

M3P65: Mathematical Logic

Lecturers: Dr. Zaniar Ghadernezhad
and
Prof. David Evans *

October 12, 2021

Contents

1	Propositional logic	3
1.1	Propositional formulas and induction	3
1.2	Valuation maps	5
1.3	Substitution	7
1.4	The Deductive approach	10
1.4.1	Formalising deduction	12
1.5	Soundness and Completeness theorems	12
1.6	Consistent and maximally consistent sets of formulae	15
2	Predicate logic	20
2.1	First order structures, languages and formulas	20
2.2	Interpretations of languages and assignments	22
2.3	Variable binding and subformulas	24
2.4	Substitution	27
2.5	Some examples of structures	28
2.6	Variable scoping	29
2.7	Natural Deduction	31
2.8	Completeness and the model existence lemma	33
2.8.1	Extension by constants and Henkin extensions	33
2.8.2	The proof of Lemma 2.35	36
2.8.3	Proof of the model existence lemma	36
2.9	Soundness of equality deduction rules	39
2.10	Some conclusions of the completeness theorem	40
2.11	Decidability and the Entscheidungsproblem	41
2.12	Dense linear orders	42
3	Set theory	44
3.0	Basic set theory (Review)	44
3.1	Cardinality	45
3.2	Axioms for set theory	47
3.3	Linear orderings	49
3.4	Ordinals	50
3.5	Transfinite induction	54
3.6	Transfinite recursion	55
3.7	The axiom of regularity (foundation)	56
3.7.1	The Zermelo-Fraenkel axioms	56
4	The axiom of choice and consequences	58
4.1	Statement of the axiom of choice and the well ordering principle	58
4.2	Cardinals and cardinality	60

* Typeset by Gautam Chaudhuri, source code available [here](#). Any typos and errors are my own, please submit corrections to the repository above.

4.3 Zorn's lemma	61
----------------------------	----

1 Propositional logic

Propositional logic deals with propositions, these are sentences/statements that are either *true* or *false*. It is fairly limited in scope, but serves as a foundation for further theories.

1.1 Propositional formulas and induction

Definition 1.1:

The language of propositional logic consists of

1. P , a nonempty set of propositions which is either finite or infinite, these are called atoms or atomic propositions.
2. Connectives: $\wedge, \vee, \rightarrow, \leftrightarrow, \neg, \perp$
3. Auxiliary symbols: $(,)$

Definition 1.2:

Let P be the set of *atoms* or *atomic propositions*, define $\mathcal{W}(P)$ to be the smallest set with the following properties.

1. $p \in \mathcal{W}(P)$ for all $p \in P$
2. if $p, q \in \mathcal{W}(P)$ then $(p \square q) \in \mathcal{W}(P)$ where $\square \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$.
3. $p \in \mathcal{W}(P) \rightarrow (\neg p) \in \mathcal{W}(P)$
4. $\perp \in \mathcal{W}(P)$

We call elements of $\mathcal{W}(P)$ the *propositional sentences* or *formulas* (built from P).

The language of propositional logic is called *meta-language* and the variables (symbols) are called *meta-variables*.

Example:

P is given, suppose $p, q \in P$ then $p \wedge q \in \mathcal{W}(P)$. Is $\neg p \in \mathcal{W}(P)$? Is $\neg \neg \wedge$ in $\mathcal{W}(P)$?

Proof. $\mathcal{W}(P) \setminus \{\neg p\}$ satisfies all the properties that $\mathcal{W}(P)$ should. Since $\mathcal{W}(P) \setminus \{\neg p\}$ is necessarily smaller, $\mathcal{W}(P)$ is no longer the smallest set that satisfies all the properties. This is a contradiction. The second question is proven similarly. ■

Theorem 1.3: Induction

Suppose A is a property about $\mathcal{W}(P)$, then $A(\phi)$ holds for all $\phi \in \mathcal{W}(P)$ if

1. $A(p)$ holds for all $p \in P$ (the atoms)
2. $A(\phi)$ and $A(\psi) \rightarrow A((\phi \square \psi))$
3. $A(\phi) \rightarrow A((\neg \phi))$
4. $A(\perp)$ holds

Proof. Clearly, the four statements hold for $X := \{\phi \in \mathcal{W}(P) : A(\phi)\}$. Then $X \subseteq \mathcal{W}(P)$ as X satisfies all the properties of $\mathcal{W}(P)$. From this we deduce that $\mathcal{W}(P) \subseteq X$ as $\mathcal{W}(P)$ is the smallest set that satisfies all the properties. Hence $X = \mathcal{W}(P)$. ■

Example:

Every statement in $\mathcal{W}(P)$ has an even number of brackets.

Proof. Show each of the inductive properties

- Every atomic proposition does not have brackets so they have an even number, so property 1 holds.
- Combining two propositions, each with an even number of brackets, will add two brackets to the total. This will result in an even number of brackets, so property 2 holds.
- Negating a statement with an even number of brackets adds two brackets to the total, this is still an even number of brackets so property 3 holds.
- \perp does not have any brackets, so property 4 trivially holds.

So by induction, the property holds for all elements of $\mathcal{W}(P)$. ■

Definition 1.4:

A sequence $\phi_1, \dots, \phi_n \in \mathcal{W}(P)$ is called a formation sequence of ϕ if:

1. $\phi_n = \phi$
2. For all i , at least one of the following holds
 - (a) ϕ_i is atomic
 - (b) $\phi_i = (\phi_j \Box \phi_k); j, k < i$
 - (c) $\phi_i = (\neg \phi_j), j < i$

Theorem 1.5:

Given P , $\mathcal{W}(P)$ is the set of all expressions that have a formation sequence.

Proof. We prove this inductively on ϕ . Let $X = \{\phi \in \mathcal{W}(P) : \phi \text{ has a formation sequence}\}$

- X clearly contains all the atoms
- The combination of two formation sequences using a connective also has a formation sequence by concatenating the two, this implies it is in X .
- The negation of a formation sequence also has a formation sequence and is thus in X .
- \perp is atomic and thus contained in X .

by the inductive theorem, $X = \mathcal{W}(P)$ ■

Example:

What is a formation sequence for $((\neg p) \wedge q)$?

- $p, (\neg p), q, ((\neg p) \wedge q)$
- $p, q, (\neg p), ((\neg p) \wedge q)$
- $p, s, q, r, (p \wedge q), (\neg p), ((\neg p) \wedge q)$

1.2 Valuation maps

Every $p \in P$ has a truth value, 1 or 0. Can we assign truth values to propositions in $\mathcal{W}(P)$?

Definition 1.6:

A mapping $v : \mathcal{W}(P) \rightarrow \{0, 1\}$ is called a *valuation map* if it satisfies the following $\forall \phi, \psi \in \mathcal{W}(P)$

1. $v((\phi \wedge \psi)) = \min\{v(\phi), v(\psi)\}$
2. $v((\phi \vee \psi)) = \max\{v(\phi), v(\psi)\}$
3. $v((\phi \rightarrow \psi)) = 0 \iff v(\phi) = 1 \text{ and } v(\psi) = 0$
4. $v((\phi \leftrightarrow \psi)) = 1 \iff v(\phi) = v(\psi)$
5. $v((\neg \phi)) = 1 - v(\phi)$
6. $v(\perp) = 0$

Lemma 1.7:

Suppose v, v' are two valuation maps s.t. $v(p) = v'(p)$ for all atomic $p \in P$, then $v(\phi) = v'(\phi)$ for all $\phi \in \mathcal{W}(P)$.

Proof, Induction on formulae. We check each case of the formation sequences, let $\phi \in \mathcal{W}(P)$.

- *atoms:* Suppose ϕ is atomic, then by assumption, $v(\phi) = v'(\phi)$.
- *negation:* Suppose $\phi = (\neg \phi_1)$ where $\phi_1 \in P$ is atomic. Then $v(\phi) = 1 - v(\phi_1) = 1 - v'(\phi_1) = v'(\phi)$
- *connectives:* Suppose $\phi = (\phi_1 \square \phi_2)$ where $\phi_1, \phi_2 \in P$ are atomic and \square is a connective. Then
 - If $\square = \wedge$

$$\begin{aligned}
 v(\phi) &= v((\phi_1 \wedge \phi_2)) \\
 &= \max\{v(\phi_1), v(\phi_2)\} \\
 &= \max\{v'(\phi_1), v'(\phi_2)\} \\
 &= v'((\phi_1 \wedge \phi_2)) \\
 &= v'(\phi)
 \end{aligned}$$

- We prove the case for $\square = \vee$ similarly.
- If $\square = \rightarrow$, then

$$\begin{aligned}
 v(\phi) &= v((\phi_1 \rightarrow \phi_2)) \\
 v(\phi) = 0 &\iff v(\phi_1) = 1 \text{ and } v(\phi_2) = 0 \\
 \iff v'(\phi_1) = 1 \text{ and } v'(\phi_2) = 0 &\iff v'((\phi_1 \rightarrow \phi_2)) = 0 \\
 \iff v'(\phi) = 0
 \end{aligned}$$

Since we have $v(\phi) = 0 \iff v'(\phi) = 0$, $v(\phi) = v'(\phi)$.

- The case for $\square = \leftrightarrow$ is clear.
- The two maps must agree on \perp by definition.

We have proven the hypothesis true for every atom, the negation of every atom, and the concatenation of any two atoms by a connective. Thus the hypothesis holds true for any formation sequence, and thus every element $\phi \in \mathcal{W}(P)$. ■

Corollary 1.8:

If $v' : P \cup \{\perp\} \rightarrow \{0, 1\}$ s.t. $v'(\perp) = 0$, then there is a unique valuation map v such that $v(p) = v'(p)$, $\forall p \in P$

Suppose P is a finite set with size n , how many valuation maps do we have for $\mathcal{W}(P)$?

Definition 1.9:

Let P be a given nonempty set of atoms, let $\phi \in \mathcal{W}(P)$.

1. ϕ is a *tautology* if $v(\phi) = 1$ for all valuation maps of $\mathcal{W}(P)$
2. Suppose $\Gamma \subseteq \mathcal{W}(P)$, then ϕ is a *semantic consequence* of Γ , written $\Gamma \models \phi$, if, for every valuation map v where $v(\psi) = 1$ for all $\psi \in \Gamma$, we have $v(\phi) = 1$. For a tautology ϕ , we write $\models \phi$.

Example:

Let P be a set of atomic propositions, v a valuation map on $\mathcal{W}(P)$ and $p, q, p_1, q_1 \in P$.

| : $\models (p \vee (\neg p))$

Proof.

$$\begin{aligned} v((p \vee (\neg p))) &= \max\{v(p), v((\neg p))\} \\ &= \max\{v(p), 1 - v(p)\} \\ &= 1 \end{aligned}$$

■

| : $\models ((p \wedge p) \rightarrow p)$

| : $\models ((p \vee p) \leftrightarrow p)$

| : $\models ((p \wedge q) \leftrightarrow (q \wedge p))$

| : $\models ((\neg(p \wedge q)) \leftrightarrow ((\neg p) \vee (\neg q)))$

| : $\{p\} \models p$

| : $\{p, (p \rightarrow q)\} \models q$

Proof. We note that $v(p) = 1$ and $v((p \rightarrow q)) = 1 \implies v(q) = 1$ (otherwise contradiction)

■

Truth Tables We can also check the veracity of statements using truth tables. Some examples are shown below.

$v(p)$	$v(\neg p)$	$v((p \vee (\neg p)))$	$v(p)$	$v(q)$	$v((p \rightarrow q))$
0	1	1	0	0	1
0	1	1	0	1	1
1	0	1	1	0	0
			1	1	1

Table 1: Truth tables for $(p \vee (\neg p))$ and $(p \rightarrow q)$

Proposition 1.10:

Let $p, q \in P$ some nonempty set of propositions

| **1.10.1:** $\models ((p \rightarrow q) \leftrightarrow ((\neg p) \vee q))$

| **1.10.2:** $\models ((p \leftrightarrow q) \leftrightarrow ((p \rightarrow q) \wedge (q \rightarrow p)))$

Proof, use the valuation map. Let $v : \mathcal{W}(P) \rightarrow \{0, 1\}$ be some valuation map

- We expand

$$v((p \rightarrow q) \leftrightarrow ((\neg p) \vee q)) = 1 \iff v((p \rightarrow q)) = v(((\neg p) \vee q))$$

Note that $v(((\neg p) \vee q)) = \max\{1 - v(p), v(q)\}$. So $v(((\neg p) \vee q)) = 0 \iff v(p) = 1 \text{ and } v(q) = 0 \iff v((p \rightarrow q)) = 0$.

- We expand

$$\begin{aligned}
 v((p \leftrightarrow q) \leftrightarrow ((p \rightarrow q) \wedge (q \rightarrow p))) &= 1 \iff v((p \leftrightarrow q)) = v(((p \rightarrow q) \wedge (q \rightarrow p))) \\
 v(((p \rightarrow q) \wedge (q \rightarrow p))) &= \min\{v((p \rightarrow q)), v((q \rightarrow p))\} \\
 v(((p \rightarrow q) \wedge (q \rightarrow p))) &= 0 \iff v((p \rightarrow q)) = 0 \text{ or } v((q \rightarrow p)) = 0 \\
 &\iff (v(p) = 1 \text{ and } v(q) = 0) \text{ or } (v(q) = 1 \text{ and } v(p) = 0) \\
 &\iff v(p) \neq v(q) \\
 &\iff v((p \leftrightarrow q)) = 0
 \end{aligned}$$

■

This statement also holds for any formula $\phi \in \mathcal{W}(P)$, but we leave this proof until later.

Remark:

The symbols $\rightarrow, \leftrightarrow$ are members of the meta language, the symbols \implies, \iff are logical implications.

1.3 Substitution

Setting Let P be a nonempty set, $n \leq |P|$, $\phi \in \mathcal{W}(P)$ and assume $p_1, \dots, p_n \in P$ are atoms appearing in ϕ (we sometimes call p_1, \dots, p_n variables appearing in $\phi(p_1, \dots, p_n)$).

Let $\psi \in \mathcal{W}(P)$, we write $\phi[\psi/p_i]$ for a proposition obtained by replacing all instances of p_i in ϕ with ψ .

Remark:

Substitution of ψ for p_i defines a map from $\mathcal{W}(P)$ to $\mathcal{W}(P)$.

•

$$\phi[\psi/p_i] = \begin{cases} \phi, & \text{if } \phi \text{ atomic and } \phi \neq p_i \\ \psi, & \text{if } \phi = p_i \end{cases}$$

- If $\phi = (\phi_1 \square \phi_2)$, then

$$\begin{aligned}
 \phi[\psi/p_i] &= (\phi_1 \square \phi_2)[\psi/p_i] \\
 &= (\phi_1[\psi/p_i] \square \phi_2[\psi/p_i])
 \end{aligned}$$

- $\phi = (\neg \phi_1)$

$$\begin{aligned}
 \phi[\psi/p_i] &= (\neg \phi_1)[\psi/p_i] \\
 &= (\neg \phi_1[\psi/p_i])
 \end{aligned}$$

Theorem 1.11: Substitution

If $\models (\phi_1 \leftrightarrow \phi_2)$, then

$$\models (\psi[\phi_1/p] \leftrightarrow \psi[\phi_2/p])$$

Proof. We use induction on formulae

- Suppose that ψ is atomic, either $\psi \neq p$ in which case $\psi[\phi_i/p] = \psi$ and the theorem trivially holds, or $\psi = p$.
 - In the latter case $\psi[\phi_i/p] = \phi_i, i = 1, 2$, by assumption we have $v(\phi_1) = v(\phi_2)$, for all valuation maps v .
 - This implies $v(\psi[\phi_1/p]) = v(\psi[\phi_2/p])$ for all valuation maps.
 - Which implies $\models (\psi[\phi_1/p] \leftrightarrow \psi[\phi_2/p])$.

- Suppose that $\psi = (\neg\psi_1)$, where ψ_1 is atomic.

$$\begin{aligned} v(\psi[\phi_1/p]) &= v((\neg\psi_1[\phi_1/p])) \\ &= 1 - v(\psi_1[\phi_1/p]) \\ &= v((\neg\psi_1[\phi_2/p])) \\ &= v(\psi[\phi_2/p]) \end{aligned}$$

- Suppose that $\psi = (\psi_1 \square \psi_2)$, where ψ_1 and ψ_2 are atomic.

$$\begin{aligned} v(\psi[\phi_1/p]) &= v((\psi_1[\phi_1/p] \square \psi_2[\phi_1/p])) \\ &= v((\psi_1[\phi_2/p] \square \psi_2[\phi_2/p])) \\ &= v(\psi[\phi_2/p]) \end{aligned}$$

■

Exercise:

Use the theorem stated above to show that:

- $\models ((\phi \rightarrow \psi) \leftrightarrow ((\neg\phi) \vee \psi))$
- $\models ((\phi \leftrightarrow \psi) \leftrightarrow ((\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)))$

Corollary:

For every $\phi \in \mathcal{W}(P)$ there is a $\psi \in \mathcal{W}(P)$ such that ψ only has connectives $\{\wedge, \vee, \neg\}$ such that $\models (\phi \leftrightarrow \psi)$

Proof, induction on formulae. We prove a much stronger result below. ■

Definition 1.12:

A *truth function* on n variables is a function $F : \{0, 1\}^n \rightarrow \{0, 1\}$.

Let $\phi \in \mathcal{W}(P)$ and assume ϕ has n variables $p_1, \dots, p_n, n \leq |P|$. Then we obtain from ϕ a truth function $F_\phi : \{0, 1\}^n \rightarrow \{0, 1\}, \bar{x} \mapsto v_{\bar{x}}(\phi)$ where $v_{\bar{x}}(p_i) = x_i, i \leq n$.

Example:

$$\bar{x} = (0, 1, \dots, 1) \implies v_{\bar{x}}(p_1) = 0, v_{\bar{x}}(p_2) = 1, \dots, v_{\bar{x}}(p_n) = 1$$

Remark:

$$\models (\phi \leftrightarrow \psi) \iff F_\phi = F_\psi$$

Definition 1.13:

We say a set A of connectives is *adequate* if for every $1 \leq n \leq |P|$, and for every truth function on n -variables F . There exists some formula $\phi \in \mathcal{W}(P)$ which is only constructed using connectives from A and the variables p_1, \dots, p_n such that $F = F_\phi$.

Theorem 1.14:

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

Proof. Let $F : \{0, 1\}^n \rightarrow \{0, 1\}$

Case 1 $F(\bar{x}) = 0, \forall \bar{x}, F_{\perp} = F_{(p \wedge (\neg p))}$ as $\models (\perp \leftrightarrow (p \wedge (\neg p)))$

Case 2 Let $T = \{\bar{x} \in \{0, 1\}^n : F(\bar{x}) = 1\} = \{\bar{x}_1, \dots, \bar{x}_r\}, \bar{x}_i = (x_i^1, \dots, x_i^n), r \geq 1.$

Let $i \leq r$, define $q_i^j = \begin{cases} p^j & \text{if } x_i^j = 1 \\ (\neg p^j) & \text{if } x_i^j = 0 \end{cases}$

Suppose $q_i^j = p^j$, then $v_{\bar{x}}(q_i^j) = v_{\bar{x}}(p^j) = x_i^j$. If $q_i^j = (\neg p^j)$, then $v_{\bar{x}}(q_i^j) = v_{\bar{x}}((\neg p^j)) = 1 - x_i^j$.

Claim: $v_{\bar{x}}(q_i^j) = 1 \iff v_{\bar{x}}(p^j) = x_i^j$. Let $\psi_i = q_i^1 \wedge \dots \wedge q_i^n$. Then

$$\begin{aligned} F_{\psi_i}(\bar{x}) = 1 &\iff v_{\bar{x}}(\psi_i) = 1 \\ &\iff v_{\bar{x}}(q_i^j) = 1, \forall j \leq n \\ &\iff v_{\bar{x}}(p^j) = x_i^j, \forall j \leq n \\ &\iff \bar{x} = \bar{x}_i \end{aligned}$$

Then let $\theta = \psi_1 \vee \dots \vee \psi_r$

$$\begin{aligned} F_{\theta}(\bar{x}) = 1 &\iff F_{\psi_i}(\bar{x}) = 1, i \leq r \\ &\iff v_{\bar{x}}(\psi_i) = 1 \\ &\iff \bar{x} = \bar{x}_i \\ &\iff \bar{x} \in T \end{aligned}$$

■

Definition 1.15:

Suppose that $\phi \in \mathcal{W}(P)$ and it is of the form of case two in the previous theorem. ϕ is then in *Disjunctive Normal Form (DNF)*.

Corollary 1.16:

Suppose $\phi \in \mathcal{W}(P)$, then there is $\psi \in \mathcal{W}(P)$ in DNF s.t. $\models (\phi \leftrightarrow \psi)$.

Proof. Take $F = F_{\phi}$ and apply the previous theorem. ■

Exercise:

The following set of connectives are adequate,

- $\{\neg, \wedge\}$
- $\{\neg, \vee\}$
- $\{\neg, \rightarrow\}$
- $\{\rightarrow, \perp\}$

1.4 The Deductive approach

Suppose $v : \mathcal{W}(P) \rightarrow \{0, 1\}$ is a valuation function, then this is uniquely determined by knowing the value of $v(p)$, $p \in P$.

Exercise:

Show that $v(\varphi \rightarrow \psi) = 1 \iff v(\varphi) \leq v(\psi)$

If P is a non-empty set of atoms, we can construct the set $\mathcal{W}(P)$, the set of all prop. formulas generated by P . We can take $\Sigma \subseteq \mathcal{W}(P)$ and derive semantic consequences from it $\Sigma \models \varphi$ using a valuation function. We can also define $\Sigma \not\models \varphi \iff \exists$ valuation map v s.t. $v(\psi) = 1, \forall \psi \in \Sigma$ and $v(\varphi) = 0$.

Non-semantic approach We want to deal with logic as a codification of exact reasoning, i.e. deduction rules. The process by which this is done is called *natural* or *formal deduction*. A deduction begins with a set of *premises*, e.g. φ and $\varphi \rightarrow \psi$. Then using a *deduction rule* (in this case *modus ponens*) we can state a *conclusion* (in this case ψ). We write this as $\frac{\varphi \quad \varphi \rightarrow \psi}{\psi}$.

Introduction rules We take our set of connectives to be $\{\rightarrow, \wedge, \perp\}$. As $\{\rightarrow, \perp\}$ was shown to be an adequate set of connectives in a previous exercise, this superset is also adequate. We need some way to introduce formulae that contain these connectives, so we define some *introduction rules*.

- Suppose we have φ and ψ as premises, then we can conclude $\varphi \wedge \psi$, this is called \wedge *introduction*. We write this as $\frac{\varphi \quad \psi}{\varphi \wedge \psi} (\wedge I)$.
- Suppose we have φ as a premise, and with *finitely* manipulations we derive ψ , then we can remove φ and conclude $\varphi \rightarrow \psi$, this is called \rightarrow *introduction*. We write this as

$$\frac{\begin{array}{c} \varphi \\ \vdots \\ \psi \end{array}}{\varphi \rightarrow \psi} (\rightarrow I)$$

Elimination rules Just as we can deduce formulae with additional connectives from certain premises, we can also eliminate connectives from premises.

- Suppose we have $\varphi \wedge \psi$ as premises, then we can conclude φ , this is called \wedge *elimination*. Note that this is a symmetric operation. We write this as $\frac{\varphi \wedge \psi}{\varphi} (\wedge E)$.
- Suppose we have φ and $\varphi \rightarrow \psi$ as premises, then we can conclude ψ , this is called \rightarrow *elimination* or *modus ponens*. We write this as $\frac{\varphi \quad \varphi \rightarrow \psi}{\psi} (\rightarrow E)$.

Additional rules There are a few additional rules we state

- Given \perp as a premise, we can conclude any φ , this is called the *falsum rule* or the *principle of explosion*. This is written $\frac{\perp}{\varphi} (\perp)$.
- Given $\neg\varphi$ as a premise, if after *finitely* many manipulations we can derive \perp , then we can conclude φ . This is called *reductio ad absurdum*. We write this as

$$\frac{\begin{array}{c} \neg\varphi \\ \vdots \\ \perp \end{array}}{\varphi} (\text{RAA})$$

Remark:

Take care not to confuse RAA with $\rightarrow I$. Note that we interpret $\neg\phi$ as $\phi \rightarrow \perp$. So

$$\frac{\begin{array}{c} \phi \\ \vdots \\ \perp \end{array}}{\neg\phi} (\rightarrow I)$$

is actually an application of \rightarrow introduction. RAA only applies when we start with $\neg\phi$ as a hypothesis.

$$\frac{\begin{array}{c} \neg\phi \\ \vdots \\ \perp \end{array}}{\phi} (RAA)$$

Example 1.17:

Some examples of formal deduction.

: $(\varphi \wedge \psi) \rightarrow (\psi \wedge \varphi)$.

Proof. • Given $\varphi \wedge \psi$ we can conclude, using \wedge elimination, φ and ψ separately.

- By \wedge introduction, we can then conclude $\psi \wedge \varphi$.
- We have shown that $\varphi \wedge \psi$ yields $\psi \wedge \varphi$ after finitely many steps, we can therefore use \rightarrow introduction to remove $(\varphi \wedge \psi)$ and conclude $(\varphi \wedge \psi) \rightarrow (\psi \wedge \varphi)$.

$$\frac{\frac{\frac{\varphi \wedge \psi}{\psi} (\wedge E) \quad \frac{\varphi \wedge \psi}{\varphi} (\wedge E)}{\psi \wedge \varphi} (\wedge I)}{(\varphi \wedge \psi) \rightarrow (\psi \wedge \varphi)} (\rightarrow I)$$

■

: $\varphi \rightarrow \neg\neg\varphi$

Proof. • We note that $\neg\varphi$ is equivalent to $\varphi \rightarrow \perp$ as the latter implies that φ cannot be true.

- We suppose φ and $\neg\varphi$ as premises, from which we conclude \perp by modus ponens.
- Using \rightarrow introduction, we can remove $\neg\varphi$ and conclude $\neg\varphi \rightarrow \perp$ given φ .
- Note that $\neg\varphi \rightarrow \perp$ is equivalent to $\neg\neg\varphi$.
- So using a second \rightarrow introduction, we remove φ and conclude $\varphi \rightarrow \neg\neg\varphi$.

$$\frac{\frac{\frac{\varphi^{(2)} \quad \neg\varphi^{(1)} : (\varphi \rightarrow \perp)}{\perp} (MP)}{(\neg\varphi) \rightarrow \perp : (\neg\neg\varphi)} (\rightarrow I, 1)}{\varphi \rightarrow \neg\neg\varphi} (\rightarrow I, 2)$$

■

Definition 1.18:

Let $\Sigma \subseteq \mathcal{W}(P)$, $\varphi \in \mathcal{W}(P)$, then $\Sigma \vdash \varphi$ if there is $\varphi_1, \dots, \varphi_n = \varphi$ such that either $\varphi_i \in \Sigma$ or φ_i is obtained from a deduction rule given above. i.e. φ_i or all the (uncancelled) assumptions are in Σ .

1.4.1 Formalising deduction

This subsection gives a slightly more formal treatment of the deduction rules we discussed previously.

\wedge introduction is as follows, suppose we have the *finite* deductions $\frac{D}{\varphi}$ and $\frac{D'}{\varphi'}$, we can then write

$$\frac{\frac{D}{\varphi} \quad \frac{D'}{\varphi'}}{\varphi \wedge \varphi'}$$

\wedge elimination is handled similarly, suppose we have the finite deduction $\frac{D}{\varphi \wedge \psi}$, we can then write both

$$\frac{\frac{D}{\varphi \wedge \psi}}{\varphi}$$

and

$$\frac{\frac{D}{\varphi \wedge \psi}}{\psi}$$

Recall that \rightarrow introduction cancels one of the premises. The intuition behind this is that a formal implication doesn't rely on the truth value of the left-hand premise. This is shown below

$$\frac{\frac{\cancel{\varphi} \quad D}{\psi}}{\varphi \rightarrow \psi}$$

for example, in the sentence *if it is snowing, the school is closed*, we don't assume that it is snowing. A deductive reasoning might start with *assume the school is closed*, but to state the final result, this can be discarded.

A diagrammatic example of modus ponens is shown below

$$\frac{\frac{D}{\phi} \quad \frac{D'}{\phi \rightarrow \psi}}{\psi}$$

The principle of explosion is simply stated as

$$\frac{D}{\perp}$$

and *reducio ad absurdum* becomes

$$\frac{\frac{\cancel{\neg\phi} \quad D}{\perp}}{\phi}$$

1.5 Soundness and Completeness theorems

Theorem 1.19: Soundness

Let $\Gamma \subseteq \mathcal{W}(P)$, $\varphi \in \mathcal{W}(P)$, $\Gamma \vdash \varphi \implies \Gamma \models \varphi$.

Theorem 1.20: Completeness

Let $\Gamma \subseteq \mathcal{W}(P)$, $\phi \in \mathcal{W}(P)$, $\Gamma \models \phi \iff \Gamma \vdash \phi$.

Proof, 1.19. Let P be a non-empty set of atoms and let $\Gamma \subseteq \mathcal{W}(P)$.

We would like to show $\Gamma \vdash \phi \implies \Gamma \models \phi$, so if there is a sequence $\phi_1, \dots, \phi_n = \phi$ such that for each ϕ_i either $\phi_i \in \Gamma$ or ϕ_i is obtained from the $\phi_1, \dots, \phi_{i-1}$ by a deductive argument; then for every valuation map that yields $v(\theta \in \Gamma) = 1$, $v(\phi) = 1$.

We do this by induction on n where n is the length of a proof. If $n = 1$, then there is no deduction so $\phi \in \Gamma \implies \Gamma \models \phi$. Thus the base case holds.

Suppose that for some $k \geq 1$, we have $\forall i < k, \Gamma_i \models \phi_i$.

\wedge introduction Let:

- $\Gamma_i \models \phi_i, \Gamma_j \models \phi_j$ s.t. $i + j < k$ and
- $\Gamma_k \supseteq \Gamma_i, \Gamma_j$ (i.e. we have $\Gamma_k \models D_k$ such that $D_k \vdash \phi_i \wedge \phi_j$).

We would like to show $\Gamma_k \models \phi_i \wedge \phi_j$.

This is more or less immediate:

- Suppose we have a valuation map $v : \mathcal{W}(P) \rightarrow \{0, 1\}$ s.t. $v(\phi \in \Gamma_k) = 1$.
- Since $\Gamma_i, \Gamma_j \subseteq \Gamma_k$ and $\Gamma_i \models \phi_i, \Gamma_j \models \phi_j$, we also have $v(\phi_i) = 1 = v(\phi_j)$.
- Thus $v(\phi_i \wedge \phi_j) = \min\{v(\phi_i), v(\phi_j)\} = 1$.

\wedge elimination Let:

- $\Gamma_i \models \phi \wedge \psi$ s.t. $i < k$ and
- $\Gamma_k \supseteq \Gamma_i$ (i.e. we have $\Gamma_k \models D_k, D_k \vdash \phi$).

We would like to show that $\Gamma_k \models \phi$.

This is again clear. Suppose we have a valuation map such that $v(\phi \wedge \psi) = \min\{v(\phi), v(\psi)\} = 1$, then we must have $v(\phi) = 1$.

\rightarrow introduction Suppose:

- $D_{2,k-1}$ is a sequence of deductions $\phi_2, \dots, \phi_{k-1}$ such that the proof

$$\frac{\begin{array}{c} \phi_1 \\ D_{2,k-1} \\ \phi_k \end{array}}{\phi_1 \rightarrow \phi_k}$$

holds,

- $\Gamma_k \models \{\phi_1\} \cup D_{2,k-1} \cup \{\phi_k\}$.

We would like to show for $\Gamma_{2,k} = \Gamma_k \setminus \{\phi_1\}$ (for which $\Gamma_{2,k} \models D_{2,k-1} \cup \{\phi_k\}$), that $\Gamma_{2,k} \models (\phi_1 \rightarrow \phi_k)$.

- Suppose for a contradiction that there exists a valuation map v s.t. $v(\phi_i \in \Gamma_{2,k}) = 1$ and $v(\phi_1 \rightarrow \phi_k) = 0$.
- The latter is true if and only if $v(\phi_1) = 1$ and $v(\phi_k) = 0$.
- This gives a valuation function on $\Gamma_k = \Gamma_{2,k} \cup \{\phi_1\}$ where $v(\theta \in \Gamma_k) = 1$ and $v(\phi_k) = 0$
- This implies $\Gamma_k \not\models \phi_k$ which contradicts $\Gamma_k \models \phi_k$.

Modus ponens Let:

- $D_{i-1} = \{\phi_1, \dots, \phi_{i-1}\}$ and $D_{i+1,k-1} = \{\phi_{i+1}, \dots, \phi_{k-1}\}$ be deductions such that the proof

$$\frac{\begin{array}{c} D_{i-1} \quad D_{i+1,k-1} \\ \phi_i \quad \phi_i \rightarrow \phi_{k-1} \end{array}}{\phi_{k-1}}$$

holds,

- $\Gamma_i \models D_{i-1} \cup \{\phi_i\}$ and
- $\Gamma_{i+1,k} \models D_{i+1,k-1} \cup \{(\phi_i \rightarrow \phi_{k-1})\}$.

We would like to show that if $\Gamma_k \supseteq \Gamma_i, \Gamma_{i+1,k}$ then $\Gamma_k \models \phi_{k-1}$ (i.e. we have $\Gamma_k \models D_k$, such that $D_k \vdash \phi_{k-1}$, $D_k = D_{i-1} \cup \{\phi_i\} \cup D_{i+1,k-1} \cup \{(\phi_i \rightarrow \phi_{k-1})\}$).

This is again clear:

- Suppose we have a valuation function v s.t. $v(\theta \in \Gamma_k) = 1$.
- In particular we have $v(\phi_i) = 1 = v((\phi_i \rightarrow \phi_{k-1}))$.
- Noting that $\phi_{k-1} \leftrightarrow (\phi_i \wedge (\phi_i \rightarrow \phi_{k-1}))$ is a tautology, we have $v(\phi_{k-1}) = 1$.
- Thus $\Gamma_k \models \phi_{k-1}$

Falsum rule Let:

- D be a set of deductions such that the proof

$$\frac{D}{\perp}$$

holds and

- Γ be such that $\Gamma \models D$ (note that $\Gamma \models \perp$ if and only if there is no valuation function such that $v(\theta) = 1$ for all $\theta \in \Gamma$).

We wish to show that for any $\phi \in \mathcal{W}(P)$, $\Gamma \models \phi$.

- Suppose for a contradiction that $\Gamma \not\models \phi$, i.e. we have a valuation function such that for every $\theta \in \Gamma$, $v(\theta) = 1$ and $v(\phi) = 0$.
- This contradicts our earlier assertion that there is no valuation function such that $v(\theta) = 1$ for all $\theta \in \Gamma$.
- And so $\Gamma \models \phi$.

Reductio ad absurdum Let:

- D be a set of deductions such that the proof

$$\frac{\neg\phi \quad D}{\perp}$$

holds,

- Γ be such that $\Gamma \models D \cup \{\neg\phi\}$ and
- $\Gamma' = \Gamma \setminus \{\neg\phi\}$.

We would like to show that if $\Gamma \models \perp$ then $\Gamma' \models \phi$.

- Suppose not, i.e. we have a valuation map v such that $v(\theta)$ for all $\theta \in \Gamma'$ and $v(\phi) = 0$.
- Then $v(\neg\phi) = 1$.
- Then $v(\theta) = 1$ for all $\theta \in \Gamma$ as $\Gamma = \Gamma' \cup \{\neg\phi\}$.
- This contradicts $\Gamma \models \perp$

- thus $\Gamma \models \phi$.

■

Remark:

An immediate corollary of the soundness theorem is that if a statement is not a semantic consequence of a set of assumptions, it cannot be deduced from those assumptions. So

$$\Gamma \not\models \phi \implies \Gamma \not\vdash \phi$$

1.6 Consistent and maximally consistent sets of formulae

The completeness theorem states that the converse of the soundness theorem is also true. That if a statement is the semantic consequence of a set of assumptions, it can be deduced from those assumptions.

To prove this, we will show that if a statement cannot be deduced from a set of assumptions, then it cannot be a semantic consequence of those assumptions. So

$$\Gamma \not\vdash \phi \implies \Gamma \not\models \phi$$

Definition 1.21:

We say $\Gamma \subseteq \mathcal{W}(P)$ is *consistent* if $\Gamma \not\vdash \perp$ (conversely, it is *inconsistent* if $\Gamma \vdash \perp$).

Lemma 1.22:

The following are equivalent

- (I) Γ is consistent
- (II) For no $\phi \in \mathcal{W}(P)$, does both $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$ hold.
- (III) $\Gamma \not\vdash \phi$ for at least one $\phi \in \mathcal{W}(P)$.

i.e.

- (I*) Γ is inconsistent
- (II*) $\exists \phi \in \mathcal{W}(P)$ for which both $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$ hold.
- (III*) $\Gamma \vdash \phi$ for all $\phi \in \mathcal{W}(P)$.

Proof. We show $I^* \implies III^* \implies II^* \implies I^*$.

$I^* \implies III^*$ If $\Gamma \vdash \perp$, then by the principle of explosion $\perp \vdash \phi, \forall \phi$. So $\Gamma \vdash \phi, \forall \phi$.

$III^* \implies II^*$ Let $\psi = \neg\phi$, then $\Gamma \vdash \psi$ and $\Gamma \vdash \phi$ so $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$.

$II^* \implies I^*$ Since $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$, $\Gamma \vdash \phi \wedge (\neg\phi) \implies \Gamma \vdash \perp$, thus Γ is inconsistent. ■

Lemma 1.23:

Suppose v is a valuation function and $\Gamma \subseteq \mathcal{W}(P)$ s.t. $v(\psi) = 1, \forall \psi \in \Gamma$. Then Γ is consistent.

Proof. Suppose not, then by soundness we have $\Gamma \vdash \perp \implies \Gamma \models \perp$ which is a contradiction. ■

Lemma 1.25:

Suppose that Γ is a consistent set of formulae.

- (a) If $\Gamma \cup \{\neg\varphi\} \vdash \perp$ then $\Gamma \vdash \varphi$.
- (b) If $\Gamma \cup \{\varphi\} \vdash \perp$ then $\Gamma \vdash \neg\varphi$.

Proof. (a) Suppose that D is a set of deductions such that $\Gamma \vdash D$ and

$$\begin{array}{c} \neg\varphi \\ D \\ \perp \end{array}$$

- By RAA we have

$$\frac{\begin{array}{c} \neg\varphi \\ D \\ \perp \end{array}}{\varphi} \text{ (RAA)}$$

- and so $\Gamma \vdash \varphi$.

(b) Suppose that D is a set of deductions such that $\Gamma \vdash D$ and that

$$\begin{array}{c} \varphi \\ D \\ \perp \end{array}$$

- Then by $\rightarrow I$ we have

$$\frac{\begin{array}{c} \varphi \\ D \\ \perp \end{array}}{\varphi \rightarrow \perp: \neg\varphi} (\rightarrow I)$$

- and so $\Gamma \vdash \neg\varphi$.

■

Definition 1.26:

Γ is called *maximally consistent* when

- Γ is consistent and
- if we have $\Gamma' \supseteq \Gamma$ such that Γ' is consistent, then $\Gamma' = \Gamma$.

So if $\Gamma' \supsetneq \Gamma$, $\Gamma' \vdash \perp$ is inconsistent.

Example 1.27:

Let v be a valuation function, then $\Gamma = \{\phi : v(\phi) = 1\}$ is maximally consistent.

Proof. Suppose that $\Gamma' \supsetneq \Gamma$.

- For $\psi \in \Gamma' \setminus \Gamma$, we have that $v(\psi) = 0$.
- Then $v(\neg\psi) = 1 \implies \neg\psi \in \Gamma \subsetneq \Gamma'$.
- $\psi, \neg\psi \in \Gamma' \implies \Gamma' \vdash \perp$ is inconsistent.

■

Lemma 1.28:

If Γ is maximally consistent, then Γ is closed under deductions. i.e. if $\Gamma \vdash \phi$, then $\phi \in \Gamma$.

Proof. Suppose for a contradiction that we have $\Gamma \vdash \phi$ but $\phi \notin \Gamma$.

- Since Γ is maximally consistent, $\Gamma \cup \{\phi\} \vdash \perp$ is inconsistent.
- By 1.25 we have that $\Gamma \vdash \neg\phi$.
- So $\Gamma \vdash \{\neg\phi, \phi\} \Rightarrow \Gamma \vdash \perp$ is inconsistent, which is a contradiction.

■

Lemma 1.29:

Let Γ be maximally consistent. Then:

(I) For all formulae ϕ , either $\phi \in \Gamma$ or $\neg\phi \in \Gamma$.

(II) For all formulae ϕ, ψ

$$(\phi \rightarrow \psi) \in \Gamma \iff (\phi \in \Gamma \implies \psi \in \Gamma)$$

Proof. (I) Suppose $\phi \notin \Gamma$.

- Then $\Gamma \cup \{\phi\}$ is inconsistent.
- By 1.25 we have that $\Gamma \vdash \neg\phi$.
- Then by 1.28 we have $\neg\phi \in \Gamma$.

$\neg\phi \notin \Gamma \implies \phi \in \Gamma$ is proven similarly.

(II) • (\implies) Suppose that $\{(\phi \rightarrow \psi), \phi\} \in \Gamma$.

- Then clearly $\Gamma \vdash \{(\phi \rightarrow \psi), \phi\}$.
- By MP and 1.28 respectively, we have $\Gamma \vdash \psi \implies \psi \in \Gamma$.

• (\impliedby) Suppose that $(\phi \in \Gamma \implies \psi \in \Gamma)$.

- Suppose for a contradiction that $\neg(\phi \rightarrow \psi) \in \Gamma$.

Exercise:

Show that $\vdash (\neg(\phi \rightarrow \psi) \leftrightarrow (\phi \wedge \neg\psi))$.

- Thus we have that $(\phi \wedge \neg\psi) \in \Gamma \implies \phi, \neg\psi \in \Gamma$.
- Which implies $\psi, \neg\psi \in \Gamma$ and thus Γ is inconsistent, which is a contradiction.
- Thus $(\phi \rightarrow \psi) \in \Gamma$.

Corollary 1.30:

If Γ is maximally consistent, then

$$\phi \in \Gamma \iff \neg\phi \notin \Gamma$$

$$\neg\phi \in \Gamma \iff \phi \notin \Gamma$$

Lemma 1.31: Lindenbaum lemma for propositional logic

For each $\Gamma \subseteq \mathcal{W}(P)$ that is consistent, there is a maximally consistent $\Gamma^* : \Gamma \subseteq \Gamma^* \subset \mathcal{W}(P)$

Proof. We can use Zorn's lemma. Suppose that $|P|$ is at most countable.

Exercise:

Show that $\mathcal{W}(P)$ is countable.

- We can then enumerate $\mathcal{W}(P)$ by $\mathcal{W}(P) = \{\phi_i : i \in \mathbb{Z}^+\}$.
- Define a non-decreasing sequence, $\Gamma_0 \subseteq \Gamma_1 \subseteq \Gamma_2 \subseteq \dots$ by setting $\Gamma_0 = \Gamma$ and

$$\Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{\phi_n\} & : \Gamma_n \cup \{\phi_n\} \text{ is consistent} \\ \Gamma_n & : \text{else} \end{cases}$$

- Note that for all $n \in \mathbb{Z}_{\geq 0}$, Γ_n is consistent.

- Let $\Gamma^* = \bigcup_{n \in \mathbb{Z}_{\geq 0}} \Gamma_n$, we claim that Γ^* is maximally consistent.
 - Suppose for a contradiction that $\Gamma^* \vdash \perp$ is inconsistent.
 - * Then there is a finite subset $\Phi = \{\phi_{i_1}, \dots, \phi_{i_m}\} \subset \Gamma^*$ such that $\Phi \vdash \perp$.
 - * Let $t = \max\{i_j : j \leq m\}$.
 - * Then $\Phi \subseteq \Gamma_t \vdash \perp$ is inconsistent, contradiction.
 - * Thus Γ^* is consistent.
 - Γ^* is clearly maximal since for $\psi = \phi_m \in \mathcal{W}(P)$, $\Gamma_m \cup \{\phi_m\}$ is either consistent or not.
 - * If it is, then $\psi = \phi_m \in \Gamma_{m+1} \subseteq \Gamma^*$.
 - * If it is not, then $\psi = \phi_m \notin \Gamma^*$ and $\{\phi_m\} \cup \Gamma_m \subseteq \{\phi_m\} \cup \Gamma^* \vdash \perp$ is inconsistent.

■

Lemma 1.32:

Suppose that Γ is consistent. Then there is a valuation function v such that $v(\theta) = 1, \forall \theta \in \Gamma$.

Proof. By 1.31 there exists a maximally consistent $\Gamma^* \supseteq \Gamma$.

- Define a valuation function v such that

$$v(p_i) = \begin{cases} 1 & : p_i \in \Gamma^* \\ 0 & : \text{else} \end{cases}$$

- We claim that $v(\phi) = 1 \iff \phi \in \Gamma^*$.

Exercise:

Prove this using induction on formulae.

■

Corollary 1.33:

If $\Gamma \not\vdash \phi$ then there is a valuation function v such that $v(\theta) = 1, \forall \theta \in \Gamma$ but $v(\phi) = 0$ (i.e. $\Gamma \not\models \phi$).

Proof. Note that by the contrapositive of 1.25 (a), we have that $\Gamma \not\vdash \perp \implies \Gamma' = \Gamma \cup \{\neg\phi\} \not\vdash \perp$ is consistent.

- Now by 1.32, we have a valuation function v such that $v(\theta) = 1, \forall \theta \in \Gamma'$.
- Thus $v(\neg\phi) = 1 \implies v(\phi) = 0$

■

Proof, 1.20, Completeness theorem. We wish to show that for $\Gamma \subseteq \mathcal{W}(P)$, $\phi \in \mathcal{W}(P)$

$$\Gamma \vdash \phi \iff \Gamma \models \phi$$

- Theorem 1.19, soundness, states that $\Gamma \vdash \phi$ implies $\Gamma \models \phi$.
- The above corollary states that $\Gamma \not\vdash \phi$ implies $\Gamma \not\models \phi$.
- Thus $\Gamma \vdash \phi \iff \Gamma \models \phi$.

■

2 Predicate logic

Predicate logic, also known as *first order logic*, allows use to express statements that use quantifiers such as \forall and \exists . As with propositional logic, we will take two approaches:

- *The semantic approach* will cover 1st order structures and formulae.
- *The syntactical approach* will give a formal system for 1st order logic and allow deductions to be performed.

The analogue of the completeness theorem for propositional logic will be *Gödel's completeness theorem*.

2.1 First order structures, languages and formulas

Definition 2.1:

Suppose A is a set and $n \in \mathbb{Z}^+$.

- An n -ary relation on A is a subset of the form $R \subseteq A^n$

$$A^n = \underbrace{A \times A \times \cdots \times A}_{n \text{ times}} = \{(a_1, \dots, a_n) : a_i \in A\}$$

- An n -ary function of A is a well-defined map $f : A^n \rightarrow A$.

Example 2.2:

Some relations and functions

- \leq is a 2-ary (binary) relation on \mathbb{R} .
- $+$ is a binary function on \mathbb{C} .
- $x \mapsto e^x$ is a 1-ary (unary) function on \mathbb{C} .
- $P \subset \mathbb{Z}$, the subset of prime numbers, is a unary relation on \mathbb{Z} .

Definition 2.3:

A first order structure \mathcal{A} consists of

- (I) A non-empty set A which is called the *domain* of \mathcal{A} .
- (II) A set of relations on A , $\{R_i \subseteq A^{n_i} : i \in I, n_i \in \mathbb{Z}^+\}$.
- (III) A set of functions on A , $\{f_j : A^{n_j} \rightarrow A \mid j \in J, n_j \in \mathbb{Z}^+\}$.
- (IV) A set of constants, $\{c_k \in A : k \in K\}$.

We can then write

$$\mathcal{A} = \langle \underbrace{A}_{\text{domain}}, \underbrace{(R_i : i \in I)}_{\text{relations}}, \underbrace{(f_j : j \in J)}_{\text{functions}}, \underbrace{(c_k : k \in K)}_{\text{constants}} \rangle$$

We call the information $(n_i, i \in I), (n_j, j \in J), K$ the *signature* of \mathcal{A} .

Example 2.4:

Some first order structures

$| : (\mathbb{Z}_{\geq 0}, \leq)$ is a structure with signature, $I = \{1\}, J = K = \emptyset, n_{\leq} = n_1 = 2$.

- : A group is a structure with:
 - equality as a binary relation,
 - a binary function (the group operation),
 - a unary function (element inversion) and
 - a constant (the identity element).
- : A field or ring is a structure with the same the same relations, functions and constants as a group plus:
 - a second binary function (multiplication)
 - a second constant (the multiplicative identity)

Definition 2.5:

A first order language L consists of:

- (I) A countable set of variables $\mathcal{V} = \{v_1, v_2, \dots\}$.
- (II) A set of connectives

$$\{\wedge, \vee, \neg, \rightarrow, \leftrightarrow, \perp, \underbrace{\forall}_{\text{for all}}, \underbrace{\exists}_{\text{there exists}}\}$$

- (III) Auxiliary symbols $\{(\cdot), \langle \text{comma} \rangle\}$.
- (IV) A set of relations $\mathcal{R} = \{R_i : i \in I\}$.
- (V) A set of functions, $\mathcal{F} = \{f_j : j \in J\}$.
- (VI) A set of constants, $\mathcal{C} = \{C_k : k \in K\}$.
- (VII) An equality symbol $=$.

Remark: A review of propositional logic

Propositional logic had a set of elements $E = P \cup \{\wedge, \vee, \rightarrow, \leftrightarrow, \neg, \perp\} \cup \{(\cdot), \langle \text{comma} \rangle\}$. We can consider the set $S_f(E) = \{ \text{all finite sequences of elements of } E \}$.

If we view E^n as the set of all n -sequences of elements of E , then $E^* = \bigcup_{n \in \mathbb{Z}_{\geq 0}} E^n = S_f(E)$.

Note that $\mathcal{W}(P) \subsetneq S_f(E)$.

Accordingly, we can look at $S_f(\mathcal{L})$ where \mathcal{L} is a first order structure.

Definition 2.6:

The set of L -terms $\text{Term}(L)$ is the smallest set $\text{Term}(L) \subseteq S_f(L)$ such that:

- (I) The set of variables $\mathcal{V} \subseteq \text{Term}(L)$, i.e. for all $i \in I, v_i \in \text{Term}(L)$.
- (II) If $f \in \mathcal{F}$ is an n_f -ary function and $t_1, \dots, t_{n_f} \in \text{Term}(L)$, then $f(t_1, \dots, t_{n_f}) \in \text{Term}(L)$. Often we will drop the brackets and write $f t_1 \dots t_{n_f}$.
- (III) The set of constants $\mathcal{C} \subseteq \text{Term}(L)$, i.e. for all $k \in K, c_k \in \text{Term}(L)$.

$L_{\text{ord}} = \{\leq\} \cup \{\wedge, \vee, \neg, \rightarrow, \leftrightarrow, \perp, \forall, \exists\} \cup \{(\cdot), \langle \text{comma} \rangle\}$, where \leq is a binary relation. Then we only have variables as terms, $V = \text{Term}(L_{\text{ord}})$.

$L_1 = \{+, f\} \cup \dots$ where $+$ is a binary function and f is a unary function. Then

- $V \subseteq \text{Term}(L_1)$ as before,
- $f t_i \in \text{Term}(L_1)$ for all $t_i \in \text{Term}(L)$ and

- $t_i + t_j \in \text{Term}(L)$ for all $t_i, t_j \in \text{Term}(L)$.

Definition 2.7:

The set of L -formulas, $\text{Form}(L)$, is the smallest subset of $S_f(L)$ such that:

- (I) $\perp \in \text{Form}(L)$.
- (II) $t_1, \dots, t_{n_R} \in \text{Term}(L) \implies R(t_1, \dots, t_{n_R}) \in \text{Form}(L)$.
- (III) $t_1, t_2 \in \text{Term}(L) \implies t_1 \dot{=} t_2 \in \text{Form}(L)$.
- (IV) $\phi, \psi \in \text{Form}(L) \implies \phi \Box \psi \in \text{Form}(L)$, for any connective \Box .
- (V) $\phi \in \text{Form}(L) \implies \neg \phi \in \text{Form}(L)$.
- (VI) $\phi \in \text{Form}(L) \implies \forall v \phi \in \text{Form}(L)$ and $\exists v \phi \in \text{Form}(L)$.

Example 2.8:

Some examples of terms and formulas.

| : Consider $L = \{\leq\} \cup \{c\}$, then $\text{Term}(L) = \mathcal{V} \cup \{c\}$ and $\text{Form}(L) = \{v_i \leq v_j, v_i \leq c\}$.

| : $L = \{+\} \cup \{0\}$, then $0 + v_i \in \text{Term}(L)$ and $0 + v_i \dot{=} v_i + v_2 \in \text{Form}(L)$.

2.2 Interpretations of languages and assignments**Definition 2.9:**

We say $\phi \in \text{Form}(L)$ is *atomic* if ϕ is of the form $R_i(t_1, \dots, t_{n_i})$ where $R_i \in \mathcal{R}$ is a n_i -ary relation and $t_1, \dots, t_{n_i} \in \text{Term}(L)$.

Example 2.10:

Suppose L has:

- a binary function f ,
- a unary relation P ,
- a binary relation R and
- constants c_1, c_2 .

Then

$$v_0, f v_0 v_1, c_1, f c_1 v_0, f f c_1 v_1 c_2$$

are all terms.

$$P(v_0), R(v_0, v_1), R(f c_1 v_0, c_2), f(c_0, v_0) \dot{=} c_2$$

are all atomic formulas.

$$\neg R(v_0, v_1), \forall v R(v, c_0)$$

are not atomic.

Definition 2.11:

Let L be a 1st order language with $R_i \in \mathcal{R}, i \in I$ a set of n_i -arity relations, $f_j \in \mathcal{F}, j \in J$ a set of m_j -arity functions and $c_k \in \mathcal{C}, k \in K$ a set of constants.

An L -structure is a first order structure

$$\mathcal{A} = \langle A, (R_i : i \in I), (f_j : j \in J), (c_k : k \in K) \rangle$$

of the same signature as L .

There is a correspondence between L and R_i, f_j, c_k in \mathcal{A} . This is called an *interpretation* of L .

Remark: Notation

We write

$$L = \mathcal{R} \cup \mathcal{F} \cup \mathcal{C} \cup \underbrace{\dots}_{\text{logical symbols and variables}}$$

to define a language. If \mathcal{A} is an L -structure, then:

- given $R_i \in \mathcal{R}$, $R_i^{\mathcal{A}}$ is the interpretation of R_i in \mathcal{A} .
- given $f_j \in \mathcal{F}$, $f_j^{\mathcal{A}}$ is the interpretation of f_j in \mathcal{A} .
- given $c_k \in \mathcal{C}$, $c_k^{\mathcal{A}}$ is the interpretation of c_k in \mathcal{A} .

Definition 2.12:

Let L be 1st order language and \mathcal{A} an L -str. A function $\beta : V \rightarrow A$ is called an \mathcal{A} -assignment.

We extend β to $\text{Term}(L)$ as follows.

- (I) $t^{\mathcal{A}}[\beta] = \beta(v_i)$ if $t = v_i$.
- (II) $t^{\mathcal{A}}[\beta] = c_k^{\mathcal{A}}$ if $t = c_k$.
- (III) $t^{\mathcal{A}}[\beta] = f_j^{\mathcal{A}}(t_1^{\mathcal{A}}[\beta], \dots, t_{m_j}^{\mathcal{A}}[\beta])$ if $t = f_j t_1 \dots t_{m_j}$.

Example 2.13: $(\mathbb{Z}_{\geq 0}, \text{succ}, 0)$

Consider $L = \mathcal{R} = \emptyset \cup \mathcal{F} = \{\text{succ}\} \cup \mathcal{C} = \{0\} \cup \dots$. We define $\beta : V \rightarrow A, v_i \mapsto i$. We define $\text{succ}^{\mathbb{Z}_{\geq 0}}(n) := n + 1$.

Then

$$\begin{aligned} \text{succ}(0)[\beta] &= \text{succ}^{\mathbb{Z}_{\geq 0}}(0[\beta]) \\ &= \text{succ}^{\mathbb{Z}_{\geq 0}}(0^{\mathbb{Z}_{\geq 0}}) \\ &= 1^{\mathbb{Z}_{\geq 0}} \end{aligned}$$

which implies $v_i[\beta] = i = \beta(v_i)$. We can also show

$$\begin{aligned} \text{succ}(\text{succ}(v_5))[\beta] &= \text{succ}^{\mathbb{Z}_{\geq 0}}(\text{succ}(v_5)[\beta]) \\ &= \text{succ}^{\mathbb{Z}_{\geq 0}}(\text{succ}^{\mathbb{Z}_{\geq 0}}(v_5[\beta])) \\ &= \text{succ}^{\mathbb{Z}_{\geq 0}}(\text{succ}^{\mathbb{Z}_{\geq 0}}(\beta(v_5))) \\ &= 7^{\mathbb{Z}_{\geq 0}} \end{aligned}$$

Remark: Notation

If \mathcal{A} is an L -str. and $\beta : V \rightarrow A$ is an \mathcal{A} -assignment. Then:

- For $a \in A$,

$$\beta(a/v) = \begin{cases} \beta(v_i) : v_i \neq v \\ a : v_i = v \end{cases}$$

- For $v \in V$,

$$\beta(a/v) : V \rightarrow A, \\ v_i \mapsto \begin{cases} \beta(v_i) : v_i \neq v \\ a : v_i = v \end{cases}$$

Definition 2.14:

Suppose that L is a language and \mathcal{A} is an L -str. Consider $\beta : \text{Term}(L) \rightarrow \mathcal{A}$ an assignment function. Then a valuation map on $\text{Form}(L)$ is a function $v^\beta : \text{Form}(L) \rightarrow \{0, 1\}$ defined inductively as

(I) $v^\beta(\perp) = 0$.

(II) For two terms $t_1, t_2 \in \text{Term}(L)$ we have $t_1 \doteq t_2 \in \text{Form}(L)$. Then

$$v^\beta(t_1 \doteq t_2) = \begin{cases} 1 & : t_1^A[\beta] = t_2^A[\beta] \\ 0 & : \text{else} \end{cases}$$

(III) For terms $t_1, \dots, t_{n_i} \in \text{Term}(L)$ we have $R_i t_1 \dots t_{n_i} \in \text{Form}(L)$. Then

$$v^\beta(R_i t_1 \dots t_{n_i}) = \begin{cases} 1 & : (t_1^A[\beta], \dots, t_{n_i}^A[\beta]) \in R_i^A \\ 0 & : \text{else} \end{cases}$$

(IV) For $\phi, \psi \in \text{Form}(L)$, $v^\beta(\phi \Box \psi)$ is defined as in the cases for propositional logic.

(V) For $\phi \in \text{Form}(L)$, $v^\beta(\neg \phi) = 1 - v^\beta(\phi)$.

(VI) For $\phi \in \text{Form}(L)$, $v^\beta(\forall w \phi) = \min\{v^{\beta(a/w)}(\phi) : a \in A\}$

(VII) For $\phi \in \text{Form}(L)$, $v^\beta(\exists w \phi) = \max\{v^{\beta(a/w)}(\phi) : a \in A\}$

- If $v^\beta(\phi) = 1$ we write $\mathcal{A} \models \phi[\beta]$.
- If for all \mathcal{A} -assignments β we have $\mathcal{A} \models \phi[\beta]$, then we write $\mathcal{A} \models \phi$ and say ϕ is *satisfied* by \mathcal{A} or \mathcal{A} is a *model* of ϕ .
- If for all L -structures \mathcal{A} we have $\mathcal{A} \models \phi$, then we write $\models \phi$ and say that ϕ is *logically valid*.

2.3 Variable binding and subformulas

Lemma 2.15:

Suppose that \mathcal{A} is an L -structure and that $\phi, \psi \in \text{Form}(L)$ are sentences. Then

(I) $\mathcal{A} \models \phi \wedge \psi \iff \mathcal{A} \models \phi \text{ and } \mathcal{A} \models \psi$.

(II) $\mathcal{A} \models \phi \vee \psi \iff \mathcal{A} \models \phi \text{ or } \mathcal{A} \models \psi$.

(III) $\mathcal{A} \models \neg \phi \iff \mathcal{A} \not\models \phi$.

(IV) $\mathcal{A} \models (\phi \rightarrow \psi) \iff (\mathcal{A} \models \phi \implies \mathcal{A} \models \psi)$.

Now suppose that $\phi \in \text{Form}(L)$ is a formula where the only free variable is $v \in V$.

(V) $\mathcal{A} \models (\forall v \phi) \iff \text{for all } a \in A, \mathcal{A} \models \phi[a]$.

(VI) $\mathcal{A} \models (\exists v \phi) \iff \text{there exists } a \in A, \mathcal{A} \models \phi[a]$.

Proof. The proofs of (I), (II) and (III) are obvious.

(IV) \implies : Suppose that $\mathcal{A} \models (\phi \rightarrow \psi), \phi$. Let $\beta : V \rightarrow A$ be any assignment. We want to show $\mathcal{A} \models \psi[\beta] \iff v^\beta(\phi) = 1$.

$$v^\beta(\phi \rightarrow \psi) = \max\{1 - v^\beta(\phi), v^\beta(\psi)\}$$

$$1 = \max\{1 - 1, v^\beta(\phi)\}$$

$$v^\beta(\phi) = 1$$

Exercise:

Prove cases (V) and (VI).

Remark: Notation

For $t \in \text{Term}(L)$ we write $t(v_1, \dots, v_n)$ where:

- the v_i s are distinct variables.
- all the variables appearing in t are $\{v_1, \dots, v_n\}$.

Lemma 2.16:

Suppose $\gamma, \beta : V \rightarrow A$ are two \mathcal{A} -assignments which agree on variables v_1, \dots, v_n , i.e. for all $i : 1 \leq i \leq n$, $\gamma(v_i) = \beta(v_i)$. Then for $t(v_1, \dots, v_n) \in \text{Term}(L)$, $t^{\mathcal{A}}[\gamma] = t^{\mathcal{A}}[\beta]$.

Exercise:

Prove the above.

Example:

Consider $L = \{R_i\}$ a language with a single binary relation. The terms are therefore $\text{Term}(L) = V$ and $R_{v_0 v_1}$ is an atomic formula. If $\Psi = \forall v_0 R_{v_0 v_1}$ is another formula, we note that ψ is a sub-formula of

$$\phi = \forall v((\underbrace{\forall v_0 R_{v_0 v}}_{\psi}) \wedge (\exists v_0 R_{v_0 v_1}))$$

Definition 2.17:

Suppose that ϕ, ψ are L -formulas and $\forall x_i \phi$ occurs as a subformula of ψ , i.e. $\psi = \dots (\forall x_i \phi) \dots$. Then we say that ϕ is the *scope* of the quantifier $\forall x_i$ in ψ .

An occurrence of a variable x_i in ψ is *bounded* if it is in the scope of a quantifier $(\forall x_i)$ in ψ (for some sub-formula ϕ).

If a variable is not in the scope of any quantifier, it has a *free occurrence*. Variables with a free occurrence in ψ are called *free variables* of ψ .

Example:

Suppose $L = \{R_1, R_2\}$ where R_i are binary relations.

$$\forall x_1 \underbrace{((R_1 x_1 x_2) \rightarrow (R_2 x_1 x_2))}_{\text{scope of } \forall x_1}$$

In the equation above:

- x_1 is bound.
- x_2 has free occurrence.

$$((\forall x_1 \underbrace{(R_1 x_1 x_2)}_{\text{scope of } \forall x_1}) \rightarrow (R_2 x_1 x_2))$$

In the equation above:

- The first x_1 is bound, the second one is free.
- As before, x_2 has free occurrence.

$$((\exists x_1 \underbrace{(R_1 x_1 x_2)}_{\text{scope of } \exists x_1}) \rightarrow (\forall x_2 \underbrace{(R_2 x_2 x_3)}_{\text{scope of } \forall x_2}))$$

- x_1 is bound.
- The first x_2 has free occurrence, the second one is bound.
- x_3 has free occurrence.

Definition 2.18:

A formula $\phi \in \text{Form}(L)$ is called a *sentence* if it has no free variables.

Example:

For $L = \{R\}$, $(\forall x, Rxx) \in \text{Form}(L)$ is a sentence.

Remark: Notation

For $\phi \in \text{Form}(L)$, write $\phi(v_1, \dots, v_n)$ when

- the v_i s are distinct variables.
- the v_i s are the only free variables appearing in ϕ .

Lemma 2.19:

Suppose β, γ are two \mathcal{A} -assignments which agree in v_1, \dots, v_n and let $\phi \in \text{Form}(L)$, $\phi(v_1, \dots, v_n)$. Then

$$v^\beta(\phi) = v^\gamma(\phi)$$

Exercise:

Prove the above using induction.

Remark: Notation

Let $\beta : V \rightarrow A$ where A is the domain of an L -str. \mathcal{A} . Let $\phi(v_1, \dots, v_n) \in \text{Form}(L)$. We write

$$\mathcal{A} \models \phi[a_1, \dots, a_n]$$

when $\mathcal{A} \models \phi[\beta]$ and $\beta(v_i) = a_i$ for all $1 \leq i \leq n$.

Recall ϕ is logically valid if for all L -structures \mathcal{A} , we have $\mathcal{A} \models \phi$.

Example:

Consider the formula $v \doteq v$ and let $\beta : V \rightarrow A, v \mapsto a$ be an assignment for some L -str. \mathcal{A} . Then as $v^{\mathcal{A}}[\beta] = v^{\mathcal{A}}[\beta] \implies v^{\beta}(v \doteq v) = 1$.

Exercise:

Show that the following formula is logically valid.

$$(\exists v_1 (\forall v_2 \phi)) \rightarrow (\forall v_2 (\exists v_1 \phi))$$

i.e. an existential quantifier can be moved behind a universal quantifier in its scope.

Corollary: (Lemma 2.19)

Suppose ϕ is a sentence, then either $\mathcal{A} \models \phi$ or $\mathcal{A} \models \neg\phi$.

For an L -structure \mathcal{A} , we write $\text{Con}(\mathcal{A}) = \{\phi \in \text{Form}(L) : \phi \text{ is a sentence, } \mathcal{A} \models \phi\}$.

2.4 Substitution

Given a non-empty set P , suppose $\theta \in \mathcal{W}(P)$ is such that p_1, \dots, p_n appear in θ . Let $\phi_1, \dots, \phi_n \in \text{Form}(L)$ and let ψ be the formula arrived at by substituting p_i in θ with ϕ_i . We claim that $\psi \in \text{Form}(L)$.

Exercise:

Prove this by inducting on $\theta \in \mathcal{W}(P)$.

Given $\beta : V \rightarrow A$ an \mathcal{A} -assignment, we can consider $(v^{\beta}(\phi_1), \dots, v^{\beta}(\phi_n))$ as $\bar{x} = (0, 1, \dots) \in \{0, 1\}^n$. Thus $v^{\beta}(\psi) = F_{\theta}(\bar{x})$.

Corollary:

If $\theta = \theta(p_1, \dots, p_n)$ is a propositional tautology and $\phi_1, \dots, \phi_n \in \text{Form}(L)$. Then $\psi \in \text{Form}(L)$ defined by replacing each p_i in θ with ϕ_i is a logically valid formula.

Similar to propositional logic, we can define a substitution.

Definition 2.20:

Suppose $t, s \in \text{Term}(L), v \in V$. Then we define $s[t/v]$ as follows:

- If $s = v_i$, then $s[t/v] = \begin{cases} t & v_i = v \\ v_i, & v_i \neq v \end{cases}$.
- If $s = c$, then $s[t/v] = c[t/v] = c$.
- If $s = f t_1 \cdots t_{m_j}$, then $s[t/v] = f t_1[t/v] \cdots t_{m_j}[t/v]$.

We can therefore see that substitution is an *operator* on $\text{Term}(L)$.

This can now be extended to formulas.

Definition 2.21:

Let $\phi \in \text{Form}(L)$, then we define $\phi[t/v]$ as follows:

- $\perp[t/v] = \perp$.
- $Rt_1 \cdots t_{n_i}[t/v] = Rt_1[t/v] \cdots t_{n_i}[t/v]$.
- $(t_1 \doteq t_2)[t/v] = t_1[t/v] \doteq t_2[t/v]$.
- $(\phi \Box \psi)[t/v] = \phi[t/v] \Box \psi[t/v]$ (as in propositional logic).
- $(\forall v_i \phi(v_i))[t/v] = \begin{cases} \forall v_i (\phi[t/v]) : v \neq v_i \\ \forall v_i \phi(v_i) : v = v_i \end{cases}$.
- $(\exists v_i \phi)[t/v]$ is defined analogously.

Example:

We must be wary of situations like below, consider

$$\psi(y) = \exists x(x < y)$$

and the substitution $\psi[x/y] = \exists x(x < x)$. We do not want this to happen so there must be further guards put in place against similarly illegal substitutions. For now, assume that substitutions can only be performed on sentences. A result on the valuation of any substitution is given by [Lemma 2.43](#)

2.5 Some examples of structures

Define $0 = \emptyset$, $n = \{0, 1, \dots, n-1\}$ and \mathbb{N} = the set of all natural numbers. Consider the structures

$$\mathcal{A}_1 = \langle m, \in, 0, +_m, s_1 \rangle$$

$$\mathcal{A}_2 = \langle \mathbb{N}, <, s_1, +, 0 \rangle$$

$$\mathcal{A}_3 = \langle \mathbb{N}, <, s_2, \cdot, 1 \rangle$$

where

- For \mathcal{A}_1
 - For $i, j \in m = \{0, 1, \dots, m-1\}$, $j \in i$ if $j \ni i$ as sets.
 - $s_1 : m \rightarrow m, a \mapsto \begin{cases} a+1 = \{0, 1, \dots, a\} : a+1 \leq m \\ 0 : \text{else} \end{cases}$.
 - $+_m$ is defined through repeated addition.
- For \mathcal{A}_2
 - $n < m$ is the usual order on \mathbb{N} .
 - $s_1 : \mathbb{N} \rightarrow \mathbb{N}, a \mapsto a+1$.
 - $+: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}, (a, b) \mapsto a+b$.
- For \mathcal{A}_3
 - $n < m$ is the usual order on \mathbb{N} .
 - $s_2 : \mathbb{N} \rightarrow \mathbb{N}, a \mapsto a+2$.
 - $\cdot : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}, (a, b) \mapsto a \cdot b$.

These are all L -structures where

$$L = \underbrace{\{R_2\}}_{\text{binary}} \cup \underbrace{\{f_1\}}_{\text{unary}} \cup \underbrace{\{f_2\}}_{\text{binary}} \cup \underbrace{\{c\}}_{\text{constant}}$$

Let $R_2(v_1, v_2) \in \text{Form}(L)$. Then

$$\begin{aligned} R_2^{A_1}(v_1, v_2) &= v_1 \dot{\in} v_2 \\ R_1^{A_1}(v_1, v_2) &= v_1 < v_2 \end{aligned}$$

Some L -terms are $c, v_i, f_1 v_i, f_1 f_1 v_i$ and $f_2 f_1 f_1 c f_1 c$. Then

$$\begin{aligned} c^{A_1} &= 0 \\ c^{A_2} &= 0 \\ c^{A_3} &= 1 \\ (f_1 c)^{A_1} &= f_1^{A_1} 0 = s_1(0) = 1 \\ (f_1 c)^{A_2} &= f_1^{A_2} 0 = s_1(0) = 1 \\ (f_1 c)^{A_3} &= f_1^{A_3} 1 = s_2(1) = 3 \end{aligned}$$

Let $\beta_1 : V = \{v_1, \dots, v_m\} \rightarrow m$ be an A_1 -assignment defined by $\beta_1(v_i) = i$. Then

$$\begin{aligned} R_2^{A_1}(v_1, v_2)[\beta_1] &= R_2^{A_1}(v_1[\beta_1], v_2[\beta_1]) \\ &= R_2^{A_1}(1, 2) \\ &= 1 \dot{\in} 2 \end{aligned}$$

thus $A_1 \not\models R_2 v_1 v_2 [\beta_1]$.

Define $\psi = (\forall x (\neg(x \dot{=} c) \rightarrow \exists y (f_1(y) \dot{=} x))) \in \text{Form}(L)$. As this has no free variables, this is a sentence. It can be shown that $A_1 \models \psi, A_2 \models \psi$. However, $A_3 \not\models \psi$ since we have $\neg(0 \dot{=} 1)$ and yet that there is no y such that $f_1^{A_3}(y) = 0$.

Exercise:

Consider the group structure $\mathcal{G} = (G, =, \cdot, ^{-1}, e)$. \mathcal{G} models the following sentences.

- $\forall x \forall y \forall z ((x \cdot y) \cdot z) = (x \cdot (y \cdot z))$.
- $\forall x x \cdot e = e \cdot x = x$.
- $\forall x x \cdot x^{-1} = e = x^{-1} \cdot x$.

A natural question is whether

$$\forall x \forall y x \cdot y = y \cdot x$$

2.6 Variable scoping

Definition 2.22:

Let $\phi \in \text{Form}(L)$, $t \in \text{Term}(L)$ and $v_i \in V$ a variable in ϕ . Then we say t is free from v_i in ϕ if for every variable $v_j \in t$, v_i is not in the scope of $\forall v_j$ nor $\exists v_j$ in ϕ .

Example:

Some examples of variable scoping.

: Consider the language $L = \{<\}$ and the formula $\phi = \forall x x < y$.

The term $t = x$ is not free from y in ϕ as y appears in the scope of $\forall x$.

: Consider the language $L = \underbrace{\{f\}}_{\text{3-ary}} \cup \underbrace{\{R\}}_{\text{binary}} \cup \underbrace{\{K\}}_{\text{unary}}$.

Let $t = f_{x_1 x_2 x_3}$ and $\phi = (\forall x_2 R_{x_2 x_4} \rightarrow K_{x_1}) \rightarrow (\forall x_1 R_{x_1 x_2})$. Then t is not free from x_2 (or x_4) in ϕ . But for $\psi = (\forall x_4 R_{x_4 x_2}) \rightarrow (\forall x_1 K_{x_1})$, t is free for x_4 in ψ .

Lemma 2.23:

Suppose $\phi \in \text{Form}(L)$, v_i is a free variable in ϕ and let t be an L -term that is free from v_i in ϕ . Then

$$\models ((\forall v_i \phi(v_i)) \rightarrow \phi(t))$$

Proof. For all L -structures \mathcal{A} and all \mathcal{A} -assignments β , we want to show

$$\mathcal{A} \models ((\forall v_i \phi(v_i)) \rightarrow \phi(t))[\beta]$$

Suppose not, then we have

$$\begin{aligned} v^\beta((\forall v_i \phi(v_i)) \rightarrow \phi(t)) &= 0 \\ \implies v^\beta(\forall v_i \phi(v_i)) &= 1 \text{ and } v^\beta(\phi(t)) = 0 \\ \implies \text{for all } a \in A, v^{\beta(a/v_i)}(\phi(v_i)) &= 1 \text{ and } v^\beta(\phi(t)) = 0 \end{aligned}$$

Letting $b = t^{\mathcal{A}}[\beta] \in A$, we have $v^{\beta(b/v_i)}(\phi(v_i)) = 1$ which is a contradiction. Thus

$$\mathcal{A} \models ((\forall v_i \phi(v_i)) \rightarrow \phi(t))[\beta]$$

which completes the proof. ■

Definition 2.24:

Suppose $\Gamma \subseteq \text{Form}(L)$, \mathcal{A} is an L -structure and β is an \mathcal{A} -assignment. We write $\mathcal{A} \models \Gamma[\beta]$ if $\mathcal{A} \models \psi[\beta]$ for all $\psi \in \Gamma$.

We write $\Gamma \models \phi$ if for all L -structures \mathcal{A} and \mathcal{A} -assignments β , $\mathcal{A} \models \Gamma[\beta] \implies \mathcal{A} \models \phi[\beta]$.

Recall [Lemma 2.15](#), we can bring an \mathcal{A} -assignment β into the statement and write, for example, $\mathcal{A} \models \neg \phi[\beta] \iff \mathcal{A} \not\models \phi[\beta]$.

The assumption in the lemma was that the consequences were sentences. However, we can modify the lemma and examine cases when the consequences are *not* sentences.

Exercise:

Find a counterexample to $\mathcal{A} \models \neg \phi[\beta] \iff \mathcal{A} \not\models \phi[\beta]$ for when ϕ is not a sentence.

Lemma 2.25:

Suppose $\phi(v_1, \dots, v_n) \in \text{Form}(L)$. Then

$$\mathcal{A} \models \phi(v_1, \dots, v_n) \iff \forall v_1 \dots v_n \phi(v_1, \dots, v_n)$$

where $\forall v_1 \dots v_n = \forall v_1 \forall v_2 \dots \forall v_n$.

Proof. If for all \mathcal{A} -assignments β , we have

$$\begin{aligned} \mathcal{A} \models \phi(v_1, \dots, v_n)[\beta] &\iff \mathcal{A} \models \phi[a_1, \dots, a_n] \text{ where } a_i = \beta(v_i) \text{ for all } \beta : V \rightarrow A \\ &\iff \mathcal{A} \models \phi[a_1, \dots, a_n] \text{ for all } (a_i)_{i=1}^n \subset \text{dom } \mathcal{A} \\ &\iff v^{\beta(a_1/v_1, \dots, a_n/v_n)}(\phi(v_1, \dots, v_n)) \\ &\iff \mathcal{A} \models \phi(v_1, \dots, v_n)[\beta(a_1/v_1, \dots, a_n/v_n)] \\ &\iff \mathcal{A} \models \forall v_1 \dots v_n \phi(v_1, \dots, v_n)[\beta] \end{aligned}$$

Which completes the proof. ■

Lemma 2.26:

Let $\phi \in \text{Form}(L)$. Then

- (i) $\models \neg \forall x \phi \leftrightarrow \exists x \neg \phi$.
- (ii) $\models \neg \exists x \phi \leftrightarrow \forall x \neg \phi$.
- (iii) $\models \forall x \phi \leftrightarrow \neg(\exists x \neg \phi)$
- (iv) $\models \exists x \phi \leftrightarrow \neg(\forall x \neg \phi)$

Proof. Easy exercise. ■

Corollary 2.27:

Let $\phi \in \text{Form}(L)$, then there exist $\psi \in \text{Form}(L)$ (with the same number of free variables) s.t. $\models \phi \leftrightarrow \psi$ and ψ has only the connectives $\{\wedge, \rightarrow, \perp, \forall\}$.

Noting that

$$\models \neg \phi \leftrightarrow (\phi \rightarrow \perp)$$

we see that this corollary holds.

Lemma 2.28:

Suppose $\phi, \theta \in \text{Form}(L)$ and $v \in V$ is not a free variable in θ . Then

$$\models \forall v (\theta \rightarrow \phi(v)) \implies \models \theta \implies \forall v \phi(v)$$

Proof. For \mathcal{A} an L -structure and β an \mathcal{A} -assignment. We want to show

$$\mathcal{A} \models \forall v (\theta \rightarrow \phi(v))[\beta] \implies \mathcal{A} \models \theta \implies \forall v \phi(v)[\beta]$$

Suppose $v^\beta(\forall v (\theta \rightarrow \phi(v))) = 1$ and that $v^\beta(\theta \rightarrow \forall v \phi(v)) = 0$. Then $v^\beta(\theta) = 1$ and $v^\beta(\forall v \phi(v)) = \min\{v^{\beta(a/v)}(\phi(v)) : a \in A\} = 0$. Thus $\exists a \in A$ s.t. $\mathcal{A} \models \neg \phi(v)[a/v]$.

By assumption we have $v^{\beta(b/v)}(\theta \rightarrow \phi(v)) = 1$ for all $b \in A$. So if $v^\beta(\theta) = 1$, we have $v^{\beta(b/v)}(\phi(v)) = 1$ for all $b \in A$. This raises a contradiction when $b = a$. ■

2.7 Natural Deduction

As in propositional logic we wish to make sense of ‘deducing’ one statement from another. This was done by defining certain rules that can be used to convert one set of statements into another. We define natural deduction for predicate logic similarly. In fact, we only introduce two additional rules to the set defined in [subsection 1.4](#).

\forall introduction If we can deduce $\frac{D}{\phi(x)}$ and x is not free for any uncanceled hypothesis of D , then we can write

$$\frac{D}{\phi(x)}$$

This is called \forall *introduction*.

\forall elimination If we can deduce $\forall v \phi(v)$, then for $t \in \text{Term}(L)$ which is free from v in ϕ , we can write

$$\frac{\forall v \phi(v)}{\phi(t)} (\forall E)$$

This is called \forall *elimination*.

We define syntactic consequence as done in [Definition 1.18](#).

Lemma 2.29: Soundness

Let $\Gamma \subseteq \text{Form}(L)$. Then

$$\Gamma \vdash \phi \implies \Gamma \models \phi$$

Proof. As before, we induct on the length of proofs.

All the rules apart from $\forall I$ and $\forall E$ are exactly the same as in propositional logic (check this as an exercise). It therefore remains to check the soundness of $\forall I$ and $\forall E$. ■

Before we prove this, we provide an additional lemma.

Lemma 2.30:

Let $\Gamma \subseteq \text{Form}(L)$.

$$\Gamma \vdash \phi(v) \implies \Gamma \vdash \forall v \phi(v)$$

if v is **not** a free variable (it cannot be a bound variable since it's free in ϕ) in Γ .

Which is an obvious conclusion of the \forall introduction rule.

Returning to the proof of soundness. We need to show $\Gamma \vdash \phi(x) \implies \Gamma \models \forall x \phi(x)$ (a conclusion of the previous lemma).

Suppose that for all \mathcal{A} an L -structure and β an \mathcal{A} -assignment we have $\mathcal{A} \models \phi(x)[\beta]$. It follows using a similar method to [Lemma 2.25](#) that $\mathcal{A} \models (\forall x \phi(x))[\beta]$.

Exercise:

Use [Lemma 1.25](#) to prove soundness for \forall introduction.

We have previously show that \forall introduction is a sound deduction rule.

Exercise:

Using the same method as [Lemma 2.23](#), show that \forall elimination is sound.

Exercise:

Some natural deductions.

| : Show $\vdash (\forall v (\phi \wedge \psi)) \rightarrow ((\forall v \phi) \wedge (\forall v \psi))$.

| : Suppose that v is not a free variable of θ . Show

$$\vdash (\forall v (\theta \rightarrow \phi(v)) \rightarrow (\theta \rightarrow \forall v (\phi(v))))$$

| : Suppose that v is not free in ϕ . Show that

$$\phi \leftrightarrow \forall v \phi$$

2.8 Completeness and the model existence lemma

As before, would like to show that $\Gamma \models \phi \implies \Gamma \vdash \phi$ (completeness). This is again done by showing that $\Gamma \not\models \phi \implies \Gamma \not\vdash \phi$. A result that we first need to assume is the *model existence lemma*.

Lemma 2.32: Model existence

If Γ is a set of L -sentences which is consistent ($\Gamma \not\vdash \perp$), then there is an L -structure \mathcal{A} (a model) where $\mathcal{A} \models \Gamma$.

Sketch proof of completeness. Recall that $\Gamma \subseteq \text{Form}(L)$ is *inconsistent* if $\Gamma \vdash \perp$ and that for Γ a consistent set of formulas $\Gamma \not\models \phi \implies \Gamma \cup \{\neg\phi\}$ is consistent ([Lemma 1.25](#)).

With this suppose that $\Gamma \not\models \phi$ i.e. $\Gamma \cup \{\neg\phi\}$ is consistent. Then we have a model \mathcal{A} such that $\mathcal{A} \models \Gamma \cup \{\neg\phi\} \implies \mathcal{A} \not\models \phi$. ■

Definition 2.31:

Let L be a language.

2.31.1: A *theory* is a collection of L -sentences $T \subseteq \text{Form}(L)$ such that

$$T \vdash \phi \implies \phi \in T$$

(T is closed under deductions).

2.31.2: Suppose T is a theory, we call $\Gamma \subseteq \text{Form}(L)$ a set of *axioms* of T if $T = \{\phi \in \text{Form}(L) \text{ a sentence} : \Gamma \vdash \phi\}$.

2.31.3: A theory T is called *Henkin theory* if for each sentence of the form $\exists x \phi(x)$, there is a constant $c \in L$ such that

$$(\exists x \phi(x) \rightarrow \phi(c)) \in T$$

Such a constant is called a *Henkin witness*.

2.31.4: Suppose T and T' are two theories of two languages L and L' respectively. T' is an *extension* of T if $T \subseteq T'$.

So to prove completeness, we just need to prove the model existence lemma. The broad steps are as follows:

- Expand the language L to L^* by constants and expand Γ to Γ^* where Γ^* is a Henkin theory in L^* .
- By the Lindenbaum lemma (an extension of [Lemma 1.31](#)) we extend Γ^* in L^* to a maximally consistent set.
- Build the structure $\mathcal{A} = \{t : t \text{ a constant term in } L^*\}$ (the model existence lemma).

2.8.1 Extension by constants and Henkin extensions

Definition 2.33:

Let T be an L -theory. For each formula of the form $\theta = \exists v \phi(v)$ in L , add a constant c_θ such that distinct θ s yield distinct c_θ s.

Call the resulting language L^* and define T^* to be the theory of axioms

$$T \cup \{\exists v \phi(v) \rightarrow \phi(c_\theta) : \theta = \exists v \phi(v) \text{ is a sentence}\}$$

We call such an extension an *extension by constants*.

We write $L^* = L \cup C$, $C = \{c_\theta : \theta \text{ is a sentence of the form } \exists v \phi(v)\}$.

Lemma 2.34:

Let T be a theory and T^* the extension by constants. Then

$$T = \text{Form}(L) \cap T^*$$

i.e. the only L -formulas (those that do not rely on the new constants c) in T^* are those already in T . To prove we use the following lemma.

Lemma 2.35:

Let Γ be a set of L -sentences. If

$$\Gamma \cup \{\exists v \phi(v) \rightarrow \phi(c)\} \vdash \psi$$

where c is a constant not occurring in Γ and ψ , then

$$\Gamma \vdash \psi$$

We leave the proof of this lemma till later.

Proof, Lemma 2.34. Since T is set of L -sentences closed under deductions ($T \vdash \psi \implies \psi \in T$), it is enough show that for an L -sentence ψ , $T^* \vdash \psi \implies T \vdash \psi$.

This implies

$$T \cup \{\sigma_1, \dots, \sigma_n\} \vdash \psi$$

where σ_i are the new axioms. Note that each σ_i will be in the form required by Lemma 2.35. Moreover since each ψ is an L -sentence, it will not contain a new constant from L^* .

Proceed by induction on n . We immediately see that this holds for $n = 0$. Suppose that $T \cup \{\sigma_1, \dots, \sigma_{k+1}\} \vdash \psi$ and let $T_k = T \cup \{\sigma_1, \dots, \sigma_k\}$.

By applying Lemma 2.35 and our inductive hypothesis, we conclude that $T \vdash \psi$. ■

Corollary:

If T is consistent, then T^* is consistent.

Proof. Suppose that T is consistent. \perp is an L -sentence which doesn't contain any constants. By Lemma 2.34 we have that if $T^* \vdash \perp$ then $T \vdash \perp$ which is a contradiction. ■

A natural question to ask at this point is whether T^* is a Henkin theory? This is not necessarily true and so we must define another kind extension that *does* yield a Henkin theory.

Definition 2.36:

Let $T_0 = T$ and inductively define $T_{n+1} = (T_n)^*$, which is a set of sentences in an L_n -language.

Define $T_\omega = \bigcup_{n \in \mathbb{Z}_{\geq 0}} T_n$ which is a theory in an L_ω -language.

Lemma 2.37:

T_ω is a Henkin theory.

Proof. We show that T_ω is in fact a theory and that it is Henkin.

(I) Suppose that $T_\omega \vdash \psi$, we want to show that $\psi \in T_\omega$.

- We then have $\{\sigma_1, \dots, \sigma_r\} \vdash \psi$ where each $\sigma_i \in T_{n_i}$.
- Note that we have $T_i \subseteq T_{i+1} \subseteq \dots$.

- Let $m = \max\{n_i : i \leq r\}$.
 - Note that for all $i \leq m$, for all $\sigma_i \in T_i$ we have $\sigma_i \in T_m$.
 - This implies $T_m \vdash \psi$ and so $\psi \in T_m$ since T_m is a theory.
 - Thus T_ω is a theory.
- (II) Suppose that $\psi = \exists v \phi(v)$ is an L_n -sentence for some $n \in \mathbb{Z}_{\geq 0}$. By the construction of the T_i s we have $c \in L_{n+1}$ such that $(\exists v \phi(v) \rightarrow \phi(c)) \in T_{n+1} \subseteq T_\omega$. Thus T_ω is a Henkin theory. ■

Lemma 2.38:

Let T be a theory, then

$$T_\omega \cap \text{Form}(L) = T$$

Proof. Again, it is enough to show that $T_\omega \vdash \psi \implies T \vdash \psi$ for an L -sentence ψ .

$T_\omega \vdash \psi \implies T_n \vdash \psi$, by induction we have $T \vdash \psi$. ■

Corollary 2.39:

When T is consistent, then there is a consistent Henkin extension T_ω of T .

Remark:

We are assuming that $|L|$ is countable which implies that $|\text{Form}(L)|$ is also countable. Thus $|L_\omega|$ is also countable.

We must now extend our theory so that it is maximally consistent, i.e. any additional sentence will render the theory inconsistent.

Lemma 2.40: Lindenbaum

Every consistent theory is contained in a maximally consistent theory within the same language.

Proof. Extend the proof of [Lemma 1.31](#). Requires use of Zorn's lemma. ■

We confirm that the Lindenbaum extension is in fact a Henkin theory.

Lemma 2.41:

Any extension of a Henkin theory in the same language is a Henkin theory.

Exercise:

Prove the above.

All that remains now is to construct the structure as required by the model existence lemma.

Let our maximally consistent Henkin extension of T be T^m . Consider $A = \{t : t \text{ is a closed } L_\omega\text{-term}\}$, i.e. t is a term without any variables. We define a structure on this set of symbols.

- For all $c \in L_\omega$, $c^A = c$.
- For all $f t_1 \cdots t_n$, $f^A t_1 \cdots t_n = f t_1 \cdots t_n$. Note this is well defined as $f : \text{Term}(L_\omega)^n \rightarrow \text{Term}(L_\omega)$ can be restricted to A^n .
- We define that $R^A t_1 \cdots t_n$ holds iff $T^m \vdash R t_1 \cdots t_n$.

We complete this proof later.

2.8.2 The proof of Lemma 2.35

Suppose $\Gamma \cup \{\exists v \phi(v) \rightarrow \phi(c)\} \vdash \psi$ where c a constant does not occur in Γ or ψ . We want to show $\Gamma \vdash \psi$.

(I) Suppose that $\Gamma \vdash \phi$ and that v is a variable not occurring in either Γ and ϕ . Then $\Gamma \vdash \phi[v/c]$.

Sketch. A proof of $\Gamma \vdash \phi$ is still valid on replacing v with c . A formal proof would utilise induction on proof length. ■

(II) If c does not occur in Γ , then $\Gamma \vdash \phi(c)$ implies $\Gamma \vdash \forall v \phi(v)$.

Proof. Apply the previous statement and the $\forall I$ rule. ■

(III) If x is not a free variable ψ , then

$$\vdash (\forall x (\phi_1(x) \rightarrow \psi)) \rightarrow (\exists x \phi_1(x) \rightarrow \psi)$$

Proof. We note that $\vdash (\theta \rightarrow \xi) \leftrightarrow (\neg \xi \rightarrow \neg \theta)$. So we will show that

$$\vdash (\forall x (\neg \psi \rightarrow \neg \phi_1(x))) \rightarrow (\neg \psi \rightarrow \neg \exists x \phi_1(x))$$

$$\frac{\frac{\frac{\forall x (\neg \psi \rightarrow \neg \phi_1(x))}{\neg \psi \rightarrow \neg \phi_1(c)} (\forall E) \quad \neg \psi}{\neg \phi_1(c)} (\text{MP}) \quad \neg \phi_1(c)}{\frac{\forall x \neg \phi_1(x) : \neg \exists x \phi_1(x)}{\neg \psi \rightarrow \neg \exists x \phi_1} (\forall I)} (\rightarrow I)$$

(IV) We are now in a position to prove the lemma. ■

Proof, Lemma 2.35. Suppose we have $\Gamma \vdash (\exists v \phi(v) \rightarrow \phi(c)) \rightarrow \psi$.

- By a previous statement, this implies $\Gamma \vdash \forall y ((\exists v \phi(v) \rightarrow \phi(y)) \rightarrow \psi)$ where y is a variable not occurring in Γ, ψ, ϕ .
- This implies $\Gamma \vdash (\exists y (\exists v \phi(v) \rightarrow \phi(y))) \rightarrow \psi$.
- With a little work, it can be show that $\Gamma \vdash (\exists v \phi(v) \rightarrow \exists y \phi(y)) \rightarrow \psi$.
- It is clear that $\vdash (\exists v \phi(v) \rightarrow \exists y \phi(y))$ and so we have $\Gamma \vdash \psi$. ■

2.8.3 Proof of the model existence lemma

Recall that given a theory $T \subseteq \text{Form}(L)$, we define $T_\omega \subseteq \text{Form}(L_\omega)$ to be the Henkin extension of the theory in the extended language L_ω . $T^m \subseteq \text{Form}(L_\omega)$ is then defined to be the maximally consistent extension of T_ω , and thus the maximally consistent Henkin extension of T .

The domain of our potential structure is defined to be $A = \{t : t \text{ is a closed } L_\omega\text{-term}\}$. Recall that a term is closed if it contains no variables.

And so we define \mathcal{A} by:

- For all $c \in L_\omega$, $c^{\mathcal{A}} = c$.
- For all $f t_1 \cdots t_n$, $f^{\mathcal{A}} t_1 \cdots t_n = f t_1 \cdots t_n$. Defined as a restriction.
- $R^{\mathcal{A}} t_1 \cdots t_n$ holds iff $T^m \vdash R t_1 \cdots t_n$.

Remark:

Note that for $\phi \in \text{Form}(L_\omega)$ a sentence

$$\mathcal{A} \models \phi \iff T^m \vdash \phi$$

is not an obvious fact.

Proof. We proceed using induction on $\text{Form}(L_\omega)$.

- The first case is \perp . This is clear because both $T^m \not\vdash \perp$ and $\mathcal{A} \not\models \perp$.
- Suppose that for $R \in L_\omega$ we have $\phi = R(t_1, \dots, t_m)$ where $t_i \in A$ for all i . We clearly have $\mathcal{A} \models \phi$, $T^m \vdash \phi$ follows from the interpretation of R in \mathcal{A} .
- Suppose $\phi = \neg\psi$ where the inductive hypothesis holds for ψ .

$$\begin{aligned}\mathcal{A} \models \phi &\iff \mathcal{A} \models \neg\psi \\ &\iff \mathcal{A} \not\models \psi \\ &\iff T^m \not\vdash \psi\end{aligned}$$

since T^m is maximally consistent, we have $T^m \vdash \neg\psi \iff T^m \vdash \phi$.

- Suppose $\phi = \theta \wedge \psi$ where both θ and ψ satisfy the inductive hypothesis.

$$\begin{aligned}\mathcal{A} \models \phi &\iff \mathcal{A} \models \theta \wedge \psi \\ &\iff \text{both } \mathcal{A} \models \{\theta, \psi\} \\ &\iff \text{both } T^m \vdash \{\theta, \psi\} \text{ (inductive hypothesis)} \\ &\iff T^m \vdash \theta \wedge \psi \\ &\iff T^m \vdash \phi\end{aligned}$$

- Suppose $\phi = \forall x \psi(x)$ where ψ satisfies the inductive hypothesis.

$$\begin{aligned}\mathcal{A} \models \phi &\iff \mathcal{A} \models \forall x \psi(x) \\ &\iff \mathcal{A} \not\models \exists x \neg\psi(x) \\ &\iff \mathcal{A} \not\models \neg\psi[a] \text{ for all } a \in A \\ &\iff \mathcal{A} \models \psi[a] \text{ for all } a \in A\end{aligned}$$

Note that $\mathcal{A} \not\models \neg\psi[a]$ for all $a \in A$ but in particular for c_θ where $\theta = \exists x \neg\psi(x)$. Since we have $\mathcal{A} \models \psi[c_\theta]$, we can use the induction hypothesis to conclude $T^m \vdash \psi(c_\theta)$. Since T^m is a Henkin theory we have $T^m \vdash \exists x \neg\psi(x) \rightarrow \neg\psi(c_\theta)$ where $\theta = \exists x \neg\psi(x)$. Consequently, we have $T^m \vdash \psi(c_\theta) \rightarrow \forall x \psi(x)$.

Thus

$$T^m \vdash \{\psi(c_\theta), \psi(c_\theta) \rightarrow \forall x \psi(x)\} \implies T^m \vdash \forall x \psi(x)$$

For the other direction, suppose $T^m \vdash \forall x \psi(x)$. Then for all t where t is free for x in $\psi(x)$ we have $T^m \vdash \psi(t)$, particularly for all closed terms t of L_ω .

Thus $\mathcal{A} \models \psi[t]$ for all $t \in A$ which implies $\mathcal{A} \models \forall x \psi(x)$.

- Consider $\phi = \theta \rightarrow \psi$ where θ, ψ satisfy the inductive hypothesis.

- $\mathcal{A} \models \theta \rightarrow \psi$ if and only if $\mathcal{A} \models \theta \implies \mathcal{A} \models \psi$.
- By the induction hypothesis we have $T^m \vdash \theta \implies T^m \vdash \psi$.
- Thus

$$\frac{\begin{array}{c} T^m \vdash \theta \\ \vdots \\ T^m \vdash \psi \end{array}}{T^m \vdash \theta \rightarrow T^m \vdash \psi : T^m \rightarrow \theta \rightarrow \psi} (\rightarrow I)$$

Which completes the proof. ■

This proof omits the completeness of the $\dot{=}$, but this is not an ordinary relation or connective. The issue is that we never defined deductive rules for $\dot{=}$.

We now define the deductive rules for $\dot{=}$.

- $\frac{}{x \dot{=} x}$ RI1
- $\frac{x \dot{=} y}{y \dot{=} x}$ RI2
- $\frac{x \dot{=} y \quad y \dot{=} z}{x \dot{=} z}$ RI3
- $\frac{x_1 \dot{=} y_1 \quad x_1 \dot{=} y_1 \cdots x_n \dot{=} y_n \quad \phi(x_1, \dots, x_n)}{\phi(y_1, \dots, y_n)}$ RI4
- The fourth rule can also be written as $\frac{x_1 \dot{=} y_1 \quad x_1 \dot{=} y_1 \cdots x_n \dot{=} y_n \quad t(x_1, \dots, x_n)}{t(x_1, \dots, x_n) \dot{=} t(y_1, \dots, y_n)}$ RI4, note that this is *equivalent* to the formulation above.

Soundness and completeness of these rules are easy and shall be proven shortly.

Another implication of these rules is that we can construct a new relation between terms. For $t, s \in A$, $t \sim s \iff T^m \vdash t \dot{=} s$. This is an equivalence relation due to the rules we have just defined, so the quotient structure is well defined.

Explicitly $A/\sim = \langle A/\sim, \dots \rangle$ where:

- For $c \in L_\omega$, $c^{A/\sim} = [c^A]$.
- For $f \in L_\omega$

$$f^{A/\sim} : A/\sim \rightarrow A/\sim$$

$$f^{A/\sim}(t_1^{A/\sim}, \dots, t_{n_f}^{A/\sim}) = [f^A(t_1^A, \dots, t_{n_f}^A)]$$

- For $R \in L_\omega$

$$R^{A/\sim}(t_1^{A/\sim}, \dots, t_{n_R}^{A/\sim}) \iff (t_1^A, \dots, t_{n_R}^A) \in R^A$$

This means that the structure defined in the model existence lemma is unique up to this equivalence. Returning to the rules of identity defined previously.

Example:

Some examples of the rules of identity.

$$\begin{aligned} & : \frac{x \dot{=} y \quad x^2 + y^2 > 2}{2y^2 > 2} \text{ RI4} \\ & : \frac{x \dot{=} y \quad x^2 + y^2 > 2}{2x^2 > 2} \text{ RI4} \end{aligned}$$

A few important sentences concerning $\dot{=}$ are given below:

- I1: $\forall x \, x \dot{=} x$
- I2: $\forall x \forall y \, x \dot{=} y \leftrightarrow y \dot{=} x$
- I3: $\underbrace{\forall x \forall y \forall z}_{\forall x y z} \, x \dot{=} y \wedge y \dot{=} z \rightarrow x \dot{=} z$
- I4: $\forall x_1 \cdots x_n \forall y_1 \cdots y_n \, (\bigwedge_{i=1}^n x_i \dot{=} y_i \rightarrow t(x_1, \dots, x_n) \dot{=} t(y_1, \dots, y_n))$ for a term t depending on variables x_i .
- Note that I4 is equivalent to writing $\forall x_1 \cdots x_n \forall y_1 \cdots y_n \, (\bigwedge_{i=1}^n x_i \dot{=} y_i \wedge \phi(x_1, \dots, x_n) \rightarrow \phi(y_1, \dots, y_n))$ for a formula ϕ depending on variables x_i .

Lemma 2.42:

For each of the Ij above, we have $\vdash Ij$.

Proof. This follows from the rules of \forall or just by using [Lemma 2.30](#) on our rules for $\dot{=}$. ■

It follows from [Lemma 2.42](#) that the relation defined in our proof of completeness is indeed an equivalence relation.

2.9 Soundness of equality deduction rules

Showing that the rules RI1 through RI3 are sound is trivial. They all rely on the fact that

$$\vdash x \dot{=} x \implies \models x \dot{=} x$$

Noting that $\models x \dot{=} x$ is equivalent to stating that for all \mathcal{A} an L -structure and all β an \mathcal{A} -assignment we have $\mathcal{A} \models x \dot{=} x[\beta]$.

Proof, soundness of RI4. We will use the formulation of RI4 for terms.

- If t is a constant then there is nothing to prove and we are done.
- For an L -structure \mathcal{A} we automatically have $f^{\mathcal{A}}(x_1, \dots, x_n) = f^{\mathcal{A}}(y_1, \dots, y_n)$ if $x_1 \dot{=} y_1 \cdots x_n \dot{=} y_n$.
- Using induction on formulas, proving RI4 is an easy exercise. ■

Recall our definition of the substitution of a term with a variable. Given $t \in \text{Term}(L)$, v a variable and $\phi \in \text{Form}(L)$. $\phi[t/v]$ is defined inductively as:

- $\perp[t/v] = \perp$.
- $Rt_1 \cdots t_{n_i}[t/v] = Rt_1[t/v] \cdots t_{n_i}[t/v]$.
- $(t_1 \dot{=} t_2)[t/v] = t_1[t/v] \dot{=} t_2[t/v]$.
- $(\phi \Box \psi)[t/v] = \phi[t/v] \Box \psi[t/v]$ (as in propositional logic).
- $(\forall v_i \phi(v_i))[t/v] = \begin{cases} \forall v_i (\phi[t/v]) : v \neq v_i \\ \forall v_i \phi(v_i) : v = v_i \end{cases}$.
- $(\exists v_i \phi)[t/v]$ is defined analogously.

There were issues in assigning semantic meaning to ϕ when it was not a sentence. These are now alleviated by the following lemma.

Lemma 2.43:

Suppose β is an \mathcal{A} -assignment for an L -structure \mathcal{A} . Then we have

$$v^{\beta}(\phi[t/v]) = v^{\beta[t^{\mathcal{A}}[\beta]/v]}(\phi(v))$$

Proof. Induction on formulas and use of the following lemma. ■

Lemma 2.44:

Given β an \mathcal{A} -assignment and $s, t \in \text{Term}(L)$ we have

$$(t[s/v])^{\mathcal{A}}[\beta] = t^{\mathcal{A}}[\beta(s^{\mathcal{A}}[\beta]/v)]$$

Proof. Induction on $\text{Term}(L)$.

- For $t = c$ a constant, the proof is clear.

$$\begin{aligned} t[s/v] &= c \\ c^{\mathcal{A}}[\beta] &= c^{\mathcal{A}} \\ &= t[\beta(s^{\mathcal{A}}[\beta]/v)] \\ &= c^{\mathcal{A}} \end{aligned}$$

- The same argument as above holds for $t = v_i$ a variable.
- Now suppose that $t = f t_1 \cdots t_m$. By the definition of substitution we have $t[s/v] = f t_1[s/v] \cdots t_m[s/v]$. Then

$$\begin{aligned}
 (t[s/b])^A[\beta] &= f^A(t_1[s/v])^A[\beta] \cdots (t_m[s/v])^A[\beta] \\
 &= f^A t_1^A[\beta(s^A[\beta]/v)] \cdots t_m^A[\beta(s^A[\beta]/v)] \\
 &= (f t_1 \cdots t_m)^A[\beta(s^A[\beta]/v)]
 \end{aligned}$$

■

Remark:

The above lemma was used in the proof of [Lemma 2.23](#)

2.10 Some conclusions of the completeness theorem

Theorem 2.45: Compactness

Suppose that Σ is a set of L -sentences. Then Σ has a model if and only if every finite subset of Σ has a model.

Proof. The only if direction is trivial. For the if direction, suppose that every finite subset of Σ has a model and suppose for a contradiction that Σ does not have a model. We then have $\Sigma \vdash \perp$ by completeness.

This implies that there is a finite subset $\Sigma_0 \subseteq \Sigma$ such that $\Sigma_0 \vdash \perp$. Which is a contradiction since an inconsistent theory cannot have a model. ■

Theorem 2.46: Downward Löwenheim-Skolem

Suppose L is a countable first order language and that \mathcal{A} is an L -structure. Let

$$T = \text{Th}(\mathcal{A}) := \{\theta : \mathcal{A} \models \theta \text{ and } \theta \in \text{Form}(L) \text{ is a sentence}\}$$

Then there is a *countable* model \mathcal{B} such that $\mathcal{B} \models T$.

Example:

Though this might seem strange, we can show that given

$$\mathcal{R} = (\mathbb{R}, +, \cdot, <, \exp, 0, 1)$$

which is interpreted as in real analysis, there is a countable model for $\text{Th}(\mathcal{R})$.

Proof, Theorem 2.46. Because L is a countable language, T is at most a countable set of L -sentences. By the proof of the completeness theorem, there is a countable model of T . ■

It can be shown by the “Lindström theorem” that predicate logic is the “strongest” logic that satisfies both compactness and Löwenheim-Skolem.

An example of the two theorems in use is by applying them to the Peano arithmetic.

Example:

Consider the structure $\mathcal{N} = \langle \mathbb{N}, <, \text{succ}, 0 \rangle$ where the interpretation is the usual one. The Peano arithmetic on a language $L = \{<\} \cup \{\text{succ}\} \cup \{0\}$ is the theory consisting of the following axioms.

$$P1 \bullet \forall x \forall y \ x \dot{=} y \leftrightarrow \text{succ}(x) \dot{=} \text{succ}(y)$$

$$P2 \bullet \forall x \ x < \text{succ}(x)$$

- $P3 \bullet \neg(\exists y \text{ succ}(y) \dot{=} 0)$
 $P4 \bullet \forall x(\neg(x \dot{=} 0) \rightarrow \exists y (\text{succ}(y) \dot{=} x))$
 $P5 \bullet \forall x \neg(\text{succ}(x) \dot{=} x)$
 $P6 \bullet \forall x \exists y \text{ succ}(x) \dot{=} y$

Consider $\text{Th}(\mathcal{N})$ defined as above. It is clear that \mathcal{N} is a model for the P_i and so we have $\text{Th}(\mathcal{N}) \models P_i$ for $i = 1, \dots, 6$.

Consider the language $L_c = L \cup \{c\}$ where c is a new constant, one not in L . Write

$$t_m(x) := \underbrace{\text{succ}(\text{succ}(\dots \text{succ}(x) \dots))}_{m \text{ applications}}$$

e.g. $t_m^{\mathcal{N}}(0) = 2$. Now consider the formula $\phi_m = (t_m(0) < c) \in \text{Form}(L_c)$ and let $\Sigma = \{\phi_m : m \in \mathbb{N}\}$. We claim that $\Sigma \cup \text{Th}(\mathcal{N})$ has a model in L_c .

Proof. If we can show that every finite subset $\Sigma_0 \subseteq \Sigma \cup \text{Th}(\mathcal{N})$ has a model, then $\Sigma \cup \mathcal{N}$ will have a model by the compactness theorem.

- Let $\Sigma_0 = \Sigma_0^1 \cup \Sigma_0^2$ where $\Sigma_0^1 \subseteq \Sigma$ and $\Sigma_0^2 \subseteq \text{Th}(\mathcal{N})$.
- Note that $\text{Th}(\mathcal{N})$ is a theory and thus closed under deductions so we naturally have $\mathcal{N} \models \Sigma_0^1$ by soundness.
- We have $\Sigma_0^1 = \{\phi_{m_1}, \dots, \phi_{m_j}\}$ for some m_j . Explicitly $\phi_{m_i} = c > t_{m_i}(0)$ for all $1 \leq i \leq j$.
- Since $m_r < m_s$ implies $t_{m_r}(0) < t_{m_s}(0)$, if we have ϕ_m for $m = \max\{m_1, \dots, m_j\}$ then all the formulas in Σ_0^1 are satisfied.
- Consider the structure $\mathcal{N}_m = \langle \mathbb{N}, \text{succ}, 0, \underbrace{m+1}_{=c} \rangle$. Since $\mathcal{N} \subseteq \mathcal{N}_m$ we have $\mathcal{N}_m \models \Sigma_0^2$. Thus $\mathcal{N}_m \models \Sigma_0$.
- $c^{\mathcal{N}_m} = m+1 > t_m(0)$ which implies that $\mathcal{N}_m \models \Sigma_0^1$.
- So for every Σ_0 we have $\mathcal{N}_m \models \Sigma_0$ and so by compactness we have a model for Σ . ■

Note that \mathcal{N} , the ordered natural numbers with a successor function is *not* a model for $\Sigma \cup \text{Th}(\mathcal{N})$.

An explicit construction of such a model is $T_1 = \langle \mathcal{N} \dot{\cup} \mathbb{Z}, \dots \rangle$ where any element of \mathbb{Z} is greater than every element of \mathbb{N} . It can be shown that $\mathbb{N} \cup \mathbb{Z} \in T_1$ and \mathbb{N} are not L -isomorphic. Note that even though $T_1 \models \Sigma \cup \text{Th}(\mathcal{N})$, if we drop the interpretation of c we have $T_1^- \models \text{Th}(\mathcal{N})$. In general any such $T_m = \mathbb{Z} \dot{\cup}_{i \in m} \mathbb{Z}_i$ is a model for this theory of “numbers beyond the naturals”.

We can conclude that the statement “in the naturals, every natural number is obtained by finitely many applications of succ to 0” cannot be expressed by a first order statement. Suppose that such a statement θ exists and that $\mathcal{N} \models \theta$. We therefore have $\theta \in \text{Th}(\mathcal{N})$. Since we have $T_1 \models \text{Th}(\mathcal{N})$ we also have $T_1 \models \theta$ which is a contradiction.

This can be thought of as a non-standard model of the natural numbers. Similarly we can use compactness to prove the existence of non-standard models of $\mathcal{R} = (\mathbb{R}, +, \cdot, 1, 0, <)$ the real numbers. One property that is not satisfied in a non-standard model of the reals is the *Archimedean property*.

$$\forall \epsilon > 0 \exists n \in \mathbb{N} \frac{1}{n} < \epsilon$$

2.11 Decidability and the Entscheidungsproblem

Suppose that θ is a propositional formula, can we determine whether $\vdash \theta$ or $\vdash \neg \theta$? By the completeness theorem for propositional logic, this is equivalent to determining whether $\models \theta$ or $\not\models \theta$. This can be known since $\not\models \theta \iff$ there is a value $\bar{x} \in \{0, 1\}^n$ such that for the truth function $F_\theta : \{0, 1\}^n \rightarrow \{0, 1\}$, $F_\theta(\bar{x}) = 0$.

This can easily be found by computing the DNF of θ , so there is an algorithm to determine whether $\vdash \theta$ or $\vdash \neg\theta$. We say that finding whether a propositional formula is a tautology is *decidable*.

The *entscheidungsproben* is a question first posed by Hilbert and Ackermann that asks whether the same is true for predicate formulas. Specifically if there is an algorithm that will determine whether any given first order formula is valid? Church and Turing independently showed that there is no such algorithm. Their proofs are both based on *Gödel's incompleteness theorem*.

Definition:

A set of natural numbers S is *recursively enumerative* (RE) if there exists an algorithm that generates all the elements of S . This algorithm does not need to terminate.

Theorem: Gödel's incompleteness theorem

No consistent set of axioms that is RE is capable of generating all valid sentences of $\langle \mathbb{N}, +, \cdot, 0 \rangle = \mathcal{N}$.

Equivalently, there is no algorithm that could determine whether a formula of \mathcal{N} is valid since this would imply that the set of valid sentences of \mathcal{N} is recursively enumerative.

Remark:

The completeness and incompleteness theorem are **not** referring to the same property. The former refers to the ability to give semantic meaning to a deductive proof and vice versa. The latter refers to the failure of a logical system to determine the validity of every sentence.

There are, however, first order theories that are decidable. An example is the theory of abelian groups. A non-example is the theory of Diophantine equations.

2.12 Dense linear orders

Suppose L is the language containing a single binary relation \leq .

Definition 2.47:

We say an L -structure \mathcal{A} is a *linear order* if it satisfies the following axioms:

$$\phi 1 \bullet \forall x_1 \forall x_2 (x_1 \leq x_2) \wedge (x_2 \leq x_1) \rightarrow (x_1 = x_2)$$

$$\phi 2 \bullet \forall x_1 x_2 x_3 (x_1 \leq x_2) \wedge (x_2 \leq x_3) \rightarrow (x_1 \leq x_3)$$

$$\phi 3 \bullet \forall x_1 x_2 (x_1 \leq x_2) \vee (x_2 \leq x_1)$$

We say that a linear order is *dense* if further

$$\phi 4 \bullet \forall x_1 \forall z_2 \exists x_3 (x_1 < x_2) \rightarrow (x_1 < x_3) \wedge (x_3 < x_2)$$

where $x_1 < x_2$ is an abbreviation for $x_1 \leq x_2 \wedge \neg(x_1 = x_2)$.

We say that a dense linear order has *no end points* if further

$$\phi 5 \bullet \forall x_1 \exists x_2 x_1 < x_2$$

$$\phi 6 \bullet \forall x_1 \exists x_2 x_2 < x_1$$

Let $\Delta = \{\phi_1, \dots, \phi_6\}$. Note that we have two models for this theory, $Q := \langle \mathbb{Q}, \leq \rangle \models \Delta$ and $R := \langle \mathbb{R}, \leq \rangle \models \Delta$. It would be nice to show that the models are equivalent in some sense, a statement made rigorous below.

Proposition 2.48:

For every θ an L -sentence,

$$R \models \theta \iff Q \models \theta$$

Intuitively, a result about dense linear orders without endpoints reached in Q will hold for R and vice versa. We will spend the rest of this section proving this result.

A result that will be useful later on is that an isomorphism between linear orders preserves the order relation.

Exercise:

If \mathcal{A} and \mathcal{B} are two linear orders that are isomorphic, then there is a bijection $f : A \rightarrow B$ such that $a \leq^{\mathcal{A}} b \implies f(a) \leq^{\mathcal{B}} f(b)$.

Remark:

Recall that an isomorphism between L -structures \mathcal{A}, \mathcal{B} is a bijection $F : A \rightarrow B$ such that:

- For every constant $c \in L$, $F(c^{\mathcal{A}}) = c^{\mathcal{B}}$.
- For every m -ary function $f \in L$, $F(f^{\mathcal{A}}(a_1, \dots, a_m)) = f^{\mathcal{B}}(F(a_1), \dots, F(a_m))$.
- For every k -ary relation $R \in L$, $R^{\mathcal{A}}(a_1, \dots, a_k) \iff R^{\mathcal{B}}(F(a_1), \dots, F(a_k))$.

Homomorphisms and embeddings are defined similarly by relaxing the rules above.

An important result for the theory of dense linear orders is that it is decidable, this result is called the Łoś-Vaught test.

Lemma 2.49: Łoś-Vaught test

For every L -sentence θ either $\Delta \vdash \theta$ or $\Delta \vdash \neg\theta$.

Proof. Suppose not, then $\Delta_1 = \Delta \cup \{\theta\}$ and $\Delta_2 = \Delta \cup \{\neg\theta\}$ are both consistent.

- by Löwenheim–Skolem theorem, both have countable models \mathcal{A}_1 and \mathcal{A}_2 .
- By Cantor's theorem (2.50) \mathcal{A}_1 and \mathcal{A}_2 are isomorphic.
- This contradicts Corollary 2.51.

Theorem 2.50: Cantor

If \mathcal{A} and \mathcal{B} are two countable dense linear orders without endpoints, then \mathcal{A} and \mathcal{B} are isomorphic.

A proof of this uses a technique called *back and forth*. The corollary follows by the same method.

Corollary 2.51:

If \mathcal{A} and \mathcal{B} are isomorphic, then for all θ an L -sentence,

$$\mathcal{A} \models \theta \iff \mathcal{B} \models \theta$$

The proof of Proposition 2.48 is now clear.

Proof, Proposition 2.48. We want to show $Q \models \theta \iff \Delta \vdash \theta$. Note that $Q \models \Delta$ so \Leftarrow is given by the completeness theorem. To show \Rightarrow consider $\Delta \not\vdash \theta$. Then we have $\Delta \vdash \neg\theta$ and $Q \not\models \neg\theta$.

We prove $R \models \theta \iff \Delta \vdash \theta$ similarly.

3 Set theory

Set theory is a first order theory that rigorously defines sets and set operations and aims to formally study them. This first order theory is of particular interest because the notion of a set is used almost universally in other areas of mathematics and is thus of foundational importance. In this section we will develop the axioms of set theory and some important theorems.

3.0 Basic set theory (Review)

We begin with a introduction to naïve set theory. This doesn't phrase the theory on an axiomatic footing, but does define some useful concepts that will be made rigorous later.

We begin by defining only one operation on sets, that of *membership*. This allows us to state the axiom of extensionality.

Definition 3.0.0: The axiom of extensionality

Two sets A and B are equal iff

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

Thus set membership can be thought of as the defining property of a set since a set is uniquely by the elements it contains.

Membership also has other uses, for example it can define an order on the natural numbers as defined below.

Example: The natural numbers

Let $0 = \emptyset$ and define $n + 1 := \{0, 1, \dots, n\}$ for $n \geq 0$. With this we can write

$$m < n \iff m \in n$$

Our next construction is the power set.

Definition 3.0.1: The axiom of power set

If A is a set, then there exists a power set $\mathcal{P}(A)$ which satisfies

$$\forall x (x \subseteq A \rightarrow x \in \mathcal{P}(A))$$

where

$$(x \subseteq A) \iff (\forall y \in x \rightarrow y \in A)$$

Another way of constructing one set from another is the ordered pair.

Definition 3.0.2: Ordered pairs (Kuratowski)

We define the ordered pair (x, y) to be the set $\{\{x\}, \{x, y\}\}$. We write $A \times B = \{(a, b) : a \in A, b \in B\}$ and A^n for the n -fold product of A with itself.

Exercise:

Show that for any x, y, z, w we have $(x, y) = (z, w) \iff x = z$ and $y = w$.

Note that the set of finite sequences of elements from A is the union $\bigcup_{n \in \mathbb{N}} A^n$.

The notion of ordered pairs allows us to define functions $f : A \rightarrow B$ as subsets $f \subseteq A \times B$ such that

$$\forall x \in A \exists! b \in B ((a, b) \in f)$$

where $\exists! c \in C \phi(c) \iff \exists c \in C (\phi(c) \wedge (\forall d \in C (\phi(d) \rightarrow (d \dot{=} c)))$.

We define $A = \text{dom}(f)$ to be the *domain* of f and $B = \text{ran}(f)$ to be the *range* of f . The set of functions from A to B is denoted B^A and is a subset of $\mathcal{P}(A \times B)$. For a subset $X \subseteq A$ we write its image under f as $f[X] = \{f(x) : x \in X\}$.

3.1 Cardinality

We now turn to the subject of cardinality which aims to give each set a “size”.

Definition 3.1.1:

We say two sets A and B are *equinumerous* (or that they have the same *cardinality*) if there is a bijection $f : A \rightarrow B$. We write $A \approx B$ or $|A| = |B|$.

With this we can now distinguish the finite and countable sets.

Definition 3.1.2:

A set is *finite* if it is equinumerous with some element of \mathbb{N} . It is *countably infinite* if it is equinumerous with \mathbb{N} . It is *countable* if it is either *finite* or *countably infinite*.

Proposition 3.1.3:

Some basic facts about cardinalities.

- (i) Every subset of a countable set is countable.
- (ii) A set A is countable iff there is an injection $A \hookrightarrow \mathbb{N}$.
- (iii) If A and B are countable, then so is $A \times B$.
- (iv) If the axiom of choice is assumed and if A_0, A_1, \dots are countable. Then $\bigcup_{n \in \mathbb{N}} A_n$ is countable.

Exercise:

Show that \mathbb{R} is not countable.

Theorem 3.1.4: Cantor

If X is any set, then there is no surjective function $f : X \rightarrow \mathcal{P}(X)$. In particular we have $X \not\approx \mathcal{P}(X)$.

Proof, Cantor's diagonal argument. Suppose that such a surjection exists.

- Let $Y = \{y \in X : y \notin f(y)\} \subseteq X \implies Y \in \mathcal{P}(X)$.
- We therefore have $z \in X : f(z) = Y$.
- If $z \in Y$, then $z \notin f(z) \implies z \notin Y$.
- If $z \notin Y$, then $z \in f(z) \implies z \in Y$.

Definition 3.1.5:

For sets A, B we write $|A| \leq |B|$ if there is an injective function $f : A \hookrightarrow B$, so A is equinumerous with a subset of B .

Exercise:

Show that this relation is transitive.

Remark:

We have $|X| \leq |\mathcal{P}(X)|$ since we have the injection $x \mapsto \{x\}$. Since we also have $X \notin \mathcal{P}(X)$ we have $|X| < |\mathcal{P}(X)|$.

Theorem 3.1.6: Cantor-Schröder-Bernstein

Suppose that A, B are sets and that $f : A \rightarrow B$ and $g : B \rightarrow A$ are both injective functions. Then $A \approx B$. i.e

$$|A| \leq |B| \text{ and } |B| \leq |A| \implies |A| = |B|$$

Proof. Let $h = g \circ f : A \rightarrow A$.

- Define $A_0 = A \setminus g[B]$ and for $n > 0$, let $A_n = h[A_{n-1}]$. Let $A^* = \bigcup_{n \in \mathbb{N}} A_n$ and $B^* = f[A^*]$.
- Note that $h[A^*] \subseteq A^*$ so $g[B^*] = g(f[A^*]) = h[A^*] \subseteq A$.
- We claim that $g[B \setminus B^*] = A \setminus A^*$.
 - If this is true, then f is a bijection between A^* and B^* and g is a bijection between $B \setminus B^*$ and $A \setminus A^*$.
 - So $k(a) = \begin{cases} f(a) : a \in A^* \\ g^{-1}(a) : a \in A \setminus A^* \end{cases}$ is a bijection from A to B .
- We now prove the claim
 - Let $a \in A \setminus A^*$.
 - * As $a \notin A_0$ we have $b \in B$ such that $g(b) = a$.
 - * Then $b \notin B^*$ as

$$\begin{aligned} b \notin B^* &\implies b \notin f[A^*] \\ &\implies g(b) \notin h[A^*] \subseteq A^* \\ &\implies a \notin A^* \end{aligned}$$
 - * Thus $g[B \setminus B^*] \supseteq A \setminus A^*$
 - Let $b \in B$.
 - * Suppose that $g(b) \in A^*$, we will show that $b \in B^*$ which implies that $g[B \setminus B^*] \subseteq A \setminus A^*$.
 - * Clearly $g(b) \notin A_0$ so $g(b) \in A_n$ for some $n > 0$.
 - * This implies we have some $a \in A_{n-1}$ such that $g(b) = h(a)$.
 - * Thus $g(b) = g(f(a)) \implies b = f(a), a \in A^*$.
 - * And so $b \in f[A^*] = B^*$.

Example 3.7:

The following sets are equinumerous:

1. $S_1 =$ The set of all sequences of 0s and 1s, $\{0, 1\}^{\mathbb{N}}$
2. $S_2 = \mathbb{R}$
3. $S_3 = \mathcal{P}(\mathbb{N})$
4. $S_4 = \mathcal{P}(\mathbb{N} \times \mathbb{N})$

5. $S_5 = \mathbb{N}^{\mathbb{N}}$ the set of all sequences of natural numbers.

Proof. We will find injective functions $f_{i,j} : S_i \rightarrow S_j$ and use the previous theorem.

- We have $\mathbb{N} \approx \mathbb{N} \times \mathbb{N}$ so $S_3 \approx S_4$.
- Moreover $S_1 \subseteq S_5 \subseteq S_4$ so we have $f_{1,5}$, $f_{5,4}$ and $f_{1,4}$.
- There is a bijection $f_{3,1} : S_3 \rightarrow S_1$ given by

$$f_{3,1}(X) = (a_n)_{n \in \mathbb{N}}, a_n = \begin{cases} 0 & n \notin X \\ 1 & n \in X \end{cases}$$

- We define $f_{1,2} : (a_n)_{n \in \mathbb{N}} \mapsto 0.a_0a_1a_2 \dots$. This decimal expansion is clearly an injective function.
- $f_{2,3}$ is defined as

$$f_{2,3} : \mathbb{R} \rightarrow \mathcal{P}(\mathbb{Q}) \xrightarrow{\sim} \mathcal{P}(\mathbb{N})$$

$$r \mapsto \{q \in \mathbb{Q} : q < r\} \mapsto \dots$$

By considering the compositions of these functions, we get the required injective $f_{i,j}$. ■

Some questions that we can now ask are:

- If A and B are sets do we have $(|A| \leq |B|) \vee (|B| \leq |A|)$? We will show that if we assume the axiom of choice, this is true.
- Is there an $X : \mathbb{N} \subset X \subset \mathbb{R}$ such that $|\mathbb{N}| < |X| < |\mathbb{R}|$? This is the continuum hypothesis and it turns out to be *independent* of the other axioms of set theory.

3.2 Axioms for set theory

The axioms for set theory as a first order theory are called the *Zermelo-Fraenkel* axioms. The language for set theory has a single binary relation, set membership, in addition to equality. The axioms themselves give instructions on how to build new sets from “existing sets”.

ZF1: Extensionality $\forall x \forall y ((x = y) \leftrightarrow \forall z ((z \in x) \leftrightarrow (z \in y)))$

ZF2: Empty set $\exists x \forall y (y \notin x)$.

The unique set with this property is the empty set, \emptyset .

ZF3: Pairing $\forall x \forall y \exists z (\forall w (w \in z \leftrightarrow ((w \in x) \vee (w \in y))))$

This axiom allows us to construct the natural numbers and ordered pairs.

ZF4: Union For any set A there is a set $B = \bigcup A = \bigcup \{z : z \in A\}$.

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

ZF5: Power set For any set A , there is a set whose elements are the subsets of A , the *power set*. We write $z \subseteq A \iff \forall y ((y \in z) \rightarrow (y \in A))$.

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

ZF6: Axiom scheme of specification Suppose that $P(x, y_1, \dots, y_r) \in \text{Form}(L)$, then

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

The axiom scheme of specification effectively gives validity to the subset specification notation from everyday mathematics. Namely $B = \{x \in A : P(x, y_1, \dots, y_r \text{ holds})\}$.

An important example is the intersection, let $C \neq \emptyset$ and $A \in C$. Then

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

The Cartesian product of nonempty sets A and B becomes

$$A \times B = \{w \in \mathcal{P}(\mathcal{P}(A \cup B)) : \phi(w)\}$$

where $\phi(w)$ is the following propositional formula

$$\phi(w) = \exists x \exists y \forall z ((x \in A) \wedge (y \in B)) \wedge (w = \{\{x\}, \{x, y\}\})$$

we note that $w = \{\{x\}, \{x, y\}\}$ is a compact notation for the slightly cumbersome expression

$$\forall z ((z \in w) \leftrightarrow \underbrace{((\forall a ((a \in z) \leftrightarrow (a \dot{=} x)))}_{z=\{x\}} \vee \underbrace{(\forall a ((a \in z) \leftrightarrow ((a \dot{=} x) \vee (a \dot{=} y))))}_{z=\{x,y\}}))$$

Definition 3.2.1:

For a set a , the *successor* of a is the set

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

We say a set is *inductive* if

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

ZF7 Infinity $\exists A ((\emptyset \in A) \wedge (\forall x ((x \in A) \rightarrow (x^+ \in A)))$

Note that our inclusion of the empty set implies that we will always have a copy of $\{\emptyset, \emptyset^+, \emptyset^{++}, \dots\}$ in our inductive set.

Definition 3.2.2:

Letting A be any inductive set and writing $\phi(A) = (\emptyset \in A) \wedge (\forall x ((x \in A) \rightarrow (x^+ \in A)))$. We can then form

$$\mathbb{N} = \{x \in A : \phi(B) \rightarrow (x \in B)\}$$

This is informally the “intersection of all inductive sets”; note that it does not depend on the choice of A .

Remark: Some notation

We can also denote \mathbb{N} as ω .

Theorem 3.2.3:

Let \mathbb{N} be defined as above.

- (1) \mathbb{N} is an inductive set. Moreover for any inductive set B , $\mathbb{N} \subseteq B$.
- (2) Suppose that $P(x)$ is a formula such that
 - (i) $P(\emptyset)$ holds and
 - (ii) $\forall k ((k \in \mathbb{N}) \rightarrow (P(k) \rightarrow P(k^+)))$ holds.

Then $\forall k ((k \in \mathbb{N}) \rightarrow P(k))$

Proof. The first statement follows more or less immediately from the definition of \mathbb{N} . To show the second statement, consider $B \subseteq \mathbb{N}$ given by $B = \{x \in \mathbb{N} : P(x)\}$.

- By properties (i) and (ii) it is clear that B is an inductive set.
- So by the first statement $\mathbb{N} \subseteq B$.
- Thus $B = \mathbb{N}$.

■

We could then develop arithmetic in \mathbb{N} using n^+ as $n + 1$ which could then lead on to defining \mathbb{Z} , \mathbb{Q} and \mathbb{R} using ZF1-7. A slightly harder task is to show that the binary relation $m \leq n \iff (m \dot{=} n) \vee (m \in n)$ for $m, n \in \mathbb{N}$ is a *well ordering* on \mathbb{N} .

3.3 Linear orderings

Definition 3.3.1:

A linear ordering (A, \leq) is a *well ordering* (or a *wo set*) if every non-empty subset of A has a least element. i.e.

$$\forall X (((X \subseteq A) \wedge \neg(X \dot{=} \emptyset)) \rightarrow \exists x ((x \in X) \wedge (\forall y ((y \in X) \rightarrow (x \leq y))))$$

Example:

Some examples and non-examples of wo sets.

- (\mathbb{N}, \leq) is a wo set.
- (\mathbb{Z}, \leq) is not a wo set.

Definition 3.3.2:

Suppose $\mathcal{A}_i = (A_i, \leq_i)$ are two linear orderings for $i = 1, 2$. We say \mathcal{A}_1 and \mathcal{A}_2 are *similar* (or *isomorphic*) if there is a bijection $\alpha : A_1 \rightarrow A_2$ such that

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

We call α a *similarity* between \mathcal{A}_1 and \mathcal{A}_2 and write $\mathcal{A}_1 \simeq \mathcal{A}_2$.

If we only have the forward implication $\forall a, b \in A_1 a \leq_1 b \implies \alpha(a) \leq_2 \alpha(b)$, we say that α is *order preserving*.

Definition 3.3.3:

Let the \mathcal{A}_i be as defined previously

3.3.3.1: The *reverse-lexicographic product* $\mathcal{A}_1 \times \mathcal{A}_2$ is the linear order $(A_1 \times A_2, \leq)$ defined by

$$(a_1, a_2) \leq (a'_1, a'_2) \iff \text{either } a_2 <_2 a'_2 \text{ or } a_2 \dot{=} a'_2 \text{ and } a_1 \leq_1 a'_1$$

3.3.3.2: Consider the disjoint union of \mathcal{A}_1 and \mathcal{A}_2 ,

$$A_1 \sqcup A_2 = \{(a, i) \in (A_1 \cup A_2) \times \{0, 1\} : ((a \in A_1) \leftrightarrow (i \dot{=} 0) \wedge (a \in A_2) \leftrightarrow (i \dot{=} 1))\}$$

We define the *sum* to be $\mathcal{A}_1 + \mathcal{A}_2 = (A_1 \sqcup A_2, \leq)$ where $a_1 \leq a_2$ for all $a_1 \in A_1$ and $a_2 \in A_2$. All other orderings are as in \mathcal{A}_1 and \mathcal{A}_2 .

Intuitively, the reverse-lexicographic product replaces each element of \mathcal{A}_1 with a copy of \mathcal{A}_1 in place.

Example:

Consider $\{0, 1\} \times \mathbb{N}$. In this case we replace each element of \mathbb{N} by the two element set in place. The result is still similar to \mathbb{N} .

For $\mathbb{N} \times \{0, 1\}$ however, we have replaced each element of the two element set by a copy of \mathbb{N} , in a sense concatenating two copies of \mathbb{N} . This is not similar to \mathbb{N} .

The sum can be thought of as a true concatenation of the two orders.

Lemma 3.3.4:

Let \mathcal{A}_1 and \mathcal{A}_2 be two linear orders. Then

- (1) $\mathcal{A}_1 + \mathcal{A}_2$ and $\mathcal{A}_1 \times \mathcal{A}_2$ are both linear orders.
- (2) If \mathcal{A}_1 and \mathcal{A}_2 are well ordered sets, then so are $\mathcal{A}_1 + \mathcal{A}_2$ and $\mathcal{A}_1 \times \mathcal{A}_2$.

Proof. The proof for the first statement is left as an exercise.

Exercise:

Prove the first statement.

We only show that $\mathcal{A}_1 \times \mathcal{A}_2$ is a wo set since the proof for $\mathcal{A}_1 + \mathcal{A}_2$ follows the same lines.

- Let $X \subseteq \mathcal{A}_1 \times \mathcal{A}_2$ be a non-empty set.
- Consider $Y = \{b \in \mathcal{A}_2 : \exists a \in \mathcal{A}_1, (a, b) \in X\} \subseteq \mathcal{A}_2$.
- Since \mathcal{A}_2 , this has a least element $d \in \mathcal{A}_2$.
- Now consider $Z = \{x \in \mathcal{A}_1 : (x, d) \in X\} \subseteq \mathcal{A}_1$.
- This also has a least element c .
- It then follows that $(c, d) \in X$ is the least element of X .

■

3.4 Ordinals

Definition 3.4.1:

Let X be a set

- (i) We say X is *transitive* if every element of X is a subset of X . i.e. $\forall x ((x \in X) \rightarrow (\forall y ((y \in x) \rightarrow (y \in X))))$.
- (ii) X is *ordinal* if:
 - (a) X is a transitive set.
 - (b) The relation $<$ on X given by $x < y \iff x \in y$ is (strict) well-ordering.

Note that for an ordinal α we have, by definition, $\alpha \notin \alpha$.

Example:

Some examples of ordinals are

1. $0 = \emptyset$
2. $1 = 0^+ = \{\emptyset\} = \{0\}$
3. $2 = 1^+ = \{0, 1\}$
4. $3 = 2^+ = \{0, 1, 2\}$

Lemma 3.4.2:

If α is an ordering, so is $\alpha^+ = \alpha \cup \{\alpha\}$.

Proof. Since α is transitive, so is α^+ . The ordering \in on α^+ is the same as on α but with α added as an extra greatest element. Thus α^+ is a well ordering. ■

Proposition 3.4.3:

Let ω be the natural numbers as defined previously.

- (i) If $n \in \omega$ then n is an ordinal.
- (ii) ω is a transitive set.

Proof. \emptyset is an ordinal. Thus each element $n \in \omega$ is an ordinal by the previous lemma and induction.

We prove the second statement similarly, it is enough to show that if $m \in n \in \omega$ then $m \in \omega$ and proceed by induction. ■

Proposition 3.4.4:

Suppose that α is an ordinal.

- (i) $\alpha \notin \alpha$.
- (ii) If $\beta \in \alpha$, then β is an ordinal.
- (iii) If β is an ordinal and $\beta \subset \alpha$, then $\beta \in \alpha$.
- (iv) $\alpha = \{\beta : \beta \text{ is an ordinal and } \beta \in \alpha\}$.

Proof. Let α be an ordinal.

- (i) Clear.
- (ii) Check the definition of an ordinal.
- (iii) $\alpha \setminus \beta \subset \alpha$ is non-empty, thus it will have a least element γ . Remains to show that $\gamma = \beta$.
- (iv) Follows from part (ii). ■

Definition 3.4.5:

If α, β are ordinals, we write $\alpha < \beta$ to mean $\alpha \in \beta$. We write $\alpha \leq \beta$ to mean $\alpha \in \beta$ or $\alpha = \beta$. Note that $\alpha \leq \beta \iff \alpha \subseteq \beta$ by part (iii) of the above proposition.

Theorem 3.4.6:

Suppose that α, β, γ are ordinals.

- (i) If $\alpha < \beta$ and $\beta < \gamma$ then $\alpha < \gamma$.
- (ii) If $\alpha < \beta$ and $\beta < \alpha$ then $\alpha = \beta$.
- (iii) Exactly one of $\alpha < \beta$, $\alpha = \beta$ or $\beta < \alpha$ holds.
- (iv) If X is a non-empty set of ordinals, then X has a least element.

The last statement can be interpreted as saying that ‘the collection of ordinals is well ordered’.

Proof. Let α, β, γ be ordinals.

- (i) We have $\alpha \subset \beta$ and $\beta \subset \gamma$ so $\alpha \in \gamma \implies \alpha < \gamma$.
- (ii) If $\alpha \leq \beta$ and $\beta \leq \alpha$ then we have both $\alpha \subseteq \beta$ and $\beta \subseteq \alpha$ thus $\alpha = \beta$.
- (iii) As $\alpha \notin \alpha$, it follows that at most one of these holds.

- It remains to show that if $\alpha \neq \beta$ then either $\alpha \subset \beta$ or $\beta \subset \alpha$.
- We first use the definition of ordinals to check that $\alpha \subset \beta$ is an ordinal.
- If we have $\alpha \not\subset \beta$, then $\alpha \cap \beta \subset \alpha$.
- As $\alpha \cap \beta$ is an ordinal, part (iv) of the previous proposition gives us $\alpha \cap \beta \in \alpha$ with a similar statement holding for $\beta \not\subset \alpha$.
- Together this gives us $\alpha \cap \beta \in \alpha \cap \beta$ which is a contradiction.

(iv) See the notes for a complete proof. ■

Corollary 3.4.7:

Some minor facts.

- (i) Suppose that X is a set of ordinals, then $\bigcup X$ is an ordinal.
- (ii) ω is an ordinal.

Proof. We have that $\bigcup X$ is a set of ordinals, \in is therefore a well ordering by the above theorem. It is easy to check that $\bigcup X$ is transitive.

The fact that ω is an ordinal is a conclusion of the previous results. ■

We can now form more ordinals such as ω^+ and ω^{++} .

We want to classify the cardinalities of sets using ordinals.

Theorem 3.4.8:

If (A, \leq) is a well ordered set, there is a unique ordinal which is similar to (A, \leq) .

Definition 3.4.9:

Suppose that (A, \leq) is a wo set. We say that $X \subseteq A$ is an *initial segment* of A if whenever $y < x \in X$ then $y \in X$. It is *proper* if $X \neq A$.

Example:

If α, β are ordinals such that $\alpha < \beta$ then α is a proper initial segment of β since $\alpha = \{\delta \in \beta : \delta < \alpha\}$.

Lemma 3.4.10:

Suppose that (A, \leq) is a wo set. If $X \subset A$ is a proper initial segment of A , there is an $x \in A$ with $X = \{a \in A : a < x\} =: A[x]$.

Proof. Let $x \in \min(A \setminus X)$.

- If $y < x$ then $y \notin A \setminus X \implies y \in X$. Thus $A[x] \subseteq X$.
 - If $z \in X$ then $z \neq x$.
 - So if $z \notin A[x]$, $z > x$ this contradicts the fact that X is an initial segment.
-

Theorem 3.4.11:

Suppose that (A, \leq) is a wo set, $f : A \rightarrow A$ is an order preserving map and $f[A]$ is an initial segment of A . Then $\forall x \in A f(x) = x$.

Proof. Suppose not. Take $x = \min(y \in A : f(y) \neq y)$. So $f|_{A[x]}$ is the identity function on $A[x] \rightarrow A[x]$.

Since f is injective and $f(x) \neq x$ we must have $f(x) > x$. So $x \notin f[A]$ which is a contradiction since $f[A]$ is an initial segment of A . ■

Corollary 3.4.12:

If $\alpha \neq \beta$ are ordinals then $\alpha \not\prec \beta$.

Proof. Assume WLOG that $\beta < \alpha$, then β is an initial segment of α . Thus $\alpha \not\prec \beta$ by the above theorem. ■

Proof, Theorem 3.4.8. We've already shown that if (A, \leq) is similar to an ordinal, it is unique. It only remains to show that such an ordinal exists.

- Consider $X = \{x \in A : A[x] \text{ is similar to an ordinal}\}$.
- By uniqueness, if $x \in X$ there is a unique ordinal α_x similar to $A[x]$.
- Let

$$S = \{\alpha_x : x \in X\} \quad (*)$$

- S is a set of ordinals, we claim that S is also an ordinal.
 - We need to show that S is transitive.
 - Consider $\beta \in \alpha_x \in S$ and let $\pi : A[x] \rightarrow \alpha_x$ be a similarity.
 - Let $y = \pi^{-1}(\beta)$, we therefore have a similarity $\pi|_{A[y]} : A[y] \rightarrow \{\delta \in \alpha_x : \delta < \beta\} = \beta$.
 - Thus we have $\beta = \alpha_y \in S$
- Denote S by α .
- If $X = A$ we are done since $x \mapsto \alpha_x$ gives a similarity between A and α .
- By our construction of X , we know it is an initial segment of A .
- So if we suppose that $X \neq A$ we have $z \in A \setminus X$ such that $X = A[z]$.
- We know that $x \mapsto \alpha_x$ gives a similarity between X and α .
- This implies that $z \in X$ by the definition of X .
- This is a contradiction, so $X = A$ and $A \approx \alpha$. ■

Our proof relies on knowing that S as defined in $*$ is in fact a set. To show this we need the Axiom of Replacement.

Definition 3.4.13:

Suppose that $F(x, y, z_1, \dots, z_r) \in \text{Form}(L)$ such that whenever s_1, \dots, s_r are fixed sets and b is a set, there is a unique set a such that $F(a, b, s_1, \dots, s_r)$ holds.

We call the s_1, \dots, s_r *parameters*. F defines a “function” on sets $b \mapsto a$. We call F an *operation* on sets where the z_1, \dots, z_r are *parameter variables*.

ZF8: **Replacement:** Suppose that $F(x, y, z_1, \dots, z_r)$ is an operation on sets and that s_1, \dots, s_r are sets. In addition suppose that B is a set. Then there is a set A such that

$$A = \{a : F(a, b, s_1, \dots, s_r) \text{ holds for some } b \in B\}$$

In the case of $(*)$ we have $F(a, b) =$ 'either b is a wo set similar to the ordinal a , or b is not a wo set (or b is a wo set and not similar to an ordinal) and $a = \emptyset$ '. By applying F to $B = \{A[x] : x \in X\}$ we get S .

3.5 Transfinite induction

Theorem 3.5.1: Transfinite induction

Suppose $P(x)$ is a property of sets. Assume that for all ordinals α we have

$$(\forall \beta ((\beta < \alpha) \rightarrow P(\beta))) \rightarrow P(\alpha) \quad (**)$$

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

Remark:

If $\alpha = \emptyset$ then $P(\beta)$ holds for all $\beta < \alpha$ immediately so we have $P(\emptyset)$.

We can use transfinite induction to prove the following results.

Theorem 3.5.2:

Suppose that α is an infinite ordinal (so $\omega \leq \alpha$), then $\alpha \approx \alpha \times \alpha$.

Corollary 3.5.3:

Infinite well ordered sets are equinumerous with their Cartesian powers.

- (1) If (A, \leq) is an infinite wo set, we have $|A| = |A \times A|$.
- (2) Assuming the axiom of choice any set A can be well ordered and so if A is infinite, $|A| = |A \times A|$.

Proof, Theorem 3.5.1. Suppose for a contradiction that there is an ordinal γ such that $P(\gamma)$ does not hold. By [Theorem 3.4.6](#), part (iv) we have that the set $\{\delta \text{ an ordinal} : \delta \leq \gamma \text{ and } P(\delta) \text{ does not hold}\}$ has a least element. Let this be α .

Then for $\beta < \alpha$ we have that $P(\beta)$ holds. So by $(**)$ we have that $P(\alpha)$ holds which is a contradiction. ■

Proof, Theorem 3.5.2. Let α be an infinite ordinal ($\omega \leq \alpha$).

- (1) Assume that if $\omega \leq \beta \leq \alpha$ we have that $\beta \approx \beta \times \beta$. If we can conclude that $\alpha \approx \alpha \times \alpha$, we are done by transfinite induction.
- (2) We may assume that α is uncountable and if $\beta < \alpha$ then $|\beta| < |\alpha|$. In fact $|\beta^+| < |\alpha|$.
- (3) It is therefore enough to show that $|\alpha \times \alpha| \leq |\alpha|$ since we can then conclude by [Theorem 3.1.6](#).

step 1. Suppose we have w.o. \leq of $A = \alpha \times \alpha$ such that for all $x \in A$

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

We claim that $|\alpha \times \alpha| \leq |\alpha|$.

Proof. By [Theorem 3.4.8](#) there is an ordinal γ which is similar to (A, \leq) . Let $f : \gamma \rightarrow A$ be the similarity.

- It is enough to show $\gamma \subseteq \alpha$ since we will then have $|\gamma| \leq |\alpha| \implies |A| \leq |\alpha|$.
- Consider $\eta \in \gamma$.

- Since f is a similarity, it yields a bijection

$$\eta = \{\delta \in \gamma : \delta < \eta\} \rightarrow A[f(\eta)]$$

- So we have that $|\eta| = |A[f(\eta)]| < |\alpha|$.
- Thus $\eta < \alpha$, if it were not we would have $\alpha \leq \eta \implies \alpha \subseteq \eta \implies |\alpha| \leq |\eta|$.
- Thus $\eta \in \alpha \implies \gamma \subseteq \alpha$.

■

step 2. We must find an ordering on $A = \alpha \times \alpha$ as in the first step.

- For $\lambda < \alpha$ let $A_\lambda = \{(\theta, \zeta) \in \alpha \times \alpha : \max(\theta, \zeta) = \lambda\}$.
- Define \leq on A by

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

or

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

and either $\zeta' < \zeta$ or $\zeta' = \lambda = \zeta$ and $\theta' < \theta$, the reverse lexicographic order on A_λ .

- We claim that \leq is a well ordering on $A = \bigcup_{\lambda \in \alpha} A_\lambda$.

Exercise:
Show this.

- It remains to show that if $x = (\theta, \zeta) \in A$ then $|A[x]| < |\alpha|$.
- Let $\lambda = \max(\theta, \zeta)$; we may assume that $\lambda \geq \omega$.
- Let $\mu = \lambda^+$. We have $\mu < \alpha$ and by the induction hypothesis $|\mu \times \mu| = |\mu| < |\alpha|$.
- Then $\{y \in A : y < x\} \subseteq \{(\theta', \zeta') \in A : \max(\theta', \zeta') \leq \lambda\} = \mu \times \mu$.
- So $|A[x]| \leq |\mu \times \mu| < |\alpha|$ as required.

■

3.6 Transfinite recursion

Transfinite recursion allows us to construct sets $G(\alpha)$, for an ordinal α , such that $G(\alpha)$ is obtained from $G(\beta)$ where $\beta < \alpha$ by applying some operation F . We write $G \upharpoonright \alpha$ for $\{G(\beta) : \beta < \alpha\}$, a function $F(G \upharpoonright \alpha) = G(\alpha)$ can then be constructed. We note that $G \upharpoonright \alpha$ is a set because of the axiom of replacement.

Remark:

For F an operation on sets, we write $F(b)$ for the result of applying F to the set b .

Theorem 3.6.1: Transfinite recursion

Suppose that F is an operation on sets. Then there is an operation G such that for all ordinals α we have

$$G(\alpha) = G(G \upharpoonright \alpha)$$

If G' is another such operation, then $G(\alpha) = G'(\alpha)$ for all ordinals α .

Proof. Non-examinable, proof on blackboard.

■

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

Lemma 3.6.2: Lindenbaum

(c.f. [Lemma 1.31](#) and [Lemma 2.40](#)). Suppose that L is a first order language whose alphabet of symbols, P , is well ordered. Suppose that Σ is a consistent set of L -sentences.

Then there is a consistent set $\Sigma^* \supseteq \Sigma$ of L -sentences such that for every L -sentence ψ either $\psi \in \Sigma^*$ or $\neg\psi \in \Sigma^*$.

As P is well-ordered, we can well order $S = \bigcup_{n \in \mathbb{N}} P^n$. We can first order by length, and then reverse lexicographic ordering each P^n . This gives a well ordering on S . Hence we obtain a well ordering of the set of L -sentences.

Let λ be the unique ordinal similar to this well ordered set. Then we can write the set of L -sentences as $\{\phi_\alpha : \alpha < \lambda\}$, a set indexed by λ .

We construct Σ^* using transfinite recursion. Define, for each ordinal α a set $G(\alpha) \supseteq \Sigma$ of L -sentences.

$$G(\alpha) = \begin{cases} \Sigma \cup \bigcup_{\beta < \alpha} G(\beta) \cup \{\phi_\alpha\} & : \alpha < \lambda \text{ and } \Sigma \cup \bigcup_{\beta < \alpha} G(\beta) \vdash \phi_\alpha \\ \Sigma \cup \bigcup_{\beta < \alpha} G(\beta) \cup \{\neg\phi_\alpha\} & : \alpha < \lambda \text{ and } \Sigma \cup \bigcup_{\beta < \alpha} G(\beta) \not\vdash \phi_\alpha \\ \Sigma \cup \bigcup_{\beta < \alpha} G(\beta) & : \alpha < \lambda \end{cases}$$

Remark:

We note that for $\beta < \alpha$, $G(\beta) \subseteq G(\alpha)$. It can be shown (using transfinite induction and results from propositional and predicate logic) that $G(\alpha)$ is consistent.

By the construction, letting $\Sigma^* = G(\lambda)$ either $\phi_\alpha \in \Sigma^*$ or $\neg\phi_\alpha \in \Sigma^*$, $\forall \alpha < \lambda$.

Similar arguments in other parts of the proof of the Model Existence lemma ([2.32](#)) to get the completeness ([2.45](#)) and compactness theorems for L .

3.7 The axiom of regularity (foundation)

The final axiom of Zermelo-Fraenkel set theory is that of regularity.

ZF9: **Regularity** $\forall x ((x \neq \emptyset) \rightarrow (\exists a ((a \in x) \wedge (a \cap x = \emptyset))))$

In particular, there is no set b such that $b \in b$. Consider $x = \{b\}$, by ZF9 we have $b \cap \{b\} = \emptyset$ so $b \notin b$.

Thus we have encountered all the axioms of Zermelo-Fraenkel set theory (ZF). We reprint them here.

3.7.1 The Zermelo-Fraenkel axioms

ZF1: **Extensionality** $\forall x \forall y ((x = y) \leftrightarrow \forall z ((z \in x) \leftrightarrow (z \in y)))$

ZF2: **Empty set** $\exists x \forall y (y \notin x)$.

ZF3: **Pairing** $\forall x \forall y \exists z (\forall w (w \in z \leftrightarrow ((w \in x) \vee (w \in y))))$

ZF4: **Union** For any set A there is a set $B = \bigcup A = \bigcup \{z : z \in A\}$.

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

ZF5: **Power set** For any set A , there is a set whose elements are the subsets of A , the *power set*. We write $z \subseteq A \iff \forall y ((y \in z) \rightarrow (y \in A))$.

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

ZF6: **Axiom scheme of specification** Suppose that $P(x, y_1, \dots, y_r) \in \text{Form}(L)$, then

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

ZF7: **Infinity** $\exists A (\emptyset \in A) \wedge (\forall x ((x \in A) \rightarrow (x^+ \in A)))$

ZF8: **Replacement**: Suppose that $F(x, y, z_1, \dots, z_r)$ is an operation on sets and that s_1, \dots, s_r are sets. In addition suppose that B is a set. Then there is a set A such that

$$A = \{a : F(a, b, s_1, \dots, s_r) \text{ holds for some } b \in B\}$$

ZF9: **Regularity** $\forall x ((x \neq \emptyset) \rightarrow (\exists a ((a \in x) \wedge (a \cap x = \emptyset))))$

4 The axiom of choice and consequences

4.1 Statement of the axiom of choice and the well ordering principle

Definition 4.1.1: Axiom of choice

Suppose A is a set of non-empty sets. Then there is a function $f : A \rightarrow \bigcup A$ with $f(a) \in a$ for all $a \in A$.

We call the ZF axioms and the axiom of choice ZFC.

Example 4.2:

Suppose that X is any non-empty set, let $A = \mathcal{P}(X) \setminus \{\emptyset\}$, the non-empty subsets of X . By the axiom of choice, there is a function $f : A \rightarrow X$ such that $f(Y) \in Y$ for every $\emptyset \neq Y \subseteq X$.

Such a function is called a choice function on X .

Remark:

If (X, \leq) is a well ordered set, then we don't need the axiom of choice to get a choice function. We can let $f(Y) = \min(Y)$ for $\emptyset \neq Y \subseteq X$.

Theorem 4.1.3: Well ordering principle (WO)

Suppose that X is a non-empty set and $f : \mathcal{P}(X) \setminus \{\emptyset\} \rightarrow X$ is a choice function. Then there is a well-ordering \leq of X , i.e. a well ordered set (X, \leq) .

Lemma 4.1.4: Hartogs'

For any set A there is an ordinal α such that there is no injective function $h : \alpha \rightarrow A$.

Proof, Theorem 4.1.3. We define $G(\alpha) = f(\underbrace{X \setminus \{G(\beta) : \beta < \alpha\}}_{\text{when } \neq \emptyset})$.

- Let α be the ordinal as yielded by Lemma 4.1.4.
- Consider $\tilde{X} = X \cup \{\infty\}$ where ∞ is some set with $\infty \in X$.
- Using transfinite recursion, for an ordinal γ , we can define an operation G

$$G(\gamma) = \begin{cases} f(X \setminus \{G(\beta) : \beta < \gamma\}) & : X \setminus \{G(\beta) : \beta < \gamma\} \neq \emptyset \\ \infty & : \text{else} \end{cases}$$

Remark:

Note that if $\infty \notin \text{im}(G \upharpoonright \gamma)$ then $G \upharpoonright \gamma$ is an injective function $\gamma \rightarrow X$.

- By Hartogs' lemma, there is some ordinal α with $G(\alpha) = \infty$
- Take the least such α .
- Thus $g = G \upharpoonright \alpha : \alpha \rightarrow X$ is an injective function and as $G(\alpha) = \infty$, the image of g is X .
- Thus g is a bijection.
- We define \leq on X by $x_1 \leq x_2 \iff g^{-1}(x_1) \leq g^{-1}(x_2)$ where the ordering on the right is given by \in on α .

■

Proof, Lemma 4.1.4, Hartogs' lemma. Consider $X = \{(Y, \leq_Y) : Y \subseteq A \text{ and } \leq_Y \text{ is a well ordering on } Y\}$. By specification, this is a set.

- Let $S = \{\beta : \beta \text{ is an ordinal similar to some } (Y, \leq_Y) \in X\}$.
- Using Theorem 3.4.8 and replacement, we find that S is a set.
- Note that $S = \{\beta : \beta \text{ is an ordinal and there is an injective function } \beta \rightarrow A\}$.
- Let $\sigma = \bigcup S$.
- This is an ordinal and $\beta \leq \sigma$ for all $\beta \in S$ by Proposition 3.4.4 and Corollary 3.4.7.
- Let $\alpha = \sigma^+$. So σ is an ordinal and for $\beta \in S, \beta \leq \sigma < \alpha$.
- Thus $\alpha \in S$.

■

Corollary 4.1.5:

Assuming the ZF axioms, then the axiom of choice is equivalent to the well ordering principle. Moreover, if A is any set, then there is $\leq_A \subseteq A \times A$ such that (A, \leq_A) is a well ordered set. i.e.

$$\text{ZF} \vdash (\text{AC} \leftrightarrow \text{WO})$$

Proof. We have $\text{AC} \implies \text{WO}$ by Theorem 4.1.3. We now show the converse.

- If A is any set of non-empty sets, let $B = \bigcup A$.
- By WO, there is a well ordering \leq_B on B .
- We define $f : A \rightarrow \bigcup A$ by $f(a) =$ the least element of a with respect to \leq_B .

■

Corollary 4.1.6:

Assuming ZFC,

- (i) *If A is any set, there is an ordinal $\alpha \approx A$.*
- (ii) *If A, B are any sets, then either $|A| \leq |B|$ or $|B| \leq |A|$.*
- (iii) *(The fundamental theorem of cardinal arithmetic) If A is any infinite set, then $|A| = |A \times A|$.*

Proof. Parts (i) and (iii) are proven by Corollary 4.1.5 and Corollary 3.5.3. We now show (ii)

- There exist ordinals α, β with $A \approx \alpha$ and $B \approx \beta$ by part (i).
- By Theorem 3.4.6, we have either $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$.

■

Lemma 4.1.7:

Assuming ZFC, suppose that A and B are sets with $A \neq \emptyset$. Then $|A| \leq |B| \iff$ there is a surjective function $h : B \rightarrow A$.

Exercise:

We leave the proof as an exercise.

4.2 Cardinals and cardinality

Definition 4.2.1:

An ordinal is a *cardinal* if it is not equinumerous with any $\beta < \alpha$.

Example:

If $n \in \omega$, n is a cardinal. Moreover ω is a cardinal. Though if β is an infinite ordinal then $\beta \approx \beta^+$ so β^+ is not a cardinal.

For the rest of this section, we assume ZFC.

Lemma 4.2.2:

Suppose A is any set. Then there is a unique cardinal $\alpha \approx A$.

Proof. By [Corollary 4.1.6](#) there is some ordinal $\gamma \approx A$. Take α to be the least such ordinal. Then α is a cardinal. ■

Definition 4.2.3:

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

Exercise:

There is an injective function $A \rightarrow B$ iff $\text{card}(A) \leq \text{card}(B)$.

This justifies the use of the notation $\text{card}(A) = |A|$.

Example:

Some examples of cardinalities.

1. If A is a countably infinite set, $|A| = \omega$.
2. If α is an ordinal, then $|\alpha| = \alpha \iff \alpha$ is a cardinal.

Definition 4.2.4:

Suppose A and B are disjoint sets with $|A| = \kappa$ and $|B| = \lambda$ (so κ, λ are cardinals). Define

- $\kappa + \lambda := |A \cup B|$
- $\kappa \cdot \lambda := |A \times B|$

Remark:

Note that this doesn't depend on the choice of A and B .

Theorem 4.2.5:

Suppose that κ, λ are cardinals, that $\kappa \leq \lambda$ and λ is infinite. Then:

- (i) $\kappa + \lambda = \lambda$ and
- (ii) $\kappa \cdot \lambda = \lambda$ if $\kappa \neq 0$.

Proof. To show (ii), note that as $\kappa \leq \lambda$, we have $\kappa \subseteq \lambda$. This implies $\kappa \times \lambda \subseteq \lambda \times \lambda$. Thus

$$\begin{aligned}\kappa \cdot \lambda &= |\kappa \times \lambda| \\ &\leq |\lambda \times \lambda| \\ &= |\lambda|(\text{FTCA}) \\ &= \lambda(\text{cardinal})\end{aligned}$$

Since $\kappa \neq 0$, there is an injective function $\lambda \rightarrow \kappa \times \lambda, \alpha \mapsto (0, \alpha)$. So $|\lambda| \leq |\kappa \times \lambda|$, i.e. $\lambda \leq \kappa \cdot \lambda$.

To show (i),

$$\begin{aligned}\lambda &\leq \kappa + \lambda \\ &\leq \lambda + \lambda \\ &\approx \{0, 1\} \times \lambda \\ &= 2 \cdot \lambda \\ &= \lambda(\text{by (ii)})\end{aligned}$$

■

Theorem 4.2.6:

Suppose A is an infinite set of cardinality λ . Suppose that each element of A is a set of cardinality $\leq \kappa$. Then $|\bigcup A| \leq \lambda \cdot \kappa$.

Example:

Suppose that P is an infinite set, $|P| = \lambda \geq \omega$. Let S be the set of finite sequences of elements of P . So $S = \bigcup_{n \in \omega} P^n$ and then $|S| = \lambda$.

Proof. For $n \in \omega, n \neq 0$; $|P^n| = \underbrace{\lambda \cdot \dots \cdot \lambda}_n = \lambda$ by (ii) in Theorem 4.2.5. By Theorem 4.2.6 we have $|A| \leq \omega \cdot \lambda = \lambda$. But $P \subseteq S$ so $|S| \geq \lambda$.

■

Proof, ??. We can assume that $\phi \in A$.

- For each $a \in A$, the set S_a of surjective functions $\kappa \rightarrow a$ is non-empty by Lemma 4.1.7.
- Assuming AC, there is a function $F : A \rightarrow \bigcup_{a \in A} S_a$ with $F(a) \in S_a$ for all $a \in A$.
- i.e. $F(a) : \kappa \rightarrow a$ is a surjective function.
- Let $h : \lambda \rightarrow A$ be a bijection.
- Define $g : \lambda \times \kappa \rightarrow \bigcup A$ by $g(\alpha, \beta) = (F(h(\alpha)))(\beta)$. We note that this is surjective.
- So by Lemma 4.1.7 $|\lambda \times \kappa| \geq |\bigcup A|$
- i.e. $|\bigcup A| \leq \lambda \cdot \kappa$

■

4.3 Zorn's lemma

A partially ordered set (poset) (A, \leq) satisfies:

$$\begin{aligned}\forall xyz \in A \quad (x \leq y \leq z \rightarrow x \leq z) \\ \forall xyz \in A \quad ((x \leq y) \wedge (y \leq x) \rightarrow (x \dot{=} y)) \\ \forall xyz \in A \quad (x \leq x)\end{aligned}$$

Example:

Consider $(\mathcal{P}(X), \subseteq)$. A chain C is a poset (A, \leq) is a subset $C \subseteq A$ such that $\forall x, y \in C ((x \leq y) \vee (y \leq x))$. An upper bound of C in A is $a \in A$ such that $a \geq c$ for all $c \in C$.

|: If $C \subseteq \mathcal{P}(X)$ is a chain, then $\bigcup C$ is an upper bound for C in $\mathcal{P}(X)$.

Definition 4.3.1: Zorn's Lemma

Suppose (A, \leq) is a non-empty poset in which every chain in A has an upper bound in A . Then A has a maximal element

$$\exists z \in A \forall x \in A ((x \geq z) \rightarrow (x = z))$$

Theorem 4.3.2:

$\text{ZF} \vdash (\text{AC} \leftrightarrow \text{ZL})$, i.e.

- (1) Assuming ZFC, then ZL holds.
- (2) Assuming ZF and ZL, then AC follows.

Example 4.3:

Assume ZFC. Suppose V is a vector space over a field F . Then V has a basis.

Proof. Use ZL. Let A be the set of linearly independent subsets of V , ordered by \subseteq . We claim that if C is a chain in A then $\bigcup C \in A$. So then the hypothesis of ZL holds for (A, \subseteq) . It remains to show that $\bigcup C$ is linearly independent, i.e. if $y_1, \dots, y_n \in C$ then y_1, \dots, y_n are linearly independent.

There are $C_1, \dots, C_n \in C$ such that $y_i \in C_i$. As C is a chain, there is $j \leq n$ with $C_i \subseteq C_j$ for all $i \leq n$.

But then $y_1, \dots, y_n \in C_j$ so y_1, \dots, y_n are linearly independent.

So by ZL, there is a maximal element B of A . As $B \in A$, B is a linearly independent set. If $v \in V \setminus B$ then $B \cup \{v\}$ is not linearly independent (as B is maximally linearly independent). This implies v is a linear combination of vectors in B and thus B spans V . i.e. B is a basis. ■

Proof, $\text{AC} \implies \text{ZL}$. Let $f : \mathcal{P}(A) \setminus \{\emptyset\} \rightarrow A$ be a choice function. Suppose for a contradiction that A has no maximal element.

- Let $C \subseteq A$ be a chain in A .
- By assumption, there is some $y \in A$ with $c \leq y$ for all $c \in C$.
- As y is not maximal, there is $z \in A$ with $z > y$.
- Therefore $z > c$ for all $c \in C$.
- We can use transfinite recursion to define an operation G such that for all ordinals α , $G(\alpha) \in A$ and $G(\alpha) = f(\{z \in A : z > G(\beta) \text{ for all } \beta < \alpha\})$.
- So $G(0) < G(1) < \dots < G(\beta) < \dots$.
- i.e. for all ordinals $\beta < \alpha$, we have $G(\beta) < G(\alpha)$.
- So for every ordinal, $G \upharpoonright \alpha$ is an injective function to A .
- This contradicts Hartogs' lemma 4.1.4. ■

Postscript

If we assume ZFC, then Gödel's completeness theorem and the compactness theorem hold for arbitrary first order languages and not just countable languages.