϕ is an $\mathcal{L}$-formula with $\vDash \phi$, then ϕ is a theorem of $K_{\mathcal{L}}$, i.e. $\vdash_{K_{\mathcal{L}}}$.*

Proof. If ϕ is closed, this follows from (2.28 with $\Sigma = \emptyset$. Suppose ϕ has free variables amongst $x_1, \dots, x_n$ and consider the closed formula ψ

$$(\forall x_1) \dots (\forall x_n)\phi.$$

As $\vDash \phi$ we obtain $\vDash \psi$. So, by the closed case it follows that $\vdash \psi$. i.e.

$$\vdash (\forall x_1) \dots (\forall x_n)\phi.$$

If θ is any formula then $\vdash ((\forall x_i)\theta \to \theta)$. So with Modus Ponens (applied n times) we obtain $\vdash_{K_{\mathcal{L}}} \phi$. □

Corollary. 2.30 (Compactness Theorem)**.** *Suppose Σ is a set of closed $\mathcal{L}$-formulas and every finite subset of Σ has a model. Then Σ has a model.*

Proof. Suppose Σ has no model, then by 2.27 Σ is inconsistent, thus there is a formula ϕ with $\Sigma \vdash \phi$ and $\Sigma \vdash (\neg\phi)$. Deductions only involve finitely many formulas in Σ. So there is a finite subset $\Sigma_0 \subseteq \Sigma$ with $\Sigma_0 \vdash \phi$ and $\Sigma_0 \vdash (\neg\phi)$. But then Σ_0 is a finite subset of Σ having no model – contradiction. $\square$

Sketch for the proof of 2.27. Series of steps; notation is cumulative. They can be found in greater detail on Blackboard.

Step 1. Let $b_0, \ldots, b_l$ be new constant symbols. Form $\mathcal{L}^+$ by adding these to the symbols of $\mathcal{L}$. Regard Σ as a set of $\mathcal{L}^+$-formulas. Check Σ is still consistent (in the formal system of $\mathcal{L}^+$). **Note:** $\mathcal{L}^+$ is still a countable language.

Step 2. (Adding witnesses)

Lemma. 2.31. *There is a consistent set of closed $\mathcal{L}^+$-formulas $\Sigma_\infty \supseteq \Sigma$ such that for every formula $\theta(x_i)$ with one free variable there is some b_j with*

$$\Sigma_\infty \vdash (\neg(\forall x_i)\theta(x_i) \to \neg\theta(b_j))$$

[Think of $\theta(x_i)$ as $\neg\chi(x_i)$. Then this formula is essentially: $(\exists x_i)\chi(x_i) \to \chi(b_j)$ - so b_j witnesses the idea that there exists some x_i that satisfies χ.]

Step 3. Use the Lindenbaum Lemma - there is a consistent set $\Sigma^* \supseteq \Sigma$ of closed $\mathcal{L}^+$-formulas such that for every closed ϕ either $\Sigma^* \vdash \phi$ or $\Sigma^* \vdash \neg\phi$.

step 4. (Building a strucure) Let $A := \{\bar{t} : t \text{ is a closed term of } \mathcal{L}^+\}$ **Note**: A term is closed if it only involves constant symbols and function symbols (no variables). We use the $\bar{}$ to distinguish when we are thinking of a term as an element of A. As $\mathcal{L}^+$ is countable, A is countable. Make A into an $\mathcal{L}^+$-structure.

- Each constant symbol c of $\mathcal{L}^+$ is interpreted as $\bar{c} \in A$.
- Suppose R is an n-ary relation symbol. Define the relation $\overline{R} \subseteq A^n$ by

$$(\overline{t_1}, \ldots, \overline{t_n} \in \overline{R} \Leftrightarrow \Sigma^* \vdash R(t_1, \ldots, t_n),$$

  where $t_1, \ldots, t_n$ are closed $\mathcal{L}^+$-terms.
- Suppose f is an m-ary function symbol. Define a function $A^m \to A$ by

$$\overline{f}(\overline{t_1}, \ldots, t_m) = \overline{f(t_1, \ldots, t_m)},$$

  for closed terms $t_1, \ldots, t_m$. Call this structure $\mathcal{A}$.

Note: If v is a valuation in $\mathcal{A}$ and t is a closed term, then $v(t) = \bar{t}$ (by steps 1 and 3 here).

Lemma. 2.32 (Main Lemma)**.** *For every closed $\mathcal{L}^+$-formula ϕ*

$$\Sigma^* \vdash \phi \Leftrightarrow A \vDash \phi.$$

Proof. By induction on the number of connectives and quantifiers in ϕ.

Base case: ϕ is atomic, i.e. ϕ is $\mathbb{R}(t_1, \ldots, t_n)$ for some closed terms t_i and relation symbol R. By definition of the structure the base step holds.

Inductive step: Assume the equivalence holds for closed formulas involving fewer connectives and quantifiers.

- Suppose ϕ is $(\neg\psi)$ then by an earlier result

  $$\mathcal{A} \models \phi \Leftrightarrow \mathcal{A} \not\models \psi$$

  which happens if and only if

  $$\Sigma^* \not\vdash \psi$$

  by the hypothesis. Thus

  $$\Sigma^* \vdash \neg\psi.$$

- ϕ being $(\psi \to \chi)$ is left as exercise.

- ϕ is $(\forall x_i)\psi$ If x_i is not free in ψ. So ψ is closed and we can use the induction hypothesis. Otherwise, x_i is free in ψ. So $\psi(x_i)$ has a single free variable. Now Suppose that $\mathcal{A} \models \phi$ and $\Sigma^* \not\vdash \phi$. Then by the definition of Σ^*, $\Sigma^* \vdash (\neg\phi)$. By step 2:

  $$\Sigma^* \vdash (\neg(\forall x_i)\psi(x_i) \to (\neg\psi(b_j)))$$

  for some constant symbol b_j. I.e. $\Sigma^* \vdash ((\neg\phi) \to (\neg\phi(b_j)))$ So $\Sigma^* \vdash \neg\psi(b_j)$. But $\neg\psi(b_j)$ is closed and we obtain

  $$\mathcal{A} \models \neg\psi(b_j).$$

  This contradicts $\mathcal{A} \models (\forall x_i)\psi$. [Take a valuation v in $\mathcal{A}$ with $v(x_i) = b_j$; then v does not satisfy ψ, by the above.] Thus the direction "$\Leftarrow$" is shown. The other direction can be found on Blackboard.

$\square$

$\square$

2.6 Equality

MISSING

Theorem. 2.33 (Loewenheim-Skolem)**.** *Suppose $\mathcal{L}^=$ is a countable first order language with equality and $\mathcal{B}$ is a normal $\mathcal{L}^=$ structure. Then there is a countable normal $\mathcal{L}^=$-structure $\mathcal{A}$ such that for every closed $\mathcal{L}^=$-formula ϕ*

$$\mathcal{B} \models \phi \Leftrightarrow \mathcal{A} \models \phi$$

Example. $\mathcal{B} + \langle \mathbb{R}, +, \cdot, 0, 1, \exp() \rangle \mathcal{A} =?$

Proof. $\Sigma = \{\text{closed } \phi : \mathcal{B} \models \phi\}$ (called the *theory of* $\mathcal{B}$. Then $\Sigma \supseteq \Sigma_E$ (axioms for equality), and Σ is consistent. By 2.27 has a countable model $\mathcal{C}$. Then $\hat{\mathcal{C}}$ is a countable normal model of Σ. (2.6.3. in the above section??) So if ϕ is closed and $\mathcal{B} \models \phi$ then $\hat{\mathcal{C}} \models \phi$. Conversely, if ϕ is closed and $\mathcal{B} \not\models \phi$ then $\mathcal{B} \models (\neg\phi)$. Thus $\hat{\mathcal{C}} \models (\neg\phi)$ so $\hat{\mathcal{C}} \not\models \phi$. $\square$

Example. 1. linear orders: Let $\mathcal{L}^=$ be a first order language with equality and a 2-ary relation symbol $\leq$.

Definition. 2.34. A *linear order* $\mathcal{A} = \langle A; \leq_A \rangle$ is a normal model of:

$$\phi_1 : \quad (\forall x_1)(\forall x_2)(((x_1 \leq x_2) \wedge (x_2 \leq x_1)) \leftrightarrow (x_1 = x_2))$$

$$\phi_2 : \quad (\forall x_1)(\forall x_2)(\forall x_3)(((x_1 \leq x_2) \wedge (x_2 \leq x_3)) \rightarrow (x_1 \leq x_3))$$

$$\phi_3 : \quad (\forall x_1)(\forall x_2)(((x_1 \leq x_2) \vee (x_2 \leq x_1))$$

It is *dense* if also

$$\phi_1 : \quad (\forall x_1)(\forall x_2)(\exists x_3)((x_1 < x_2) \rightarrow ((x_1 < x_3) \wedge (x_3 < x_2))))$$

where $x_1 < x_2$ is the obvious abbreviation. It is *without endpoints* if

$$\phi_5 : \quad (\forall x_1)(\exists x_2)(x_1 < x_2)$$

and

$$\phi_6 : \quad (\forall x_1)(\exists x_2)(x_2 < x_1) \quad .$$

Let $\Delta = \{\phi_1, \ldots, \phi_6\}$.

Theorem. 2.35. *1. For every closed $\mathcal{L}^=$-formula ϕ*

$$\mathbb{Q} \vDash \phi \Leftrightarrow \mathbb{R} \vDash \phi$$

2. There is an algorithm to decide, given a closed $\mathcal{L}^=$-formula ϕ, whether $\mathbb{Q} \vDash \phi$ or $\mathbb{Q} \vDash (\neg\phi)$.

Definition. 2.36. Linear orderings $\mathcal{A}$ and $\mathcal{B}$ are *isomorphic* if there is an order preserving bijection $\alpha : A \rightarrow B$ i.e. for all $a, a‘ \in A$

$$a \leq_A a‘ \Leftrightarrow \alpha(a) \leq_B \alpha(a‘) \quad .$$

If $\mathcal{A}, \mathcal{B}$ are isomorphic and ϕ is closed then $\mathcal{A} \vDash \phi \Leftrightarrow \mathcal{B} \vDash \phi$.

Theorem. 2.37 (Cantor)**.** *If $\mathcal{A}, \mathcal{B}$ are countable dense linear order without endpoints, then $\mathcal{A}, \mathcal{B}$ are isomorphic.*

Lemma. 2.38 (Los-Vaught Test)**.** *Let $\Sigma = \Sigma_E \cup \Delta$. Then for every closed $\mathcal{L}^=$-formula ϕ we have either*

$$\Sigma \vdash \phi \quad \textit{or} \quad \Sigma \vdash (\neg\phi)$$

[Say Σ is complete.]

Proof. Suppose not. Then because Σ is consistent (we know it has models) we can use 1.22 to get

$$\Sigma_1 = \Sigma \cup \{(\neg\phi\}$$

and

$$\Sigma_2 = \Sigma \cup \{(\neg\neg\phi\}$$

which are consistent. By earlier results it follows that Σ_1, Σ_2 have countable normal models $\mathcal{A}_1, \mathcal{A}_2$. So $\mathcal{A}_1, \mathcal{A}_2$ are countable linear orders without endpoints and $\mathcal{A}_1 \vDash (\neg\phi)$ and $\mathcal{A}_2 \vDash \phi$. This is a contradiction due to Cantor‘s Theorem above. $\square$

Proof of 2.35. Show that $\mathbb{Q} \vDash \phi \Leftrightarrow \Sigma \vdash \phi$. "$\Leftarrow$": As $\mathbb{Q} \vDash \Sigma$ we already know this direction. "$\Rightarrow$": If $\Sigma \nvdash \phi$ then by the lemma above $\Sigma \vdash (\neg\phi)$. Thus $\mathbb{Q} \vDash (\neg\phi)$, so $\mathbb{Q} \nvDash \phi$. Similarly $\mathbb{R} \vDash \phi \Leftrightarrow \Sigma \vdash \phi$ proving the result. □

We want an algorithm deciding - given a closed formula θ - whether

$$\langle \mathbb{Q}; \leq \rangle \vDash \theta$$

or

$$\langle \mathbb{Q}; \leq \rangle \vDash \neg\theta$$

Remark and Definition. 2.39. Σ is a *recursively enumerable* set of formulas: i.e. we can write an algorithm to systematically generate the formulas in Σ. Note that the set of axioms for $K_{\mathcal{L}}$ is recursively enumerable. So the set of deductions from Σ is also recursively enumerable. So the set of consequences of Σ is recursively enumerable.

Method: Run the algorithm from the remark above, that generates all consequences of Σ. By $\Sigma \vdash \theta \Leftrightarrow \langle \mathbb{Q}; \leq \rangle \vDash \theta$ at the same point, we will see either θ or $(\neg\theta)$. At this point the method stops.

Remark. This method depends on the completeness theorem and the axiomatizability of Δ for $\langle \mathbb{Q}; \leq \rangle \vDash \theta$.

It works for some other structures **BUT** there is no such algorithm in general. Example: $\langle \mathbb{N}; +; \cdot; 0 \rangle$.(Goedel's Incompleteness Theorem)

3 Set Theory

3.1 Basic set theory

Definition. 3.1. 1. *extensionality* Sets A, B are *equal* iff

$$\forall x((x \in A) \leftrightarrow (x \in B))$$

2. *natural Numbers*

$$\mathbb{N} = \{0, 1, \dots\}$$

and since in set theory we like our objects to be sets we think of

$$0 = \emptyset, 1 = \{0\} = \{\emptyset\}, \dots, n = \{0, \dots, n\}, \dots$$

. Note that this gives $m < n \Leftrightarrow m \in n \Leftrightarrow m \subseteq n$.

3. *ordered pair* The ordered pair (x, y) is the set { {x}, {x,y} }.
Exercise. For any x, y, z, w we have $(x, y) = (z, w)$ iff $x = z$ and $y = w$.

If A, B are sets then $A \times B = \{(a, b) \mid a \in A, b \in B\}$ $A^2 = A \times A$ and so forth. The set of *finite sequences* of elements of A is the set

$$\bigcup_{n \in \mathbb{N}} A^n$$

(also $A^0 = \{\emptyset\}$).

4. *functions* Think of a function $f : A \to B$ as a subset of $A \times B$. Where $A = \mathrm{dom} f$ (*domain*), $B = \mathrm{ran} f$ (*range*). If $X \subseteq A$

   $$f[X] = \{f(a) \mid a \in X\} \subseteq B.$$

   The set of functions from A to B is $B^A \subseteq P(A \times B)$ (powerset).

3.2 Cardinality

Definition. 3.2. Sets A, B are *equinumerous* (or of the *same cardinality*) if there is a bijection $f : A \to B$. Write $A \approx B$ or $|A| = |B|$.

Definition. 3.3. A set A is *finite* if it is equinumerous with some element of $\mathbb{N}$. A set is countably infinite if it is equinumerous with $\mathbb{N}$. *Countable* if it is either of the above.

Remark. The following are facts:

1. Every subset of a countable set is countable.
2. A set A is countable iff there is an injective function $f : A \to \mathbb{N}$.
3. If A, B are countable then $A \times B$ is countable.
4. If $A_0, A_1, \ldots$ are countable then $\bigcup_{i \in \mathbb{N}} A_i$ is countable (using Axiom of choice).

Exercise. 1. $\mathbb{Q}$ is countable.

2. $\bigcup_{i \in \mathbb{N}} A^i$ is countable.
3. (Cantor) $\mathbb{R}$ is not countable.

Theorem. 3.4 (Cantor)**.** *(If X is any set then $\mathbb{P}(X)$ is the set of all subsets of X.) There is no surjective function $f : X \to \mathbb{P}(X)$.*

Proof. Suppose there is such a function. Let $Y = \{y \in X \mid y \notin f(y)\}$. There is $z \in X$ with $f(z) = Y$.

- If $z \in Y$ then $z \notin f(z) = Y$
- If $z \notin Y$ then $z \notin f(z)$ so $z \in Y$.

– Contradiction. □

Definition. 3.5. For sets A, B write

$$|A| \leq |B|$$

if there is an injective function $f : A \to B$.

Remark. $|X| \leq |\mathbb{P}(X)|$, since $x \to \{x\}$ is injective. Since $|X| \neq |\mathbb{P}(X)|$ this means that there is no injective function $\mathbb{P}(X) \to X$. (see next theorem) We write $|X| < |\mathbb{P}(X)$.

Exercise. If $|A| \leq |B|$ and $|B| \leq |C|$ then $|A| \leq |C|$.

Theorem. 3.6 (Schroeder-Bernstein)**.** *Suppose A, B are sets and $f : A \to B$ and $g : B \to A$ are injective, then $|A| = |B|(A \approx B)$.*

Proof. Let $h := g \circ f : A \to A$. Let $A_0 = A \setminus g[B]$ and for $n > 0$ let $A_n = h[A_{n-1}]$. Let

$$A^* = \bigcup_{n \in \mathbb{N}} A_n \text{ and } B^* = f[A^*].$$

Note that, $h[A^*] \subseteq A^*$ so $g[B^*] = h[A^*] \subseteq A^*$. **Claim:** $g[B \setminus b^*] = A \setminus A^*$. Once we have this, f gives a bijection $A^* \to B^*$ and g gives a bijection $B \setminus B^* \to A \setminus A^*$. So

$$k(a) = \begin{cases} f(a) \text{ if } a \in A^* \\ g(a) \text{ otherwise} \end{cases} .$$

Proof of the claim:

1. Let $a \in A \setminus A^*$. As $a \notin A_0$ there is $b \in B$ with $g(b) = a$. hen $b \notin B^*$ as

   $$b \in B^* \Rightarrow b \in f[A^*] \Rightarrow g(b) \in h[A^*] \subseteq A^*$$

   but this would imply that

   $$a \in A^*$$

   which is a contrdiction. Therefore

   $$g[B \setminus B^*] = A \setminus A^*.$$

2. Let $b \in B$; suppose $g(b) \in A^*$. Show $b \in B^*$. As $g(b) \notin A_0$ we have $g(b) \in A_n$ for some $n > 0$. So $g(b) = h(a)$, for some $a \in A_{n-1}$. Thus $g(b) = g(f(a))$ and therefore $b = f(a)$ for some $a \in A^*$. Thus $b \in f[A^*]$ and the other direction of the claim is proven as well.

□

Example. The following sets are equinumerable:

1. $S_1 =$ the set of all sequences of $0, 1 = \{0, 1\}^{\mathbb{N}}$.
2. $S_2 = \mathbb{R}$
3. $S_3 = \mathbb{P}(\mathbb{N})$
4. $S_4 = \mathbb{P}(\mathbb{N} \times \mathbb{N})$
5. $S_5 =$ set of all sequences of natural numbers $= \mathbb{N}^{\mathbb{N}}$

Proof. We find injective functions $f_{i,j} : S_i \to S_j$, where $i, j \in \{1, \ldots, 5\}$ and then use the Schroeder-Bernstein Theorem.3.6.

As $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$ we get that $S_3 \approx S_4$. Also $S_1 \subseteq S_5 \subseteq S_4$. We have a bijection

$$f_{3,1} : \mathbb{P}(\mathbb{N}) \to S_1.$$

For $X \subseteq \mathbb{N}$ we define $f_{3,1}(X) = (a_n)_{n\in\mathbb{N}}$, where

$$\begin{cases} 0 \text{ if } n \notin X \\ 1 \text{ otherwise} \end{cases}$$

Furthermore we take $f_{1,2} : (a_n)_{n\in\mathbb{N}} \mapsto 0.a_0a_1, dots$ to be the map onto the decimal expansion, which is obviously injective.

For $f_{2,5}$ we map the decimal expansion onto a sequence

$$f_{2,5}(\pm n.m_1m_2 \cdots \mapsto (0 \text{or} 1, n, m_1, \ldots$$

where the first value stands for $\pm$. □

Remark. **If A, B are sets is one of**

$$|A| \leq |B| \text{ or } |B| \leq |A|?$$

If we assume the Axiom of Choice (AC) then this question has a positive answer. Also: **Is there $X \subseteq \mathbb{R}$ with $|\mathbb{N}| < |X| < |\mathbb{R}|$?** [Continuum Hypothesis says 'no' ...]

3.3 Axioms for Set Theory

Zermelo-Fraenkel Axioms say how we are allowed to 'build' sets. All can be expressed in a first order language (with =) using a single 2-ary relation symbol $\in$.

We have to avoid the *Russel Paradox*:

$$S = \{x : x \text{ is a set and } x \notin x\}$$

If this is a set, is $S \in S$?. $[(\exists S)(\forall x)((x \in S) \leftrightarrow (x \notin x))$ leads to inconsistency.]

Definition. 3.7. The following are the *Zermelo-Fraenkel Axioms* 1-6

1. *extensionality*
   $$(\forall x)(\forall y)((x = y) \leftrightarrow (\forall z)((z \in x) \leftrightarrow (z \in y)))$$
   'Two sets are equal iff they have the same elements'.
2. *Empty set axiom*
   $$(\exists x)(\forall y)(y \notin x)$$
   There is a unique set with this property $\emptyset$.

3. *Pairing axiom*

$$(\forall x)(\forall y)(\exists z)(\forall w)\ ((w \in z) \leftrightarrow (w = x) \vee (w = y))$$

Given sets x, y we can form $\{x, y\}$.
Remark. Using the pairing axiom twice we can form ordered pairs

$$(x, y) = \{\{x\}, \{x, y\}\}$$

Also, form

$$0 = \emptyset, 1 = \{\emptyset\}, 2 = \{0, 1\}.$$

But – we can not form $3 = \{0, 1, 2\}$ using only the axioms above.

4. *Union axiom*

$$(\forall A)(\exists B)(\forall x)\ ((x \in B) \leftrightarrow (\exists z)((z \in A) \wedge (x \in z))$$

For any set A we want to form the set $B = \bigcup A = \bigcup\{z \mid z \in A\}$.

5. *Power set axiom*

$$(\forall A)(\exists B)(\forall z)\ ((z \in B) \leftrightarrow (z \subseteq A))$$

where

$$z \subseteq A \text{ means } (\forall y)((y \in z) \rightarrow (y \in A).$$

For any set A we form its powers set $\mathbb{P}(A)$.

6. *Axiom scheme for specification* Suppose $P(x, y_1, \ldots, y_r)$ is a formula in our language. Then we have an axiom:

$$(\forall A)(\forall y_1)\ldots(\forall y_r)(\exists B)(\forall x)\ ((x \in B) \leftrightarrow ((x \in A) \wedge P(x, y_1, \ldots, y_r)).$$

So this guarantees that we can for the subset of $B \subseteq A$, $B = \{x \in A \mid P(x, y_1, \ldots, y_n)\}$ for any given set A and any given $y_1, \ldots, y_n$.
Example. a) Let C be any non empty set and $A \in C$. Then

$$\bigcap C = \{x \in A \mid (\forall z)((z \in C) \rightarrow (x \in z))\}$$

Where $(\forall z)((z \in C) \rightarrow (x \in z))$ is a formula $P(x, C)$.

b)

$$A \times B = \{w \in \mathbb{P}(\mathbb{P}(A \cup B)) \mid (\exists a)(\exists b)((a \in A) \wedge (b \in B) \wedge w = \{\{a\}, \{a, b\}\})\}.$$

7. *Axiom of infinity*
Definition. 3.8. For a set a the *successor* of a is

$$a^+ := a \cup \{a\}$$

A set A is *inductive* if

$$(\emptyset \in A) \wedge (\forall x)((x \in A) \rightarrow (x^+ \in A)).$$

Now, the *axiom of infinity* is written as the following

$$(\exists A)((\emptyset \in A) \wedge (\forall x)((x \in A) \rightarrow (x^+ \in A)))$$

guaranteeing that there exists an inductive set.

Definition. 3.9. Let A be an inductive set. We can form (using specification) the set

$$\mathbb{N} = \{x \in A \mid \text{ if } B \text{ is an inductive set, then } x \in B\}.$$

(Informally this is the intersection of all inductive sets.) This does not depend on the choice of A. Also denote this set be ω.

Theorem. 3.10. *a) $\mathbb{N}$ is an inductive set, and if B is an inductive set, then $\mathbb{N} \subseteq B$.*

b) (Proof by induction works) Suppose $P(x)$ is a property of sets (i.e. a formula) such that

i. $P(\emptyset)$ holds

ii. For every $k \in \mathbb{N}$, if $P(k)$ holds, then $P(k^+)$ holds.

Then $P(n)$ holds for all $n \in \mathbb{N}$.

Proof. a) Exercise, use definition.

b) Consider $B \subseteq \mathbb{N}$ given by $B = \{x \in \mathbb{N} \mid P(x)\text{holds}\}$. The conditions i, ii tell us that B is an inductive set, thus it contains $\mathbb{N}$. Therefore $B = \mathbb{N}$.

□

Now, we could develop arithmetic in $\mathbb{N}$ (using n^+ as $n+1$,etc.), $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ 'in the natural way' using the axioms above.

Exercise (Hard exercise). For $m, n \in \mathbb{N}$ write $m \leq n$ to mean $(m = n) \vee (m \in n)$. This is a well-ordering on $\mathbb{N}$.

3.4 Well-orderings

Definition. 3.11. A linear ordering $(A; \leq)$ is a *well ordering* (or a *well ordered set*) if every non-empty subset of A has a *least element.*

$$(\forall X)((X \subseteq A) \wedge (X \neq \emptyset) \rightarrow (\exists x)((x \in X) \wedge (\forall y \in X)(x \leq y).$$

Example. $(\mathbb{Z}; \leq)$ not a well ordering with the usual $\leq$. But $(\mathbb{N}; \leq)$ is.

Definition. 3.12. Suppose $\mathcal{A}_1 = (A_1; \leq_1), \mathcal{A}_2 = (A_2; \leq_2)$ are linear ordered sets. Say that $\mathcal{A}_1, \mathcal{A}_2$ are *similar* (or *isomorphic*) if there is a bijection $\alpha : \mathcal{A}_1 \rightarrow \mathcal{A}_2$ such that

$$\forall a, b \in A_1 \ a \leq_1 b \Leftrightarrow \alpha(a) \leq_2 \alpha(b)$$

Write $\mathcal{A}_1 \simeq \mathcal{A}_2$. Say α is a *similarity.* If α suffices the condition except for being bijective, then call it *order-preserving.*

Definition. 3.13. 1. The *reverse-lexicographic product*

$$\mathcal{A}_1 \times A_2 = (A_1 \times A_2; \leq)$$

where

$$(a_1, a_2) \leq (a_1', a_2') \Leftrightarrow (a_2 \leq a_2' \text{ or } ((a_2 = a_2') \wedge (a_1 \leq a_1'))$$

2. *Sum*: Regard A_1, A_2 as disjoint (w.l.o.g) and define

$$\mathcal{A}_1 + \mathcal{A}_2 = (A_1 \cup A_2; \leq)$$

where $\leq$ is the union of $\leq_1, \leq_2$ together with $a_1 \leq a_2$ for all $a_1 \in A_1, a_2 \in A_2$.

Example. 1. $\mathbb{N} + \mathbb{N}$ is just two copies of $\mathbb{N}$ put one after the other.

2. $\{0, 1\} \times \mathbb{N} \simeq \mathbb{N}$

3. $\mathbb{N} \times \{0, 1\} \simeq \mathbb{N} + \mathbb{N}$

Lemma. 3.14. *1. $\mathcal{A}_1 + \mathcal{A}_2$ and $\mathcal{A}_1 \times \mathcal{A}_2$ are linearly ordered sets.*

2. If $\mathcal{A}_1, \mathcal{A}_2$ are well-orderings then so are $\mathcal{A}_1 + \mathcal{A}_2$ and $\mathcal{A}_1 \times \mathcal{A}_2$

Proof. We show that $\mathcal{A}_1 \times \mathcal{A}_2$ is a well-ordering if $\mathcal{A}_1, \mathcal{A}_2$ are. The rest is left as exercise. Let $\emptyset \neq X \subseteq A_1 \times A_2$. Let

$$Y = \{b \in A_2 \mid \text{ there is } a \in A \text{ with } (a, b) \in X\} \subseteq A_2$$

Let y be the least element in Y. Let

$$Z = \{z \in A_1 \mid (z, y) \in X\}$$

This has a least element x. Then (x, y) is the least element of X. □

Definition. 3.15. Suppose $\mathcal{A} = (A, \leq)$ is a linearly ordered set. A subset $X \subseteq A$ is an *initial segment* of A if

$$(\forall x \in X)(\forall a \in A)((a \leq x) \rightarrow (a \in X)).$$

It is *proper* if $X \neq A$.

Example. 1. Let $b \in A$ and let

$$A[b] = \{a \in A \mid a < b\}$$

is a proper initial segment

2. $$A[\leq b]\{x \in A \mid x \leq b\}$$

is an initial segment.

Lemma. 3.16. *If $\mathcal{A} = (A; \leq)$ is a well ordered set, then every proper initial segment X is of the form $A[b]$ for some $b \in A$.*

Remark. This is not true in general e.g. $\{x \in \mathbb{Q} \mid x \leq \pi\}$

Proof. Let b be the minimal element of $A \setminus X$. □

Theorem. 3.17. *Suppose $\mathcal{A}_1 = (A_1, \leq_1), \mathcal{A}_2 = (A_2, \leq_2)$ are well ordered sets. Then exactly one of the following is true:*

1. *$\mathcal{A}_1, \mathcal{A}_2$ are similar;*
2. *$\mathcal{A}_1$ is similar to a proper segment of $\mathcal{A}_2$;*
3. *$\mathcal{A}_2$ is similar to a proper segment of $\mathcal{A}_1$.*

In each case the similarity is unique.

Proof. First show: (for uniqueness) Suppose we have order preserving

$$\beta, \alpha : \mathcal{A}_1 \to \mathcal{A}_2$$

whose images are initial segments of $\mathcal{A}_2$. Show $\alpha = \beta$.

Step 1 Check that if $b \in A_1$ then $\alpha(\mathcal{A}[b]) = \mathcal{A}_2[\alpha(b)]$.

Step 2 If $\alpha \neq \beta$ take $b \in A_1$ minimal with $\alpha(b) \neq \alpha(b)$. So $\alpha \restriction \mathcal{A}_1[b] = \beta \restriction \mathcal{A}_1[b]$. By step 1

$$\mathcal{A}_2(\alpha(b)] = \mathcal{A}_2[\beta(b)]$$

and concludes $\alpha(b) = \beta(b)$

This shows

- By taking $\mathcal{A}_1 = \mathcal{A}_2$ and $\alpha = \mathrm{id}_{\mathcal{A}_1}$, then $\mathcal{A}_1$ is not similar to a proper initial segment of itself.
- It follows that we cannot have two of 1.,2.,3. holding.

Then show: (for the existence) Suppose $\mathcal{A}_2$ is not similar to an initial segment of $\mathcal{A}_1$. Show $\mathcal{A}_1$ is similar to a proper initial segment of $\mathcal{A}_2$. Look at

$$C = \{c \in A_1 \mid \text{ there is a similarity between } \mathcal{A}_1[\leq c] \text{ and an initial segment of } \mathcal{A}_2.$$

If $c \in C$ there is a unique $\alpha_c : \mathcal{A}_1[\leq c] \to \mathcal{A}_2$ with the image being an initial segment of $\mathcal{A}_2$ (by the uniqueness part. **Note:** C is an initial segment of $\mathcal{A}_1$. If $c_1 < c_2 \in C$, then α_{c_1} is the restriction of α_{c_2} to $\mathcal{A}_1[\leq c_1]$.Let $\alpha = \bigcup\{\alpha_c \mid c \in C\}$. Then α is a similarity between C and an initial segment of $\mathcal{A}_2$. If $C = \mathcal{A}_1$ we are done. Otherwise, let a be the minimal element of $\mathcal{A}_1 \setminus C$. Now, $\alpha(C) \neq \mathcal{A}_2$ since then $\mathcal{A}_2 \simeq C$. So $\alpha(C) = \mathcal{A}_2[b]$ for some $b \in A_2$. Now we can extend α by sending $a \mapsto b$ and get that $a \in C$ which is a contradiction.

□

Remark. In the notation of the theorem above, we have, in particular, an injective function

$$A_1 \to A_2 (\text{cases 1), 2)})$$

or an injective function

$$A_2 \to A_1$$

. So either $|A_1| \leq |A_2|$ or $|A_2| \leq |A_1|$. The axiom of choice will imply that every set can be well-ordered. In particular, every two sets are comparable by size and $\leq$ is a total ordering on the sets.

3.5 Ordinals

The *ordinals* are special well-ordered sets, generalizing the notion of natural number:

$$0 = \emptyset$$

$$1 = \{0\}$$

$$2 = \{0, 1\}$$

$$\vdots$$

For $m, n \in \omega$ then $m < n \Leftrightarrow m \in n$.

Definition. 3.18. 1. A set X is *transitive* if every element of X is a subset of X. (I.e. $y \in x, x \in X$ then $y \in X$.)

2. A set α is an *ordinal* if
   - α is a transitive set;
   - the relation $<$ on α given by
     $$x < y \Leftrightarrow x \in y$$
     is a strict well-ordering of α

Remark. If α is an ordinal, then by definition $\alpha \notin \alpha$.

Lemma. 3.19. *If α is an ordinal then so is $\alpha^+ = \alpha \cup \{\alpha\}$.*

Proof. As α is transitive, so is α^+. The ordering $\in$ on α^+ is just the same as on α with the extra element α added as the greatest element. □

Corollary. 3.20. *If $n \in \omega$ then n is an ordinal.*

Proof. $\emptyset$ is an ordinal, so this follows from the previous lemma with induction. (3.10 □

Proposition. 3.21. *1. If α is an ordinal then $\alpha \notin \alpha$.*

2. If α is an ordinal and $\beta \in \alpha$ then β is an ordinal.

3. If α, β are ordinals and $\alpha \subsetneqq \beta$ then $\alpha \in \beta$.

4. If α is an ordinal then $\alpha = \{\beta \mid \beta$ is an ordinal and $\beta \in \alpha\}$

Proof. 1. definition implies this.

2. same as above.

3. As $\alpha \subsetneqq \beta$ so $\beta \setminus \alpha \neq \emptyset$ so it has a least element γ. Show that $\gamma = \alpha$.

4. by 2)

All are in more detail on blackboard. (As is a big part of this subsection in general. □

Definition. 3.22. If α, β are ordinals, write

$$\alpha < \beta$$

to mean

$$\alpha \in \beta$$

.

Theorem. 3.23. *If α, β, γ are ordinals.*

1. *If $\alpha < \beta$ and $\beta < \gamma$ then $\alpha < \gamma$.*
2. *If $\alpha \leq \beta$ and $\beta \leq \alpha$, then $\alpha = \beta$.*
3. *Exactly one of $\alpha < \beta$, $\alpha = \beta$ and $\beta < \alpha$ holds.*
4. *If X is an non-empty set of ordinals then X has a least element δ. Moreover, $\delta = \bigcap X$. "The collection of ordinals is well-ordered".*

Proof.

1. The first one is easy to check.
2. By the first and $\alpha \notin \alpha$.
3. Show that, if $\alpha \neq \beta$ then $\alpha \subsetneqq \beta$ or $\beta \subsetneqq \alpha$. Consider

   $$\alpha \cap \beta$$

   and show that this is an ordinal. If $\alpha \not\subsetneqq \beta$ then $\alpha \cap \beta \subsetneqq \alpha$ and the symmetric case. By 3.21 we get that $\alpha \cap \beta \in \alpha$ and $\alpha \cap \beta \in \beta$. Thus $\alpha \cap \beta \in \alpha \cap \beta$ which is a contradiction.
4. See notes.

□

Corollary. 3.24.

1. *If X is a non-empty set of ordinals, then $\bigcup X$ is an ordinal.*
2. *ω is an ordinal.*

Proof.

1. By 3.23, X is well-ordered by $\in$. So it is enough to show that $\bigcup X$ is a transitive set (exercise).
2. Check that $\bigcup \omega = \omega$; then apply the first part.

□

Remark. Now we can form some new ordinals by taking succesors:

$$\omega^+ = \{0, 1, \ldots, \omega\}$$

$$(\omega^+)^+ = \{0, 1, \ldots, \omega, \omega^+\}$$

$$\vdots$$

Theorem. 3.25. *If $(A; \leq)$ is a well-ordered set, then there is a **unique** ordinal α which is similar to $(A; \leq)$.*

Proof. We have to show uniqueness and existence.

Uniqueness: Suppose $(A; \leq)$ is similar to ordinals α, β. So α, β are similar. W.l.o.g. $\alpha \leq \beta$, so if $\alpha \neq \beta$ then $\alpha < \beta$. α is therefore a proper initial segment of β (3.21). But by 3.17 no well-ordered set is similar to a proper initial segment of itself. Contradiction.

Existence: Let

$$X = \{x \in A \mid \text{ the initial segment } A[x] \text{ is similar to an ordinal}\}.$$

Note: If $A \neq \emptyset$ then the least element of A is in X. By uniqueness above, if $x \in X$ there is a unique ordinal α_x similar to $A[x]$. Let

$$S = \{\alpha_x \mid x \in X\}$$

Claim: S is an ordinal. We need to show that S is a transitive set (since previous results already say that it is well ordered). I.e. we need to show that every $\beta \in \alpha_x \in S$ is in fact an element of $S \ni b$. Let $\theta : A[x] \to \alpha_x$ be the similarity. Let $A \ni y\theta^{-1}(\beta)$. Then θ restricted to $A[y]$ gives a similarity $\theta \restriction A[y] : A[y] \to \{\delta \in \alpha_x \mid \delta < \beta\}$. Therefore $\beta = \alpha_y$, which proves the claim. (Every ordinal is the set of smaller ordinals.) Denote S by α (as it is an ordinal by the claim above.) If we know $X = A$ we are done, since then

$$x \mapsto \alpha_x \text{gives a similarity} A \to \alpha.$$

By what we already did, X is an initial segment. So if $X \neq A$ it is proper i.e. there is a $z \in A \setminus X$ with $X = A[z]$ (3.17). We know that $x \mapsto \alpha_x$ gives a similarity $X \to \alpha$. So $z \in X$, by definition of X. That is a contradiction. Thus $X = A$ and therefore $A \simeq \alpha$.

□

Remark. Why is S a set? We need an extra axiom to justify this – the axiom of replacement.

Definition. 3.26. Suppose $F(x, y, z_1, \ldots, z_r)$ is a formula (in our first order language) with the following property:
whenever $s_1, \ldots, s_r$ are sets and b is a set, there is a unique set a such that $F(a, b, s_1, \ldots, s_r)$ holds. The $s_1, \ldots, s_r$ are referred to as *parameters*, and $F(x, y, s_1, \ldots, s_r)$ gives us a "function on sets" $b \mapsto a$. F is called an *operation on sets* (with $z_1, \ldots, z_r$ being *parameter variables*).

Remark. This is very similar to the definition of a function. But there we need the domain to be a set, and the set of all sets does not fulfill this condition.

Example. 1. No parameter variables: $F(a, b)$ says "a is the power set of b".

2. $F(a, b, c)$ (with parameter c) says "a is the set of all functions from b to c".

Definition. 3.27 (Axiom 8 of the ZF axioms)**.** Suppose $F(x, y, z_1, \ldots, z_r)$ is an operation on sets as above and $s_1, \ldots, s_r$ are sets and B is a set. Then there is a set A such that

$$A = \{a \mid F(a, b, s_1, s \ldots, s_r) \text{ holds for some } b \in B\}.$$

Remark. In the proof of (3.25 we obtain S from X using the operation $F(a, b)$ which says "either b is a well-ordered set similar to the ordinal a or b is not a well-ordered set and $a = \emptyset$. Apply this to $B = \{A[x] \mid x \in X\}$ to get S.

3.6 Transfinite induction

Theorem. 3.28 (Transfinite induction)**.** *Suppose $P(x)$ is a property of sets. Assume that for all ordinals α*

$$\text{if } P(\beta) \text{ holds for all ordinals } \beta < \alpha, \text{ then } P(\alpha) \text{ holds.} \tag{1}$$

Then $P(\gamma)$ holds for all ordinals γ.

Proof. Note that if $\alpha = 0 = \emptyset$ then $P(\beta)$ holds for all $\beta < \alpha$. So by (1) $P(0)$ holds. Suppose for a contradiction that there is some ordinal γ such that $P(\gamma)$ does not hold. There is a least such γ: call it α. Thus for any ordinal $\beta < \alpha$ $P(\beta)$ holds. But then by (1) $P(\alpha)$ holds – contradiction. □

Theorem. 3.29. *Suppose α is an infinite ordinal (i.e. $\omega < \alpha$). Then $\alpha \approx \alpha \times \alpha$.*

Corollary. 3.30.

1. *If $(A; \leq)$ is an infinite well-ordered set, then $|A| = |A \times A|$.*
2. *Assuming the axiom of choice (or the equivalent statement that every set can be well ordered), if A is any infinite set then $|A| = |A \times A|$.*

Proof.

1. By (3.25) there is an ordinal α similar to A and the result is an immediate consequence of (3.29).
2. Since A can be well ordered under our assumptions this is an immediate consequence itself.

□

Proof of the theorem.

1. Assume: if $\omega \neq \beta < \alpha$ then $\beta \approx \beta \times \beta$. We want to show the statement via transfinite induction.
2. We may assume: if $\beta < \alpha$ then $|\beta| < |\alpha|$. Implies $|\beta^+| < |\alpha|$.
3. Enough to show $\alpha \approx \alpha \times \alpha$ by an earlier result.

□

Step 1: Prove the following.
Suppose we have a well ordering $\leq$ of $A = \alpha \times \alpha$ such that for all $x \in A$

$$|A[x]| < |\alpha|.$$

Then $|\alpha| = |\alpha \times \alpha|$.

Proof. By 3.25 there is an ordinal γ which is similar to $(A; \leq)$. Let $f : \gamma \to A$ be the similarity. Show: $\gamma \subseteq \alpha$.
Let $\mu \in \gamma$ so $\mu < \gamma$. As f is a similarity, it gives a bijection:

$$\mu = \{\delta \in \gamma \mid \delta < \mu\} \to A[f(\mu)].$$

Thus $|\mu| = A[f(\mu)] < |\alpha|$. Thus $\mu < \alpha$ and furthermore $\mu \in \alpha$. □

Step 2: Find an ordering $\leq$ on $\alpha \times \alpha = A$ as in step 1.

Proof. For $\lambda < \alpha$ let

$$A_\lambda = \{(\theta, \zeta) \in \alpha \times \alpha \mid \max(\theta, \zeta) = \lambda\}.$$

Define $\leq$ on A by:

$$(\theta', \zeta') < (\theta, \zeta) \leftrightarrow \max(\theta', \zeta') < \max(\theta, \zeta)$$

or

$$\max(\theta', \zeta') = \lambda = \max(\theta, \zeta)$$

and one of the following

1. $\zeta = \zeta' = \lambda$ and $\theta' < \theta$
2. $\zeta' < \zeta = \lambda$
3. $\theta = \theta' = \lambda$ and $\zeta < \zeta'$

Check: $\leq$ is a well-ordering of A. **show:** property in step 1 holds. Let $x = (\theta, \zeta) \in A$. Let $\lambda = \max(\theta, \zeta)$; may assume $\lambda \geq \omega$. Let $\mu = \lambda^+$ so $\mu < \alpha$ and by the induction hypothesis

$$|\mu \times \mu| = |\mu| < |\alpha|.$$

$$\{y \in A \mid y < x\} \subseteq \{(\theta', \zeta') \in A : \max(\theta', \zeta') \leq \lambda\} \subseteq \mu \times \mu.$$

Therefore

$$|A[x]| \leq |\mu \times \mu| < |\alpha|.$$

□

3.7 Transfinite recursion

Allows us to construct, for ordinals α, sets $G(\alpha)$ so that $G(\alpha)$ is obtained from sets $G(\beta)$, $\beta < \alpha$ by applying an operation F.

$$G(0), G(1), \ldots, G(\omega), G(\omega^+), \ldots, G(\beta), \ldots (\beta < \alpha)$$

$$\xrightarrow{F} G(\alpha)$$

$$G \restriction \alpha : \alpha \to \{G(\beta) \mid \beta < \alpha\}$$

where the former is a set, by the axiom of replacement. **Notation:** F is an operation on sets. Denote by $F(b)$ the result of applying F to the set b.

Theorem. 3.31 (Transfinite recursion)**.** *Suppose F is an operation. Then there is an operation G such that for all ordinals α we have*

$$G(\alpha) = F(G \restriction \alpha).$$

If G' is another such operation then

$$G(\alpha) = G'(\alpha)$$

for all ordinals α.

Proof. Omitted. Can be found on blackboard. □

Remark. In practice, we usually do not write F down as a first order formula.

Example. Application:
Lindenbaum-Lemma: Suppose $\mathcal{L}$ is a first order language whose alphabet of symbols can be well-ordered. Suppose Σ is a consistent set of closed $\mathcal{L}$-formulas. Then there is a consistent set $\Sigma^* \supseteq \Sigma$ of closed $\mathcal{L}$-formulas such that for every closed $\mathcal{L}$ formula ψ either

$$\Sigma^* \vdash \psi$$

or

$$\Sigma^* \vdash (\neg\psi)$$

Proof. The set of $\mathcal{L}$-formulas can be well-ordered. Any subset of a well-ordered set us well-ordered so the set of closed $\mathcal{L}$-formulas is well-ordered. Any well-ordered set is similar to some ordinal. (3.25) Thus we can write the set of closed $\mathcal{L}$-formulas as

$$\{\phi_\alpha \mid \alpha < \lambda\}$$

for some ordinal λ. Define for any ordinal α a consistent set $G(\alpha) \geq \Sigma$ of closed $\mathcal{L}$-formulas

$$G(\alpha) = \begin{cases} \Sigma \cup \dots \\ \Sigma \cup \dots \end{cases}$$

□

Remark. Can use similar arguments at other parts in the proof of the completeness theorem to get that completeness holds. The same for compactness etc.

Definition. 3.32 (*Axiom of regularity(Foundation)*)**.** ZF9:

$$(\forall x)((x \neq \emptyset) \to (\exists a)((a \in x) \wedge (a \cap x = \emptyset))$$

In particular, there is no set b with $b \in b$. We will not use this axiom.

Definition. 3.33. The 9 axioms we wrote down until now, are the axioms for *Zermelo-Fraenkel set theory.*

4 Axiom of Choice and Consequences

4.1 Statement and well-orderings

Definition. 4.1. The *Axiom of Choice (AC)* is the formula stating the following:
Suppose A is a set of non-empty sets. Then there is a function

$$f : A \to \bigcup A$$

such that $f(a) \in a$ for every $a \in A$.

Definition. 4.2. The set of the Zermelo-Fraenkel axioms together with the Axiom of Choice is denoted by ZFC.

Example. Suppose X is any non-empty set. Let

$$A = \mathbb{P}(X) \setminus \emptyset.$$

By AC there is a function

$$f : A \to X, f(Y) \in Y \text{ for all } \emptyset \neq Y \subseteq X.$$

Such a function is called a *choice function* (on the subsets of X).

Remark. If $(X; \leq)$ is a well-ordered set we do not need the axiom of choice to get a choice function on X: Define for $\emptyset \neq Y \subseteq X$

$$f(Y) = \min(Y).$$

$$f = \{(Y, a) \in (\mathbb{P} \setminus \emptyset) \times X \mid a = \min(Y)\}$$

Theorem. 4.3. *Suppose X is a non-empty set and $f : \mathbb{P} \setminus \emptyset \to X$ is a choice function. Then there is a well-ordering $\leq$ of X, i.e. a well ordered set $(X; \leq)$*

Proof. Chose one element of the set. Declare it as your smallest element. Then chose from the remaining elements, etc. For a more formal and exact proof we use the following Lemma:

Lemma. 4.4 (Hartogs' Lemma)**.** *For any set X there is an ordinal α such that there is no injective function*

$$h : \alpha \to X$$

More formally, define $\tilde{X} = X \cup \{\infty\}$. Use transfinite recursion and the choice function to define an operation G. For an ordinal γ:

$$G(\gamma) = \begin{cases} f(X \setminus \{G(\beta) \mid \beta < \gamma\} \text{ if } X \setminus \{G(\beta) \mid \beta < \gamma\} \neq \emptyset \\ \infty \text{ otherwise.} \end{cases}$$

Note: If $\infty \notin (G \restriction \gamma)$ then $G \restriction \gamma$ is an injective function $\gamma \to X$. By Hartogs' Lemma there is an ordinal α such that $G(\alpha) = \infty$. Take the least such ordinal and call in β. Then

$$g = G \restriction \beta : \beta \to X$$

is injective and surjective. So $g : \beta \to X$ is a bijection. Define $\leq$ on X by

$$x_1 \leq x_2 \Leftrightarrow g^{-1}(x_1) \leq g^{-1}(x_2)$$

$\square$

Proof of Hartogs' Lemma. Given A any set, find an ordinal α with no injective function $\alpha \to X$. Consider:

$$Y = \{(Z; \leq_Z) \mid Z \subseteq X \text{ and } \leq_| \text{ is a well-ordering on } Z\}.$$

This is a set by specification. Let

$$S = \{\beta \mid \beta \text{ is an ordinal similar to some}(Z; \leq_Z) \in Y\}$$

SO

$$S = \{\beta \mid \ \beta \text{ is an ordinal and there is an injective function } \beta \to X\}.$$

Let $\sigma = \bigcup S$. This is an ordinal and $\beta < \sigma \ldots$ $\square$

Corollary. 4.5 (Assuming ZF)**.** *AC is equivalent to to the following* Well-Ordering Principle *(WO): If A is any set there is $\leq_A \subseteq A \times A$ such that $(A; \leq_A)$ is a well-ordered set.*

Corollary. 4.6 (Assuming ZFC)**.**

1. *If A is any set there s an ordinal α with $\alpha \approx A$.*
2. *If A, B are sets then one of*
$$|A| \leq |B| \quad |B| \leq |A|$$
holds.
3. *(Fundamental Theorem of Cardinal Arithmetic – FTCA) If A is any infinite set, then $|A \times A| = |A|$.*

Lemma. 4.7 (Assuming ZFC)**.** *For sets A, B with $A \neq \emptyset$ we have*

$$|A| \leq |B| \leftrightarrow \text{ there is a surjective function } h : B \to A.$$

Proof. Was in the problem class. $\square$

4.2 Cardinals and Cardinal Arithmetic

Assume ZFC.

Definition. 4.8. An ordinal α is a *cardinal* if it is not equinumerous with any $\beta < \alpha$. (E.g. ig $n \in \omega$ then n is a cardinal; if β is infinite then β^+ is not a cardinal as $\beta \approx \beta^+$.)

Lemma. 4.9.

1. *Suppose γ is any ordinal. There is a (unique) cardinal α with $\alpha \approx \gamma$.*
2. *Suppose A is any set. There is a unique cardinal α such that $A \approx \alpha$.*

Proof.

1. Take the least element of
$$\{\beta \leq \gamma \mid \beta \approx \gamma\}.$$
2. We know that there is an ordinal γ with $\gamma \approx A$. Use 1. .

$\square$

Definition. 4.10. The unique cardinal equinumerous with A is called the *cardinality* of A. Denote it by $\mathrm{card}(A)$ or $|A|$.

Remark. Note that this notation coincides with the previous usage of $|A|$. There is an injective function $f : A \to B$ iff $\mathrm{card}(A) \leq \mathrm{card}(B)$

Remark. If α is an ordinal then α is a cardinal iff $|\alpha| = \alpha$. E.g. if A is a countably infinite set, then $|A| = \omega$.

Definition. 4.11. Suppose A, B are disjoint sets with $|A| = \kappa$ and $|B| = \lambda$. Let $\kappa + \lambda$ be $|A \cup B|$ and $\kappa \cdot \lambda$ be $|A \times B|$.

Remark. This does not depend on the choice of A, B.

Theorem. 4.12. *Suppose κ, λ are cardinals, $\kappa \leq \lambda$ and λ is infinite. Then*

1. $\kappa + \lambda = \lambda$.
2. $\kappa \cdot \lambda = \lambda$ *(if* $\kappa \neq \emptyset$.

Proof. 1. We use part 2.:

$$\lambda \leq \kappa + \lambda \leq \lambda + \lambda = 2 \cdot \lambda = \lambda$$

2. As $\kappa \subseteq \lambda$ we have $\kappa \times \lambda \subseteq \lambda \times \lambda$. Thus $\kappa \cdot \lambda \leq \lambda$ (By FTCA 3) Since $\kappa \neq \emptyset$ obviously $\lambda \leq |\kappa \times \lambda| = \kappa \cdot \lambda$.

□

Theorem. 4.13. *Suppose A is an infinite set of cardinality λ. Suppose each element of A is a set of cardinality $\leq \kappa$. Then*

$$|\bigcup A| \leq \lambda \cdot \kappa.$$

Proof. We can assume $\emptyset \notin A$. For each $a \in A$ the set S_a of surjective functions $\kappa \to a$ is non-empty. Assuming AC there is a function $F : A \to \bigcup_{a \in A} S_a$ with

$$F(a) : \kappa \to a \text{ a surjective function.}$$

Denote this function by F_a. Let $h : \lambda \to A$ be a bijection. Now define $G : \lambda \times \kappa \to \bigcup A$ by

$$G(\alpha, \beta) = F_{h(\alpha)}(\beta).$$

This is surjective. So by an earlier result

$$|\lambda \times \kappa| \geq |\bigcup A|$$

i.e.

$$|\bigcup A| \leq \lambda \cdot \kappa.$$

□

Example. 1. Suppose A is an infinite set, $|A| = \lambda \geq \omega$. Let S be the set of finite sequences of elements of A. So

$$S = \bigcup_{n \in \omega} A^n.$$

Then $|S| = |A| = \lambda$.

Proof. For $n \in \omega$, $|A|^n = |A|$ by repeated use of FTCA. Then by the result above $|S| \leq \omega \cdot |A| = \lambda$. $A \subseteq S$, so $|S| \geq \lambda$. □

2. Consider $\mathbb{R}$ as a vector space over $\mathbb{Q}$. Suppose $B \subseteq \mathbb{R}$ such that the linear span of B (over $\mathbb{Q}$) is $\mathbb{R}$. Then $|B| = |\mathbb{R}|$. Obviously, $|B| \leq |\mathbb{R}|$. If $x \in \mathbb{R}$ there are $n \in \mathbb{N}$ such that for some $q_0, \ldots, q_n \in \mathbb{Q}$ and $b_0, \ldots, b_n \in B$ with

$$x = \sum_{i=0}^{n} q_i b_i.$$

Let S_1 be the set of finite sequences from $\mathbb{Q}$ and S_2 the set of all sequences of B. Consider $S_1 \times S_2 \supseteq T = \{((q_0, \ldots, q_n), (b_0, \ldots, b_n)\}$. We have a surjection

$$T \to \mathbb{R} \quad ((q_0, \ldots, q_n), (b_0, \ldots, b_n) \mapsto \sum_{i=0}^{n} q_I b_i$$

. By cardinal arithmetic we have

$$|\mathbb{R}| \leq |T| \leq |S_1 \times S_2| = \omega \cdot |B|$$

Since $|\mathbb{R}| > \omega$ we have $|B| = |\mathbb{R}|$.

4.3 Zorn's Lemma

Definition. 4.14. A *partially ordered set* (poset) $(A; \leq)$ satisfies

$$\forall x, y, z \in A \quad x \leq y \leq z \to x \leq z$$

$$\wedge (x \leq y \wedge y \leq x) \to x = y$$

$$\wedge (x \leq x).$$

A *chain* C in a poset $(A; \leq)$ is a subset $C \subseteq A$ such that

$$\forall x, y \in C \quad (x \leq y) \vee (y \leq x)$$

An *upper bound* of C in A is $a \in A$ such that

$$\forall x \in C \quad x \leq a.$$

Example. $(\mathbb{P}(X); \subseteq)$ is a poset. If $C \subseteq \mathbb{P}(X)$ is a chain than $\bigcup C$ is an upper bound for C in $\mathbb{P}(X)$.

Definition. 4.15. *Zorn's Lemma* is: Suppose $(A; \leq)$ is a non-empty poset in which every chain has an upper bound in A. Then A has a maximal element.

Theorem. 4.16. ZF $\vdash$ (AC $\leftrightarrow$ ZF) *or stated differently:*

1. *Assuming ZFC, then ZL holds.*
2. *Assuming ZF and ZL, then AC holds.*

Proof. Let $f : \mathbb{P}(A) \setminus \{\emptyset\} \to A$ be a choice function. Suppose for a contradiction that A has no maximal element. Let $C \subseteq A$ be a chain. By assumption there is an upper bound for C in A. Call it y. Since we assume that A has no maximal element there is some $z \in A$ such that $z > y$. Therefore $z > c$ for all $c \in C$.
Use transfinite recursion to define an operation G such that for all ordinals α

$$G(\alpha) \in A \text{ and } G(\alpha) = f(\{z \in A \mid z > G(\beta) \text{ for all } \beta < \alpha\}).$$

So $G(0) < G(1) < \cdots < G(\beta) < \ldots$ i.e. for all ordinals $\beta < \alpha$

$$G(\beta) < G(\alpha).$$

By Hartogs' Lemma this is impossible. $\square$

Example. Assume ZFC. Suppose V is a vector space over a field F. Then V has a basis over F. Use ZL. Let A be the set of linearly independent subsets of V, ordered by $\subseteq$. Let C be a chain in A – we need to find an upper bound in A. In fact $\bigcup C$ is such an upper bound (exercise). Thus there exists a maximal element B of A. So B is a linearly independent set and if $v \in V \setminus B$ then $B \cup \{v\}$ is not linearly independent. Hence v is a linear combination of elements in B and B spans V.

5 Postscript

Assume ZFC. For a first order language $\mathcal{L}$ the *cardinality* of the set of $\mathcal{L}$ formulas (= cardinality of the alphabet). Assuming ZFC the following hold for $\mathcal{L}$ of arbitrary cardinality:

1. Goedel Completeness Theorem in $\mathcal{L}$ and results leading to it
2. compactness Theorem

(and the same for normal models in $\mathcal{L}^=$.