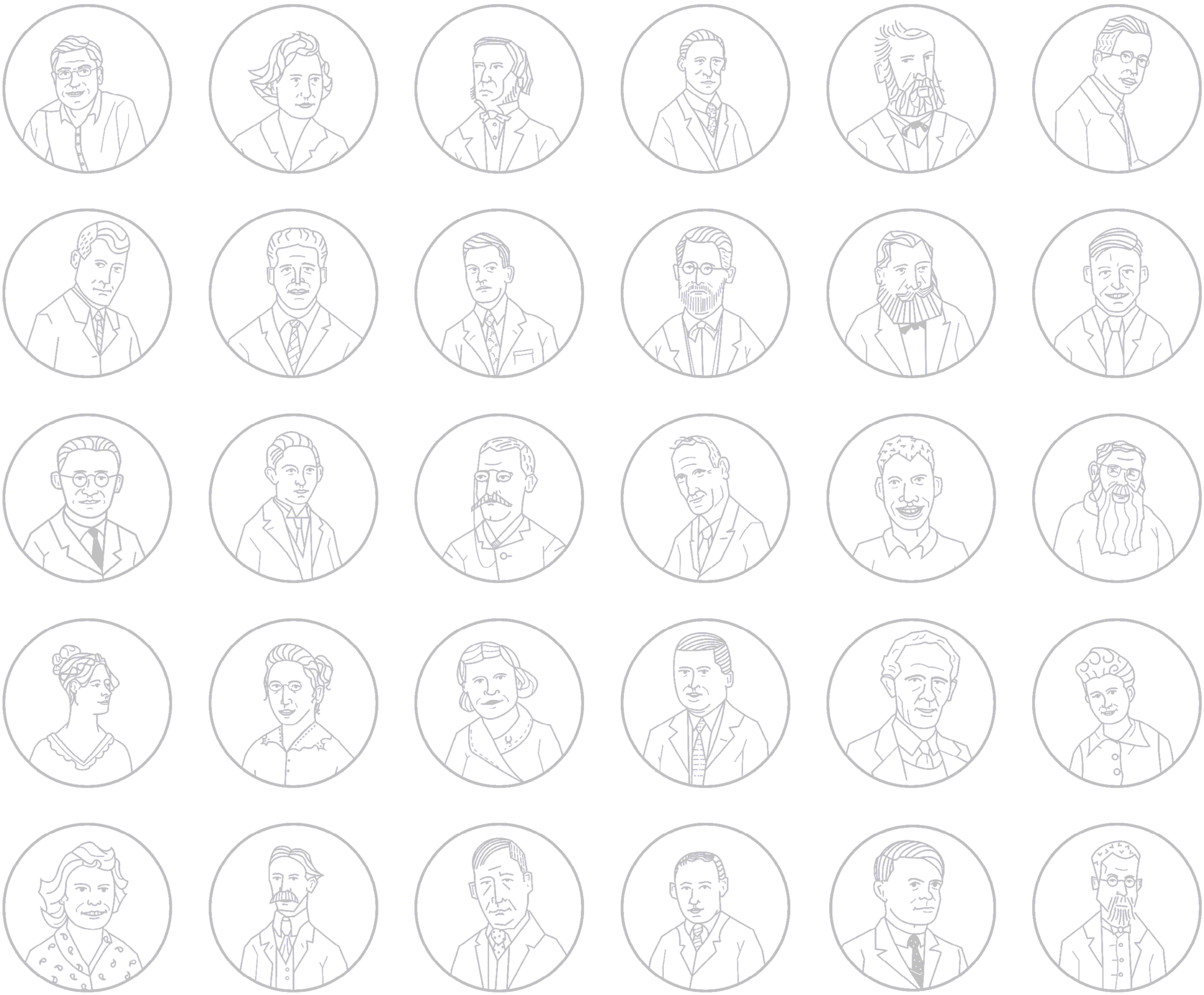
Logic and Its Limits

Lavinia Picollo



Artist: Matthew Leadbeater

Contents

Chapter 1

Syntax for First-Order Languages

The well-formed formulae (wff) of a first-order language are strings of symbols borrowing from the first-order alphabet.

Definition 1.1 (first-order alphabet)**.** The first-order alphabet is a set consisting of the following basic symbols:

- Logical terms: $\neg, \rightarrow, \forall, =$
- Individual variables: $x_0, x_1, x_2, \ldots$
- Individual constants: $f_0^0, f_1^0, f_2^0, \ldots$
- Function symbols: $f_0^1, f_1^1, \ldots, f_0^2, f_1^2, \ldots, f_0^n, f_1^n, \ldots$
- Predicate/relation symbols: $P_0^1, P_1^1, \ldots, P_0^2, P_1^2, \ldots, P_0^n, P_1^n, \ldots$
- Auxiliary symbols: (,), ,

The superscripts in function and predicate symbols indicate their arity. Thus, f_m^n and P_m^n have arity n; they are an n-place function symbol and an n-place predicate symbol, respectively. Sometimes we omit subscripts and superscripts for readability, and also write $x, y, z, \ldots$ for $x_0, x_1, x_2, \ldots$.

A first-order language $\mathscr{L}$ is given by an (effectively decidable) subset of individual constants, function symbols, and predicate/relation symbols of the first-order alphabet. We note it '$V(\mathscr{L})$'. $V(\mathscr{L})$ is the set of non-logical vocabulary of $\mathscr{L}$.

The universal first-order language, $\mathscr{L}_U$, is such that (s.t.) $V(\mathscr{L}_U)$ is the set of all individual constants, all function symbols, and all predicate symbols of the first-order alphabet.

In what follows, let $\mathscr{L}$ be a first-order language.

Definition 1.2 (term of $\mathscr{L}$)**.**

- Individual variables are terms of $\mathscr{L}$.
- Individual constants in $V(\mathscr{L})$ are terms of $\mathscr{L}$.
- If f is an n-place function symbol in $V(\mathscr{L})$ and $t_1, \ldots, t_n$ are terms of $\mathscr{L}$, $f(t_1, \ldots, t_n)$ is also a term of $\mathscr{L}$.
- Nothing else is a term of $\mathscr{L}$.

Note that function symbols in $V(\mathscr{L})$ by themselves don't constitute terms of $\mathscr{L}$.

Definition 1.3 (well-formed formula of $\mathscr{L}$)**.**

- If t_1 and t_2 are terms of $\mathscr{L}$, then $t_1 = t_2$ is a wff of $\mathscr{L}$.
- If P is an n-place predicate symbol in $V(\mathscr{L})$ and $t_1, \ldots, t_n$ are terms of $\mathscr{L}$, $P(t_1, \ldots, t_n)$ is a wff of $\mathscr{L}$.
- If φ is a wff of $\mathscr{L}$, $\neg\varphi$ is also a wff of $\mathscr{L}$.
- If φ and ψ are wff of $\mathscr{L}$, then $(\varphi \to \psi)$ is also a wff of $\mathscr{L}$.
- If φ is a wff of $\mathscr{L}$ and v is a variable, $\forall v\, \varphi$ is also a wff of $\mathscr{L}$.
- Nothing else is a wff of $\mathscr{L}$.

In what follows, let φ, ψ, χ range over wff of $\mathscr{L}$ in the metalanguage, Γ, Δ, Σ over sets of wff of $\mathscr{L}$, c over individual constants of $\mathscr{L}$, s, t over terms of $\mathscr{L}$ and u, v, over individual variables, possibly with subscripts.

Definition 1.4 (free and bound variables)**.**

- The scope of an occurrence of a quantifier in a wff is the wff that occurs immediately after the occurrence of the quantifier.
- An occurrence of a variable v in a wff is bound if and only if (iff) it occurs in the scope of an occurrence of the quantifier $\forall v$, and free otherwise.
- φ is a sentence iff no variable occurs free in φ.
- t is a closed term iff it contains no variables; otherwise it's open.
- t is free for v in φ iff v doesn't occur in φ in the scope of a quantifier whose variable occurs in t.

By $\varphi[t/v]$ we denote the wff that results from replacing all free occurrences of v with t in φ if t is free for v in φ; otherwise, $\varphi[t/v]$ is just φ.

For convenience, we may use the following abbreviations:

- $(\varphi \wedge \psi)$ for $\neg(\varphi \to \neg\psi)$
- $(\varphi \vee \psi)$ for $(\neg\varphi \to \psi)$

- $(\varphi \leftrightarrow \psi)$ for $\neg((\varphi \to \psi) \to \neg(\psi \to \varphi))$
- $\exists v\, \varphi$ for $\neg \forall v\, \neg \varphi$
- $t_1 \neq t_2$ for $\neg t_1 = t_2$

The usual bracketing conventions apply.

Example 1.1 (the language of first-order Peano arithmetic)**.** $\mathscr{L}_{\mathsf{PA}}$ is the first-order language s.t. $V(\mathscr{L}_{\mathsf{PA}}) = \{f_0^0, f_0^1, f_0^2, f_1^2\}$. For convenience, we write 0 for f_0^0, S for f_0^1, + for f_0^2, and $\times$ for f_1^2. We also write $(t_1 + t_2)$ and $(t_1 \times t_2)$ instead of $+(t_1, t_2)$ and $\times(t_1, t_2)$ for readability, and we omit outer brackets when possible.
Some terms of $\mathscr{L}_{\mathsf{PA}}$ are 0, S0, SSx, $(y + \mathrm{S}0) \times \mathrm{SS}z$.
Some wff of $\mathscr{L}_{\mathsf{PA}}$ are $0 = x \times 0$, $\forall y\, (y \neq y + \mathrm{SS}0 \to y \neq 0)$, $\neg \exists x\, 0 \neq z$.

Exercise 1.1. *Decide which of the following are terms of $\mathscr{L}_{\mathsf{PA}}$ (notational conventions may apply). Indicate which are open and which are closed terms.*

1. SSSSS0
2. S000
3. x
4. $1 + 0$
5. $y \times x$
6. $+$
7. $0 \times 0 = 0$
8. $\mathrm{SS}0 \times (y + (\mathrm{S}0 \times x))$

Appendix: Induction

Induction for natural numbers

Recall the natural numbers are $0, 1, 2, \ldots$ For each predicate P,[1] the following is true:

Definition 1.5 (standard induction)**.** If 0 is P and, for every natural number, if it being a P entails that its successor is also P, then all natural numbers are P.

This principle is schematic on P, that is, it consist of all the instances that result from replacing the schematic letter P with a predicate. It isn't logically true, albeit true of the natural numbers, due to their structure: each natural number can be obtained by finitely many applications of the successor operation to 0.

Breaking standard induction down into steps into argument form:

- *Base step:* 0 is P.

[1] Strictly speaking, only well-behaved predicates can substitute P in an induction principle. Non-well-behaved predicates include vague predicates, such as 'if the number of hairs on Bob's head is . . . , then Bob is bald'.

- *Inductive step:* For every natural number x, if x is P, then $x+1$ is P.
- *Conclusion:* Every natural number is P.

Of course this can be adapted to the positive integers: we just start with 1 instead of 0. Similarly, we can adapt it to every subset of natural numbers equal to or greater than a given n, i.e. $\{n, n+1, n+2, \dots\}$.

Example 1.2. For every natural number x, $3^{x+1}-1$ is even.
We prove it by standard induction, replacing P with 'is s.t. $3^{x+1}-1$ is even'.
Base step: If $x=0$, $3^{x+1}-1 = 3^1-1 = 2$, which is an even number.
Inductive step: Since we have to prove a conditional, we assume the antecedent – i.e. the inductive hypothesis – and try to arrive at the consequent. Thus, assume $3^{x+1}-1$ is even – this is the inductive hypothesis. Now we need to show that $3^{(x+1)+1}-1$ is even too. Note that $3^{(x+1)+1}-1 = 3\times 3^{x+1}-1 = (2+1)\times 3^{x+1}-1 = 2\times 3^{x+1}+3^{x+1}-1$. This number is even, since $2\times 3^{x+1}$ is obviously even and, by inductive hypothesis, so is $3^{x+1}-1$ (and the addition of two even numbers is always even).
Conclusion: Every natural number x is s.t. $3^{x+1}-1$ is even.

Exercise 1.2. *Show that, for every natural number x, $0+1+\dots+x = \frac{x\times(x+1)}{2}$.*

Exercise 1.3. *Give an induction principle for the positive integers.*

Standard induction allows one to conclude that all natural numbers are P, provided that one is able to show that 0 is a P and that if a number is a P then its successor is so too, namely, the base and the inductive step. Note that if it's not true that all natural numbers are P, then it won't be possible to prove both the base and the inductive step (because induction is a *true* principle about the natural numbers).

To prove the inductive step, one needs to prove a general claim about all natural numbers. But this is usually easier than proving the conclusion of induction itself, because the inductive step has conditional form. This means we can assume the antecedent and then try to derive the consequent from this assumption. In other words, one can assume that an (arbitrary) number x is P – the inductive hypothesis – and, from this assumption, try to prove that $x+1$, the successor of x, is also P. Since x was arbitrary, one can conclude that, for all x, if x is P then $x+1$ is P too.

The following principle is also true for each (well-behaved) predicate P:

Definition 1.6 (total induction)**.** If 0 is P and, for every natural number x, if all natural numbers smaller than x being P entails that x is P too, then all natural numbers are P.

Total induction for natural numbers is also a schematic principle and it's *equivalent* to standard induction. When to use one or the other depends on the nature of the statement to be proved. As standard induction, it can be adapted to prove universal claims about the positive integers and, more generally, about every final segment of the natural numbers.

Breaking total induction down into steps into argument form:

- *Base step:* 0 is P.
- *Inductive step:* For every natural number x, if all natural numbers smaller than x are P, then x is P.
- *Conclusion:* Every natural number is P.

Exercise 1.4. *Give a total induction principle for a set of natural numbers equal to or greater than a given number n.*

Induction for well-formed formulae

If every object of a given domain has a unique natural number (or positive integer) associated with it, one can appeal to this association to prove universal claims about the given domain by proving universal claims about numbers.

Let the logical complexity of a wff φ of $\mathscr{L}$ be the number of occurrences of logical connectives – i.e. $\neg, \rightarrow, \forall$ – in φ. Atomic formulae have logical complexity 0, their negations have logical complexity 1, conditionals between the latter have logical complexity 3, etc. Since wff are always finite, their logical complexity is always a natural number. So we can use the following instance of total induction for natural numbers to prove universal claims about wff of $\mathscr{L}$:

Definition 1.7 (total induction for wff by proxy)**.**

- *Base step:* All wff of $\mathscr{L}$ of logical complexity 0 are P.
- *Inductive step:* For every natural number x, if all wff of $\mathscr{L}$ of logical complexity lower than x are P, then all wff of $\mathscr{L}$ of logical complexity x are P.
- *Conclusion:* For every natural number x, all wff of $\mathscr{L}$ of logical complexity x are P. (In other words, every wff of $\mathscr{L}$ is P.)

Note that the base step doesn't require you to prove that only one object has a property, but that all atomic wff of $\mathscr{L}$ have the property. It's a bit more demanding.

For the inductive step, the easiest way is to proceed by *cases*, i.e. to consider all possible forms a wff φ of logical complexity x might take (atomic wff have been dealt with already in the base step):

1. φ is a negation, that is, φ is $\neg\psi$, for some wff ψ. Note that, since φ is of logical complexity x, the logical complexity of ψ is lower than x ($x - 1$, to be precise). By inductive hypothesis, ψ is P, and usually this turns out to be useful to prove that φ is also P.

2. φ is a conditional, that is, φ is $(\psi \to \chi)$, for some wff ψ and χ. Since φ is of logical complexity x, the logical complexity of ψ and χ is lower than x. By inductive hypothesis, ψ and χ are both P, and usually this turns out to be useful to prove that φ also is P.

3. φ is a universal claim, that is, φ is $\forall v\, \psi$, for some wff ψ and some variable v. Again, since φ is of logical complexity x, the logical complexity of ψ is lower than x. By inductive hypothesis, ψ is P, and usually this turns out to be useful to prove that φ also is P.

Example 1.3. Every formula of $\mathscr{L}$ has more atomic subformulae than occurrences of $\to$.
We prove it by total induction for wff by proxy, replacing P with 'has more atomic subformulae than occurrences of $\to$'.
Base step: The wff of $\mathscr{L}$ of logical complexity 0, the atomic wff, have no occurrences of $\to$ and exactly one atomic subformula each, namely, themselves. Thus, they have more atomic subformulae than occurrences of $\to$.
Inductive step: Assume all wff of lower logical complexity than φ's have more atomic subformulae than occurrences of $\to$ (this is the inductive hypothesis). (Alternatively, we can write: "Assume all wff of logical complexity less that x have more atomic subformulae than occurrences of $\to$, and let φ have logical complexity x.") We reason by cases:

- φ is $\neg\psi$, for some ψ. Since the logical complexity of ψ is lower than φ's/x, by inductive hypothesis, ψ has more atomic subformulae than occurrences of $\to$. Since φ has the same atomic subformulae and the same number of occurrences of $\to$ as ψ, φ also has more atomic subformulae than occurrences of $\to$.

- φ is $(\psi \to \chi)$, for some ψ and χ. Since the logical complexity of ψ and χ is lower than φ's/x, by inductive hypothesis, ψ and χ both have more atomic subformulae than occurrences of $\to$. But, then, ψ and χ together have at least two more atomic subformulae than occurrences of $\to$, which means that $(\psi \to \chi)$ – i.e. φ – has at least one more atomic subformulae than occurrences of $\to$.

- The case in which φ is a universal claim can be proved just like the first case.

Conclusion: For every natural number x, all wff of logical complexity x have more atomic subformulae than occurrences of $\to$.

The natural numbers or the positive integers are not the only domains that support induction. Other domains have a similar structure and have analogous inductive principles. This is the case, for instance, of the integers. Just like all natural numbers can be obtained by finitely many applications of the successor operation to 0, the integers can be all obtained by finitely many applications of one of the following two operations to 0: successor and predecessor.

Another case is that of the wff of a first-order language. Every wff of a first-order language can be obtained by adding finitely many negations, conditionals, and universal quantifiers to atomic wff. This structural feature of the wff of any first-order language $\mathscr{L}$ entails the following induction principle:

Definition 1.8 (tailored induction for wff)**.**

- *Base step:* All atomic wff of $\mathscr{L}$ are P.
- *Inductive step 1:* For every wff φ of $\mathscr{L}$, if φ is P, then $\neg\varphi$ is also P.
- *Inductive step 2:* For every wff φ and ψ of $\mathscr{L}$, if φ and ψ are P, then $(\varphi \to \psi)$ is also P.
- *Inductive step 3:* For every wff φ of $\mathscr{L}$, if φ is P, then $\forall v\, \varphi$ is also P.
- *Conclusion:* Every wff of $\mathscr{L}$ is P.

Exercise 1.5. *Every formula of $\mathscr{L}$ contains at least as many terms as predicate symbols.*

Exercise 1.6. *Give an induction principle for the integers.*

Exercise 1.7. *Give an induction principle for the terms of a given first-order language $\mathscr{L}$.*

Exercise 1.8. *Show that every term of $\mathscr{L}$ contains as many left as right brackets.*

Exercise 1.9. *Show that every wff of $\mathscr{L}$ contains as many left as right brackets by tailored induction for wff.*

Chapter 2

Proof Theory for First-Order Logic

A calculus for a logic consists of a set of principles together with an indication how they are to be combined, that is, a definition of *derivation* (or *proof*). Each derivation is a proof of a valid argument in the logic or of one of its logical truths.

In this chapter we will introduce a Hilbert-style calculus for first-order logic with identity. Hilbert-style calculi consist mostly of *axioms* and very few *rules of inference*. Derivations are lists of wff in which axioms and rules feature prominently. Then, we will show how to 'extract' a Fitch-style natural deduction calculus from our Hilbert-style calculus. Fitch-style natural deduction calculi consist of rules of inference, some of which require introducing and discharging assumptions, and possibly a few axioms. We will show that these rules and axioms are all derivable – or, as we usually say, *admissible* – in our Hilbert-style calculus.

2.1 A calculus for first-order logic with identity

Definition 2.1 (logical axioms)**.** The logical axioms of the calculus for $\mathscr{L}$ consist of all the instances of the following schemata given by expressions of $\mathscr{L}$:

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

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

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

A4 $\forall v\, (\varphi \to \psi) \to (\forall v\, \varphi \to \forall v\, \psi)$

A5 $\forall v\, \varphi \to \varphi[t/v]$

A6 $\varphi \to \forall v\, \varphi$, if v is not free in φ

A7 $\forall v\, v = v$

A8 $\forall u \forall v\, (u = v \to (\varphi[u/v] \to \varphi))$

Definition 2.2 (derivation)**.** A derivation of φ from Γ in the calculus is a finite sequence of wff of $\mathscr{L}$ whose last member is φ s.t. each of them is either a logical axiom, a member of Γ, or follows by one of the following two rules from previous entries in the sequence:

Modus Ponens From φ and $\varphi \to \psi$, infer ψ.[1]

Generalisation From φ infer $\forall v\, \varphi$, if v is not free in any member of Γ occurring in the sequence.

Example 2.1. The following is a derivation of $\varphi \to \chi$ from $\{\varphi \to \psi, \psi \to \chi\}$

1.	$\varphi \to \psi$	
2.	$\psi \to \chi$	
3.	$(\psi \to \chi) \to (\varphi \to (\psi \to \chi))$	A1
4.	$\varphi \to (\psi \to \chi)$	MP 2,3
5.	$(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$	A2
6.	$(\varphi \to \psi) \to (\varphi \to \chi)$	MP 4,5
7.	$\varphi \to \chi$	MP 1,6

Example 2.2. The following is a derivation of $\varphi \to \varphi$ from $\varnothing$:

1.	$\varphi \to ((\varphi \to \varphi) \to \varphi)$	A1
2.	$(\varphi \to ((\varphi \to \varphi) \to \varphi)) \to ((\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi))$	A2
3.	$(\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi)$	MP 1,2
4.	$\varphi \to (\varphi \to \varphi)$	A1
5.	$\varphi \to \varphi$	MP 3,4

Definition 2.3 (proof-theoretic consequence)**.** $\Gamma \vdash \varphi$ iff there is a derivation of φ from Γ.

Example 2.3 (hypothetical syllogism)**.** $\varphi \to \psi, \psi \to \chi \vdash \varphi \to \chi$
See Example 2.1.

Definition 2.4 (logical theoremhood)**.** $\vdash \varphi$ iff $\varnothing \vdash \varphi$.

Example 2.4. $\vdash \varphi \to \varphi$
See Example 2.2.

We often write $\Gamma, \varphi \vdash \psi$ for $\Gamma \cup \{\varphi\} \vdash \psi$, for simplicity.

Theorem 2.1 (structural properties)**.**

[1] Also known as 'elimination of the conditional', or $\to$E.

- *Reflexivity.* $\Gamma, \varphi \vdash \varphi$.
- *Transitivity.* *If* $\Gamma \vdash \varphi$ *and* $\Delta, \varphi \vdash \psi$, *then* $\Gamma \cup \Delta \vdash \psi$.
- *Monotonicity.* *If* $\Gamma \vdash \varphi$, *then* $\Gamma \cup \Delta \vdash \varphi$.

Proof. All properties follow from the definition proof-theoretic consequence and of derivation.

For reflexivity, note that the sequence consisting only of φ is a derivation of φ from Γ, φ.

For transitivity, assume $\Gamma \vdash \varphi$ and $\Delta, \varphi \vdash \psi$. Then, there are derivations of φ from Γ and of ψ form Δ, φ. Thus, the sequence that results from concatenating these two derivations in order is itself a derivation of ψ from $\Gamma \cup \Delta$. Therefore, $\Gamma \cup \Delta \vdash \psi$.

The proof of monotonicity is left as an exercise. □

Exercise 2.1. *Finish the proof of Theorem 2.1.*

Theorem 2.2 (deduction). *If* $\Gamma, \varphi \vdash \psi$, *then* $\Gamma \vdash \varphi \to \psi$.[2]

Gloss: If there is a derivation of ψ from Γ together with φ, then there's a derivation of $(\varphi \to \psi)$ from Γ alone.

Proof. Assume $\Gamma, \varphi \vdash \psi$. By the definition of proof-theoretic consequence, there must be a derivation of ψ from Γ, φ. By *induction* on the length of the derivation of ψ from Γ, φ – i.e. induction on derivations by proxy (see the appendix to this chapter), we will show that there is also a derivation of $(\varphi \to \psi)$ from Γ, i.e. that $\Gamma \vdash \varphi \to \psi$. In other words, we will show the following:

> For every positive integer x, if there is an x-step derivation of ψ from Γ, φ, then $\Gamma \vdash \varphi \to \psi$.

Base step: Assume there is a one-step long derivation of ψ from Γ, φ. Then, there are only three possibilities: ψ is an instance of an axiom of the calculus, a member of Γ, or ψ is φ. In the first two cases, we can give the following derivation of $(\varphi \to \psi)$ from Γ:

1.	$\psi \to (\varphi \to \psi)$	A1
2.	ψ	axiom/member of Γ
3.	$\varphi \to \psi$	MP 1,2

If ψ is φ, by Example 2.4, $\vdash \varphi \to \psi$ and, by Theorem 2.1 (Monotonicity), $\Gamma \vdash \varphi \to \psi$.

Inductive step: Assume that, if a formula χ can be derived from Γ, φ in fewer than x steps, then $\Gamma \vdash \varphi \to \chi$. If there is an x-step derivation of ψ from Γ, φ, there are now five possibilities: ψ is an instance of an axiom of the calculus, a

[2] Also known as 'introduction of the conditional' or $\to$I.

member of Γ, ψ is φ, ψ was obtained by Modus Ponens, or it was obtained by Generalisation. The first three cases can be dealt with as in the base step.

Assume ψ was obtained by Modus Ponens. Then, there must be wff χ and $(\chi \to \psi)$ in the derivation that occur before φ. Since these wff are derivable from Γ, φ in fewer than x steps, by inductive hypothesis we have that $\Gamma \vdash \varphi \to \chi$ and $\Gamma \vdash \varphi \to (\chi \to \psi)$. That is, there are derivations of $(\varphi \to \chi)$ and of $(\varphi \to (\chi \to \psi))$ form Γ. Pasting these two derivations together and then extending the result as follows delivers a derivation of $(\varphi \to \psi)$ from Γ:

	$\vdots$	
$m.$	$\varphi \to \chi$	
	$\vdots$	
$n.$	$\varphi \to (\chi \to \psi)$	
$n+1.$	$(\varphi \to (\chi \to \psi)) \to ((\varphi \to \chi) \to (\varphi \to \psi))$	A2
$n+2.$	$(\varphi \to \chi) \to (\varphi \to \psi)$	MP $n, n+1$
$n+3.$	$\varphi \to \psi$	MP $m, n+2$

Assume now that ψ was obtained by Generalisation, and let Δ be the set of wff in $\Gamma \cup \{\varphi\}$ that occur in the derivation. Then, there must be a wff χ occurring before in the derivation s.t. ψ is $\forall v\, \chi$ and v is not free in any member of Δ. Either $\varphi \in \Delta$ or $\varphi \notin \Delta$.

If $\varphi \notin \Delta$, we have that $\Delta \vdash \psi$, which, by A1, entails that $\Delta \vdash \varphi \to \psi$. By monotonicity, $\Gamma \vdash \varphi \to \psi$.

If $\varphi \in \Delta$, since χ is derivable from Δ in fewer than x steps, by inductive hypothesis we know there is a derivation of $(\varphi \to \chi)$ from $\Delta - \{\varphi\}$, and v is not free in φ. Thus, we can extend the derivation of $(\varphi \to \chi)$ from $\Delta - \{\varphi\}$ as follows:

	$\vdots$	
$m.$	$\varphi \to \chi$	
$m+1.$	$\forall v\, (\varphi \to \chi)$	Gen m
$m+2.$	$\forall v\, (\varphi \to \chi) \to (\forall v\, \varphi \to \forall v\, \chi)$	A4
$m+3.$	$\forall v\, \varphi \to \forall v\, \chi$	MP $m+1, m+2$
$m+4.$	$\varphi \to \forall v\, \varphi$	A6
	$\vdots$	
$n.$	$\varphi \to \forall v\, \chi$	Example 2.1 $m+3, m+4$

Thus, $\Delta - \{\varphi\} \vdash \varphi \to \forall v\, \chi$. By monotonicity, $\Gamma \vdash \varphi \to \forall v\, \chi$.

By induction, we can then conclude from the base and the inductive steps that, if $\Gamma, \varphi \vdash \psi$, no matter how long the derivation of ψ from Γ, φ is, $\Gamma \vdash \varphi \to \psi$. $\square$

Example 2.5 (ex falso quodlibet)**.** $\varphi, \neg\varphi \vdash \psi$

1.	φ	premise
2.	$\neg\varphi$	premise
3.	$(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)$	A3
4.	$\neg\varphi \to (\neg\psi \to \neg\varphi)$	A1
5.	$\neg\psi \to \neg\varphi$	MP 2, 4
6.	$\varphi \to \psi$	MP 3, 5
7.	ψ	MP 1,6

Example 2.6 (consequentia mirabilis)**.** $\neg\varphi \to \varphi \vdash \varphi$

1.	$\neg\varphi \to \varphi$	premise
2.	$\neg\varphi$	assump.
3.	φ	MP 1,2
4.	$\neg(\neg\varphi \to \varphi)$	EFQ 2,3
5.	$\neg\varphi \to \neg(\neg\varphi \to \varphi)$	$\to$I 2-4
6.	$(\neg\varphi \to \neg(\neg\varphi \to \varphi)) \to ((\neg\varphi \to \varphi) \to \varphi)$	A3
7.	$(\neg\varphi \to \varphi) \to \varphi$	MP 5,6
8.	φ	MP 1,7

Proposition 2.1 (natural deduction rules)**.**

- *Double Negation* ($\neg$E). $\neg\neg\varphi \vdash \varphi$
- *Reductio ad Absurdum* ($\neg$I.) *If* $\Gamma, \varphi \vdash \psi$ *and* $\Gamma, \varphi \vdash \neg\psi$, *then* $\Gamma \vdash \neg\varphi$
- ($\forall$E). $\forall v\, \varphi \vdash \varphi[t/v]$

Proof. For $\neg$E, consider the following derivation:

1.	$\neg\neg\varphi$	premise
2.	$\neg\varphi$	assump.
3.	φ	EFQ 1,2
4.	$\neg\varphi \to \varphi$	$\to$I 2-3
5.	φ	CM 4

For $\neg$I, assume $\Gamma, \varphi \vdash \psi$ and $\Gamma, \varphi \vdash \neg\psi$. By the Deduction Theorem, there are derivations of $\varphi \to \psi$ and of $\varphi \to \neg\psi$ from Γ, which we can paste together and extend the result to a derivation of $\neg\varphi$ from Γ as follows:

	$\vdots$	
$m.$	$\varphi \to \psi$	
	$\vdots$	
$n.$	$\varphi \to \neg\psi$	
$n+1.$	$\neg\neg\varphi$	assump.
$n+2.$	φ	$\neg$E $n+1$
$n+3.$	ψ	MP $m, n+2$
$n+4.$	$\neg\psi$	MP $n, n+2$
$n+5.$	$\neg\varphi$	EFQ $n+3, n+4$
$n+6.$	$\neg\neg\varphi \to \neg\varphi$	$\to$I $n+1$-$n+5$
$n+7.$	$\neg\varphi$	CM $n+6$

□

Exercise 2.2. *Finish the proof of Proposition 2.1.*

Exercise 2.3. *Prove the following claims:*

- *Modus Tollens.* $\varphi \to \psi, \neg\psi \vdash \neg\varphi$
- *Converse Double Negation.* $\varphi \vdash \neg\neg\varphi$
- *Consequentia Mirabilis. If* $\Gamma, \neg\varphi \vdash \psi$ *and* $\Gamma, \neg\varphi \vdash \neg\psi$*, then* $\Gamma \vdash \varphi$
- ($\wedge$I). $\varphi, \psi \vdash \varphi \wedge \psi$
- ($\wedge$E1). $\varphi \wedge \psi \vdash \varphi$
- ($\wedge$E2). $\varphi \wedge \psi \vdash \psi$
- ($\vee$I1). $\varphi \vdash \varphi \vee \psi$
- ($\vee$I2). $\psi \vdash \varphi \vee \psi$
- ($\vee$E). $\varphi \vee \psi, \varphi \to \chi, \psi \to \chi \vdash \chi$
- *Disjunctive Syllogism.* $\varphi \vee \psi, \neg\varphi \vdash \psi$
- ($\leftrightarrow$I). $\varphi \to \psi, \psi \to \varphi \vdash \varphi \leftrightarrow \psi$
- ($\leftrightarrow$E1). $\varphi \leftrightarrow \psi, \varphi \vdash \psi$
- ($\leftrightarrow$E2). $\varphi \leftrightarrow \psi, \psi \vdash \varphi$
- ($\forall$I). *If* $\Gamma \vdash \varphi[t/v]$*, then* $\Gamma \vdash \forall v\, \varphi$*, if* t *is either an individual constant not occurring in* Γ *or an individual variable not free in* Γ.[3]

[3]Tip: Use induction on the length of the derivation of $\varphi[t/v]$ from Γ.

- ($\exists$I). $\varphi[t/v] \vdash \exists v\, \varphi$
- ($\exists$E). *If* $\Gamma, \exists v\, \varphi, \varphi[t/v] \vdash \psi$*, then* $\Gamma, \exists v\, \varphi \vdash \psi$*, if* t *is either an individual constant not occurring in* Γ *or* ψ *or an individual variable not free in* Γ *or* ψ*.*

One of the standard Fitch-style natural deduction calculus for first-order logic with identity consists of the inference rules $\neg$I, $\neg$E, $\to$I, $\to$E, $\forall$I, $\forall$E, and the axioms A7 and A8 to govern identity.

Exercise 2.4. *Prove the following claims:*

- *Law of Excluded Middle.* $\vdash \varphi \vee \neg\varphi$
- *Law of Non-Contradiction.* $\vdash \neg(\varphi \wedge \neg\varphi)$
- *Symmetry of Identity.* $\vdash \forall x \forall y\, (x = y \to y = x)$
- *Transitivity of Identity.* $\vdash \forall x \forall y \forall z\, (x = y \wedge y = z \to x = z)$
- $\vdash \exists! x\, x = c$[4]
- $\vdash \forall x \exists! y\, y = f(x)$
- $\vdash \forall x_1 \ldots \forall x_n \exists! y\, y = f(x_1, \ldots, x_n)$

Definition 2.5 (negation consistency)**.** Γ is negation consistent iff there is no φ s.t. φ and $\neg\varphi$ are derivable from Γ.

Exercise 2.5. *Give an example of a negation inconsistent set and show it's negation inconsistent.*

Definition 2.6 (absolute consistency)**.** Γ is absolutely consistent iff there is a φ s.t. $\Gamma \nvdash \varphi$.

Exercise 2.6. *Give an example of an absolutely inconsistent set and show it's absolutely inconsistent.*

Proposition 2.2. Γ *is negation consistent iff it's absolutely consistent.*

Exercise 2.7. *Prove Proposition 2.2.*

Since both notions are extensionally equivalent, we can drop the qualifiers and simply talk of *(in)consistent* sets of wff.

Exercise 2.8. *Show that* $\Gamma \nvdash \varphi$ *iff* $\Gamma \cup \{\neg\varphi\}$ *is consistent.*

---

[4] $\exists! v\, \varphi$ is short for $\exists v\, (\varphi \wedge \forall u\, (\varphi[u/v] \to u = v))$; it says that there is a unique object of which φ holds.

2.2 Theories and definitions

Definition 2.7 (theory)**.** A theory T formulated in $\mathscr{L}$ is a set of sentences of $\mathscr{L}$ s.t., for every sentence φ, if $T \vdash \varphi$, then $\varphi \in T$. Each member of T is a theorem of T. Any Γ s.t., for every sentence φ, $\Gamma \vdash \varphi$ iff $\varphi \in T$ is an axiomatisation of T.

Note that, if Γ is an axiomatisation of T, then $\Gamma \subseteq T$.

Example 2.7 (first-order Peano arithmetic)**.** PA is the theory formulated in $\mathscr{L}_{\mathsf{PA}}$ axiomatised by the following:

PA1. $\forall x\, 0 \neq \mathrm{S}x$

PA2. $\forall x \forall y\, (\mathrm{S}x = \mathrm{S}y \rightarrow x = y)$

PA3. $\forall x\, x + 0 = x$

PA4. $\forall x \forall y\, x + \mathrm{S}y = \mathrm{S}(x + y)$

PA5. $\forall x\, x \times 0 = 0$

PA6. $\forall x \forall y\, x \times \mathrm{S}y = x \times y + x$

Induction Schema. $\varphi(0) \wedge \forall x\, (\varphi(x) \rightarrow \varphi(\mathrm{S}x)) \rightarrow \forall x\, \varphi(x)$[5]

Example 2.8. $\mathsf{PA} \vdash \mathrm{S}0 \neq \mathrm{SS}0$

1.	$\mathrm{S}0 = \mathrm{SS}0$	assump.
2.	$\forall x \forall y\, (\mathrm{S}x = \mathrm{S}y \rightarrow x = y)$	PA2
3.	$\mathrm{S}0 = \mathrm{SS}0 \rightarrow 0 = \mathrm{S}0$	$\forall$E 2
4.	$0 = \mathrm{S}0$	MP 1,3
5.	$\forall x\, 0 \neq \mathrm{S}x$	PA1
6.	$0 \neq \mathrm{S}0$	$\forall$E 5
7.	$\mathrm{S}0 \neq \mathrm{SS}0$	$\neg$I 1,4,6

Example 2.9. $\mathsf{PA} \vdash \forall x\, 0 + x = x$

[5] We write $\varphi(t)$ to indicate that the term t occurs in φ. If t is a variable, $\varphi(t)$ also indicates that t is occurs free in φ. Additionally, $\varphi(v_1, \ldots, v_n)$ indicates that $v_1, \ldots, v_n$ are the only variables that occur free in φ.

1.	$\forall x\, x + 0 = x$	PA3
2.	$0 + 0 = 0$	$\forall$E 1
3.	$0 + x = x$	assump.
4.	$\forall x \forall y\, x + \mathrm{S}y = \mathrm{S}(x + y)$	PA4
5.	$\forall y\, 0 + \mathrm{S}y = \mathrm{S}(0 + y)$	$\forall$E 4
6.	$0 + \mathrm{S}x = \mathrm{S}(0 + x)$	$\forall$E 5
7.	$\forall y \forall z\, (y = z \to (0 + \mathrm{S}x = \mathrm{S}y \to 0 + \mathrm{S}x = \mathrm{S}z))$	A8
8.	$\forall z\, (0 + x = z \to (0 + \mathrm{S}x = \mathrm{S}(0 + x) \to 0 + \mathrm{S}x = \mathrm{S}z))$	$\forall$E 7
9.	$0 + x = x \to (0 + \mathrm{S}x = \mathrm{S}(0 + x) \to 0 + \mathrm{S}x = \mathrm{S}x)$	$\forall$E 8
10.	$0 + \mathrm{S}x = \mathrm{S}(0 + x) \to 0 + \mathrm{S}x = \mathrm{S}x$	MP 3,9
11.	$0 + \mathrm{S}x = \mathrm{S}x$	MP 6,10
12.	$0 + x = x \to 0 + \mathrm{S}x = \mathrm{S}x$	$\to$I 3-11
13.	$\forall x\, (0 + x = x \to 0 + \mathrm{S}x = \mathrm{S}x)$	$\forall$I 12
14.	$0 + 0 = 0 \wedge \forall x\, (0 + x = x \to 0 + \mathrm{S}x = \mathrm{S}x)$	$\wedge$I 2,13
15.	$0 + 0 = 0 \wedge \forall x\, (0 + x = x \to 0 + \mathrm{S}x = \mathrm{S}x) \to \forall x\, 0 + x = x$	IS
16.	$\forall x\, 0 + x = x$	MP 14,15

Or in a sloppier but less cumbersome way:

1.	$0 + 0 = 0$	$\forall$E PA3
2.	$0 + x = x$	assump.
3.	$0 + \mathrm{S}x = \mathrm{S}(0 + x)$	$\forall$E PA4
4.	$0 + \mathrm{S}x = \mathrm{S}x$	$\forall$E, MP, A8, 3
5.	$\forall x\, (0 + x = x \to 0 + \mathrm{S}x = \mathrm{S}x)$	$\to$I 2-4, $\forall$I
6.	$\forall x\, 0 + x = x$	IS, MP 1,5

Exercise 2.9. *Show that* $\mathsf{PA} \vdash \forall x \forall y\, x + y = y + x$.[6]

Exercise 2.10. *What is the smallest theory formulated in a given first-order language $\mathscr{L}$? Provide an axiomatisation of this theory. What is the smallest set that axiomatises this theory?*

Exercise 2.11. *What is the largest theory formulated in a given first-order language $\mathscr{L}$? Provide an axiomatisation of this theory that contains only one axiom.*

Next, we define what it means to introduce a non-logical symbol (predicate, function, constant) by 'definition' to a theory. These definitions are both explicit and conventional. They are not meant to provide a deep analysis of the

[6] Tip: You need to use an instance of the Induction Schema.

concept expressed by the introduced symbol – i.e. the *definiendum* (that what is being defined) – or to capture our usage of this symbol. They are rather abbreviations of the *definiens* (the expression we use to define the *definiendum*) we stipulate for convenience. These definitions are *stipulative*.

In what follows, let T be a theory formulated in $\mathscr{L}$ and let Γ be an axiomatisation of T. To introduce a symbol ξ to T by definition one first needs to *expand* $\mathscr{L}$ to $\mathscr{L}^+$ by adding ξ to $V(\mathscr{L})$ – i.e. $V(\mathscr{L}^+) = V(\mathscr{L}) \cup \{\xi\}$. Next, we extend Γ by adding a definition of ξ, obtaining Γ^+. Finally, T is extended to the theory T^+, formulated in $\mathscr{L}^+$ and axiomatised by Γ^+. T^+ will now contain all the sentences of the expanded language that follow from the extended set of axioms.

The following definitions provide a precise account of what counts as a definition of a symbol ξ for a given axiomatisation Γ of a theory.

Definition 2.8 (predicate definition)**.** The following can be added as a definition of the n-place predicate symbol P to Γ iff P occurs neither in Γ nor in φ, and $v_1, \dots, v_n$ are distinct variables (and the only free variables in φ):

$$\forall v_1 \dots \forall v_n \, (P(v_1, \dots, v_n) \leftrightarrow \varphi(v_1, \dots, v_n))$$

In such definitions, P is the *definiendum* and φ is the *definiens*.

Definition 2.9 (constant definition)**.** The following can be added as a definition of the individual constant c to Γ iff c occurs neither in Γ nor in φ and $\Gamma \vdash \exists! v \, \varphi(v)$:

$$\forall v \, (v = c \leftrightarrow \varphi(v))$$

In such definitions, c is the *definiendum* and φ is the *definiens*.

Definition 2.10 (function symbol definition)**.** The following can be added as a definition of the n-place function symbol f to Γ iff f occurs neither in Γ nor in φ, $v_1, \dots, v_n, u$ are distinct variables, and $\Gamma \vdash \forall v_1 \dots \forall v_n \exists! u \, \varphi(v_1, \dots, v_n, u)$:

$$\forall v_1 \dots \forall v_n \forall u \, (u = f(v_1, \dots, v_n) \leftrightarrow \varphi(v_1, \dots, v_n, u))$$

In such definitions, f is the *definiendum* and φ is the *definiens*.

A symbol occurring in the axioms of a theory that hasn't been introduced by definition is called a 'primitive symbol' of the theory.

Exercise 2.12. *Show that the following can be successively added to the axioms of* PA *as definitions of the new corresponding symbols:*

1. $\forall x \forall y \, (x < y \leftrightarrow \exists z \, y = x + \mathrm{S}z)$
2. $\forall x \, (x = 1 \leftrightarrow x = \mathrm{S}0)$
3. $\forall x \, (x = 2 \leftrightarrow x = \mathrm{S}1)$
4. $\forall x \forall y \, (\mathrm{half}(x) = y \leftrightarrow x = 2 \times y \vee x = 2 \times y + 1)$

5. $\forall x \forall y \forall z\, (\mathrm{div}(x, y) = z \leftrightarrow (y = 0 \land z = 0) \lor (y \neq 0 \land \exists w\, (w < y \land x = y \times z + w)))$

Exercise 2.13. *Define the* 2*-place function symbol* f *(or '$-$', if you prefer) in* PA *to express the subtraction operation between natural numbers; i.e. if* $x < y$, $x - y = 0$.

Theorem 2.3 (conservativeness)**.** *If* φ *can be added as a definition of* ξ *to* Γ*, then, for every* ψ *not containing* ξ*, if* $\Gamma, \varphi \vdash \psi$*,* $\Gamma \vdash \psi$*.*

Gloss: Introducing a new symbol by definition to a theory doesn't allow us to prove new theorems not containing the new symbol (i.e. new theorems expressed in the old language).

Proof. We show it for the case in which ξ is an n-place predicate symbol. Let $\chi(v_1, \ldots, v_n)$ be the *definiens* of ξ in φ and assume $\Gamma, \varphi \vdash \psi$. Then replace in a derivation of ψ from Γ, φ all occurrences of $\xi(t_1, \ldots, t_n)$ with $\chi(t_1, \ldots, t_n)$. It can be shown by induction on the length of this derivation that the result is a derivation of ψ from Γ. $\square$

Exercise 2.14. *Carry out the proof by induction hinted at in the proof of Theorem 2.3.*

Corollary 2.1. *If* φ *can be added as a definition of* ξ *to* Γ*, then* $\Gamma \cup \{\varphi\}$ *is inconsistent just in case* Γ *is so too.*

Exercise 2.15. *Prove Corollary 2.1.*

Exercise 2.16. *Give examples to show that, if one of the following holds, then adding* $\forall x\, (P(x) \leftrightarrow \varphi(x))$ *to a set* Γ *could result in an inconsistent set:*

1. φ *contains* P

2. P *occurs in* Γ

Exercise 2.17. *Show that if* Γ *entails one of the following, adding* $\forall x\, (x = c \leftrightarrow \varphi(x))$ *to* Γ *results in an inconsistent set:*

1. $\neg \exists x\, \varphi(x)$

2. $\exists x \exists y\, (\varphi(x) \land \varphi(y) \land x \neq y)$

Appendix: Induction for derivations

Note that every derivation in our Hilbert-style calculus has a finite length. The number of steps in each derivation is always a positive integer. Thus, we can use the following instance of total induction for positive integers to prove universal claims about derivations in our calculus:

Definition 2.11 (total induction for derivations by proxy)**.**

- *Base step:* All 1-step long derivations are P.
- *Inductive step:* For every positive integer x, if all derivations with fewer than x steps are P, then all x-step long derivations are P.
- *Conclusion:* For every positive integer x, all x-step long derivations are P. (In other words, every derivation is P.)

To establish the base step, one can proceed by cases, by considering all possible 1-step long derivations. Note that these will always consist either of an axiom or a premise.

For the inductive step, the easiest way is to proceed by cases, i.e. to consider all possible ways in which the last wff φ of the derivation was obtained:

1. φ is an axiom or a premise. These cases have already been considered in the base step.
2. φ's been obtained by Modus Ponens, that is, there are previous entries ψ and $(\psi \to \varphi)$ in the derivation, for some wff ψ. Note that, since φ was obtained in the x-th step, ψ and $(\psi \to \varphi)$ must have been derived in fewer than x steps. By inductive hypothesis, the derivations of ψ and $(\psi \to \varphi)$ are P. Usually this turns out to be useful to prove that the derivation of φ is also P.
3. φ's been obtained by Gen, that is, φ is $\forall v\, \psi$ for some wff ψ and some variable v, ψ occurs earlier in the derivation, and v doesn't occur free in any premises. Note that, since φ was obtained in the x-th step, ψ must have been derived in fewer than x steps. By inductive hypothesis, the derivation of ψ is P, which usually turns out to be useful to prove that the derivation of φ is also P.

Derivations are also entitled to their own, tailored induction principle. Just like all positive integers can be obtained by finitely many applications of the successor operation to 1, every derivation can be obtained by finitely many applications of Modus Ponens and Gen to the premises or axioms of the calculus. This structural feature of the derivations of our calculus entails the following induction principle:

Definition 2.12 (tailored induction for derivations)**.**

- *Base step:* All 1-step long derivations are P.
- *Inductive step 1:* For every derivation $\mathcal{D}$, if $\mathcal{D}$ is P, then the derivations that result from extending $\mathcal{D}$ with an axiom or a premise are also P.
- *Inductive step 2:* For every derivation $\mathcal{D}$, if $\mathcal{D}$ is P, then the derivations that result from extending $\mathcal{D}$ with a wff by Modus Ponens are also P.
- *Inductive step 3:* For every derivation $\mathcal{D}$, if $\mathcal{D}$ is P, the derivations that result from extending $\mathcal{D}$ with a wff by Generalisation are also P.
- *Conclusion:* Every derivation is P.

Chapter 3

Set Theory

Sets are collection of objects, called the 'elements' or 'members' of the set. They are said to *belong to* or to *be in* the set. Sets are characterised *exclusively* by their members: same members, same set.

'Set theory' is an umbrella term for several first-order theories – all expressed in the (an expansion of the) same language, $\mathscr{L}_\in$ – with a common core.

3.1 The language

Definition 3.1 (the language of pure set theory)**.** $\mathscr{L}_\in$ is the first-order language s.t. $V(\mathscr{L}_\in) = \{P^2_0\}$, i.e. it has just one non-logical symbol, the 2-place predicate P^2_0. For convenience, we write '$\in$' for P^2_0.

We write $t_1 \in t_2$ instead of $\in(t_1, t_2)$, for perspicuity. The intuitive reading is 't_1 is an element of t_2'. If t_1 is not an element of t_2, we write $t_1 \notin t_2$ as short for $\neg t_1 \in t_2$. For convenience, we will often use constants (in the metalanguage) to name certain sets. We will normally use low case Roman letters $a, b, \ldots$

Effectively, we will work with an expansion of $\mathscr{L}_\in$, $\mathscr{L}^+_\in$, that contains extra non-logical terms to be specified that will be introduced by definition for brevity. We will also informally add vocabulary to talk about sets of urelements – i.e. objects that are not sets, such as numbers, letters, formulae, etc. – instead of just pure sets – i.e. sets whose only elements are other pure sets.

3.2 Naive Set Theory

A theory of sets typically consists of axioms indicating what sets exist and which don't. Naive Set Theory consists of one axiom ruling out the existence of sets, namely, Extensionality, and another axiom-schema asserting the existence of sets, namely, Naive Comprehension.

Axiom (Extensionality)**.** $\forall x \forall y\, (\forall z\, (z \in x \leftrightarrow z \in y) \to x = y)$

Gloss: If x and y have the same members, then they are the same set. In other words, there are no two sets with the same members. Extensionality expresses one of the fundamental properties of sets: sets are exclusively characterised by their members.

Exercise 3.1. *The converse of Extensionality can be formalised as*

$$\forall x \forall y\, (x = y \rightarrow \forall z\, (z \in x \leftrightarrow z \in y))$$

Is it true? Why?

When trying to answer the question what sets exist, one might naively think that every (meaningful) predicate determines a set. There is a set of all humans, a set of all natural numbers and, in general, a set of all Ps, for each predicate P. This can be expressed in $\mathscr{L}_\in$ by the following schematic principle, where φ a wff of the language:

Axiom (Naive Comprehension)**.** $\exists x \forall y\, (y \in x \leftrightarrow \varphi(y))$

Gloss: For each wff $\varphi(y)$, there is a set x whose members are exactly the things that satisfy φ/that are φ.

Unfortunately, Naive Comprehension is *inconsistent*, as *Russell's Paradox* shows.[1] Consider the predicate 'is a set that is not a member of itself'. By Naive Comprehension, there must be a set, r, that has as members exactly those sets that are not members of themselves. There are then two possibilities: either r is a member of itself or r is not a member of itself. If r belongs to itself, since the members of r are exactly those sets that are not members it itself, then r must not be a member of itself, which contradicts our initial assumption. Thus, it must be that r is not a member of itself. But then, r must be a member of r after all. Again, we reached a contradiction. In any case, we obtain a contradiction.

More formally, we have the following:

1.	$\exists x \forall y\, (y \in x \leftrightarrow y \notin y))$	Naive Comprehension
2.	$\forall y\, (y \in x \leftrightarrow y \notin y))$	assump.
3.	$x \in x \leftrightarrow x \notin x$	$\forall$E 2
4.	$\exists x\, (x \in x \leftrightarrow x \notin x)$	$\exists$I 3
5.	$\exists x\, (x \in x \leftrightarrow x \notin x)$	$\exists$E 1,2-4

Since Naive Comprehension must be given up, we must find other, consistent ways of establishing the existence of sets. The most popular theory of sets is Zermelo-Fraenkel set theory.

[1] Russell's Paradox was a big blow especially for Frege who, in his *Grundgesetze der Arithmetik*, employed Naive Comprehension – which he called 'Basic Law V' – to formulate a theory of extensions (entities similar to sets in many respects), which was part of his logic.

3.3 Zermelo-Fraenkel Set Theory

In this section we introduce the axioms of Zermelo-Fraenkel set theory (ZF) and explore some of their consequences. The axioms of ZF are Extensionality, Empty Set, Separation, Pairing, Union, Power Set, Infinity, Replacement, and Regularity. We also consider the axiom of Choice, which gives us the stronger system known as ZFC.

Naive Set Theory and ZF share the axiom of Extensionality.

Proposition 3.1. $\mathsf{ZF} \vdash \exists x \forall y\,(y \in x \leftrightarrow \varphi(y)) \to \exists! x \forall y\,(y \in x \leftrightarrow \varphi(y))$ *(for every wff $\varphi(y)$ of $\mathscr{L}_\in^+$)*

Gloss: If there's a set of φs, then it is unique.

Proof.

1.	$\exists x \forall y\,(y \in x \leftrightarrow \varphi(y))$	assump.
2.	$\forall y\,(y \in x \leftrightarrow \varphi(y))$	assump.
3.	$\forall y\,(y \in z \leftrightarrow \varphi(y))$	assump.
4.	$\forall y\,(y \in x \leftrightarrow y \in z)$	2,3
5.	$x = z$	Ext 4
6.	$\forall z\,(\forall y\,(y \in z \leftrightarrow \varphi(y)) \to z = x)$	$\to$I, $\forall$I 3-5
7.	$\forall y\,(y \in x \leftrightarrow \varphi(y)) \wedge \forall z\,(\forall y\,(y \in z \leftrightarrow \varphi(y)) \to z = x)$	2,6
8.	$\exists x(\forall y\,(y \in x \leftrightarrow \varphi(y)) \wedge \forall z\,(\forall y\,(y \in z \leftrightarrow \varphi(y)) \to z = x))$	$\exists$I 7
9.	$\exists x(\forall y\,(y \in x \leftrightarrow \varphi(y)) \wedge \forall z\,(\forall y\,(y \in z \leftrightarrow \varphi(y)) \to z = x))$	$\exists$I 1,2-8
10 .	$\exists! x \forall y\,(y \in x \leftrightarrow \varphi(y))$	notation 9
11.	$\exists x \forall y\,(y \in x \leftrightarrow \varphi(y)) \to \exists! x \forall y\,(y \in x \leftrightarrow \varphi(y))$	$\to$I 1-10

□

By Definition 2.9 and Proposition 3.1, whenever we can prove in ZF that the set of things that have the property expressed by $\varphi(y)$ exists – i.e. $\exists x \forall y\,(y \in x \leftrightarrow \varphi(y))$ – we can introduce an individual constant to denote this set. By default, this constant will be $\{v : \varphi(v)\}$, or $\{t_1, \dots, t_n\}$ if φ is of the form $v = t_1 \vee \cdots \vee v = t_n$, although in some cases we will introduce special names.[2]

For instance, $\{v : v$ is a natural number smaller than $2\}$ is the set whose members are 0 and 1, and $\{0, 1, 2\}$ is the set whose members are 0, 1, and 2. We will often use the following special names for sets of numbers:

- $\mathbb{N} = \{x : x$ is a natural number$\} = \{0, 1, 2, 3, \dots\}$

[2]The new constant c is defined by $\forall x\,(x = c \leftrightarrow \forall y\,(y \in x \leftrightarrow \varphi(y)))$ which, by extensionality, is equivalent to $\forall y\,(y \in c \leftrightarrow \varphi(y))$.

- $\mathbb{Z} = \{x : x \text{ is an integer }\} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$
- $\mathbb{Q} = \{x : x \text{ is a rational number }\} = \{\ldots, -1, -\frac{1}{2}, 0, \frac{1}{3}, \frac{3}{4}, 1, \ldots\}$
- $\mathbb{R} = \{x : x \text{ is a real number }\} = \{\ldots, -e, -1, -\frac{1}{2}, 0, \frac{3}{4}, 1, \sqrt{2}, 3, \pi \ldots\}$

Exercise 3.2. *List the elements of the following sets:*

1. $\{x : x$ *is an even natural number smaller than 6*$\}$
2. $\{x : x$ *is an odd natural number smaller than 6*$\}$
3. $\{x : x$ *is a prime number smaller than 10*$\}$
4. $\{x : x$ *is the capital city of France*$\}$

Exercise 3.3. *Show that* $\mathsf{ZF} \vdash$

$$\forall x\,(c = x \leftrightarrow \forall y\,(y \in x \leftrightarrow \varphi(y))) \leftrightarrow \forall y\,(y \in c \leftrightarrow \varphi(y))^3$$

Let $\vec{z}$ abbreviate the string of variables $z_1, \ldots, z_n$ (possibly of length 0) and similarly for other terms and variables. We can generalise Proposition 3.1 as follows:

Proposition 3.2. $\mathsf{ZF} \vdash \forall\vec{z}\exists x\forall y\,(y \in x \leftrightarrow \varphi(y, \vec{z})) \to \forall\vec{z}\exists! x\forall y\,(y \in x \leftrightarrow \varphi(y, \vec{z}))$.

Exercise 3.4. *Prove Proposition 3.2.*

Exercise 3.5. *Formulate Naive Comprehension in $\mathscr{L}_\in$ with parameters.*

By Definition 2.10 and Proposition 3.2, whenever we can prove in ZF that, for each $\vec{z}$, the set of things that have the property expressed by $\varphi(y, \vec{z})$ exists – i.e. $\forall\vec{z}\exists x\forall y\,(y \in x \leftrightarrow \varphi(y, \vec{z}))$ – we can introduce an n-place function symbol which, by default, will be $\{v : \varphi(v, \vec{z})\}$, or $\{t_1, \ldots, t_n\}$ if φ is of the form $v = t_1 \vee \cdots \vee v = t_n$, although in some cases we will introduce special symbols.[4]

Example 3.1. $\mathsf{ZF} \vdash \{0, 1\} = \{1, 0\}$

1.	$x \in \{0, 1\} \leftrightarrow x = 0 \vee x = 1$	def. of $\{0, 1\}$
2.	$x \in \{1, 0\} \leftrightarrow x = 1 \vee x = 0$	def. of $\{1, 0\}$
3.	$x \in \{0, 1\} \leftrightarrow x \in \{1, 0\}$	1,2
4.	$\forall x\,(x \in \{0, 1\} \leftrightarrow x \in \{1, 0\})$	$\forall$I 3
5.	$\{0, 1\} = \{1, 0\}$	Ext 4

Exercise 3.6. *Show that* $\mathsf{ZF} \vdash \{0, 1\} = \{0, 0, 1\}$

[3]Tip: This follows by Extensionality.

[4]The new function symbol $f(\vec{z})$ is defined by $\forall\vec{z}\forall x\,(f(\vec{z}) = x \leftrightarrow \forall y\,(y \in x \leftrightarrow \varphi(y, \vec{z})))$ which, by extensionality, is equivalent to $\forall\vec{z}\forall y\,(y \in f(\vec{z}) \leftrightarrow \varphi(y, \vec{z}))$.

Axiom (Empty set)**.** $\exists x \forall y\,(y \in x \leftrightarrow y \neq y)$

Gloss: There is a set with no elements. By Proposition 3.1 it's unique, so we can add the constant $\{x : x \neq x\}$ to denote this set but, since it's a special set, we also have the following:

Definition 3.2 (empty set)**.** $\varnothing = \{\} = \{x : x \neq x\}$

Exercise 3.7. *Indicate whether the following are true.*

1. $\{a, b\} = \{b, a\}$
2. $\varnothing = 0$
3. $\varnothing = \{x : x \in \mathbb{N} \wedge x \notin \mathbb{N}\}$
4. $\{\varnothing\} = \{\varnothing, \{\varnothing\}\}$

Exercise 3.8. *Indicate whether the following are true.*

1. $\{a, b\} \in \{a, b\}$
2. $a \in \varnothing$
3. $\varnothing \in \varnothing$
4. $\varnothing \in \{\{\varnothing\}\}$
5. $\{\varnothing\} \in \{\{\varnothing\}\}$
6. $\{1\} \in \{1, 2, \{1\}\}$
7. $\{1\} \in \mathbb{N}$
8. $\varnothing \in \mathbb{Z}$

Proposition 3.3. $\mathsf{ZF} \vdash$

1. $\forall x\, x \notin \varnothing$
2. $\varnothing \neq \{\varnothing\}$
3. $\{\varnothing\} \neq \{\{\varnothing\}\}$

Proof.

1. Since $x = x$ (A7), by Def. of $\varnothing$, $x \notin \varnothing$. Thus, $\forall x\, x \notin \varnothing$ (by Gen).
2. Assume $\varnothing = \{\varnothing\}$. Since $\varnothing \in \{\varnothing\}$, then $\varnothing \in \varnothing$. By Def. of $\varnothing$, this implies that $\varnothing \neq \varnothing$, a contradiction.

$\square$

Exercise 3.9. *Finish the proof of Proposition 3.3.*

Exercise 3.10. *Is it true that if $x \in y$ and $y \in z$, then $x \in z$?*

Definition 3.3 (inclusion)**.** $\forall x \forall y\,(x \subseteq y \leftrightarrow \forall z\,(z \in x \to z \in y))$

Gloss: $x \subseteq y$ means that all members of x are members of y. In that case, we say that x is a *subset* of y or that x is *included* in y. If $a \subseteq b \wedge a \neq b$ – which we abbreviate as $a \subsetneq b$ – we say that a is a *proper* subset of b. Also, we often write $a \nsubseteq b$ as short for $\neg a \subseteq b$.

Example 3.2 (inclusions)**.**

- $\{0\} \subseteq \{0,1\}$ (but $\{0\} \notin \{0,1\}$!)
- $\{0,2\} \subsetneq \mathbb{N}$
- $\mathbb{Z} \subseteq \mathbb{Q}$
- $\{0,2\} \nsubseteq \{0,1\}$
- $\varnothing \subseteq \{\varnothing\}$ (and $\varnothing \in \{\varnothing\}$)
- $\varnothing \subseteq \{\{\varnothing\}\}$

Exercise 3.11. *Determine which of the following statements are true.*

1. $\varnothing \in \{\varnothing, \{\varnothing\}\}$
2. $\varnothing \subseteq \{\varnothing, \{\varnothing\}\}$
3. $1 \subseteq \{1,2,3\}$
4. $\{1\} \subseteq \{1,2,3\}$
5. $\{\varnothing, 1\} \nsubseteq \{1,2,3\}$
6. $\{\varnothing\} \nsubseteq \{\varnothing, \{\{\varnothing\}\}\}$
7. $\{\{\varnothing\}\} \subseteq \{\varnothing, \{\varnothing\}\}$
8. $\{1,2,3\} \subseteq \{1,2,3\}$
9. $\{1,2,3\} \subseteq \mathbb{Z}^+$
10. $0 \subseteq \mathbb{N}$
11. $\varnothing \subseteq \varnothing$
12. $\varnothing \in \varnothing$
13. $\mathbb{Z} \subseteq \mathbb{Q}$
14. $\mathbb{Z} \subseteq \mathbb{R}$

Proposition 3.4. $\mathsf{ZF} \vdash$

1. $\forall x\, x \subseteq x$
2. $\forall x\, \varnothing \subseteq x$
3. $\forall x \forall y\, (x \subseteq y \wedge y \subseteq x \rightarrow x = y)$
4. $\forall x \forall y \forall z\, (x \subseteq y \wedge y \subseteq z \rightarrow x \subseteq z)$

Proof. 2. Since $y \in \varnothing$ is always false (by Proposition 3.3.1), $y \in \varnothing \rightarrow y \in x$ is always true. Thus, $\forall x \forall y\, (y \in \varnothing \rightarrow y \in x)$. By Def. of $\subseteq$, $\forall x\, \varnothing \subseteq x$. ☐

Exercise 3.12. *Finish the proof of Proposition 3.4.*

Axiom (Separation). $\forall \vec{z} \forall w \exists x \forall y\, (y \in x \leftrightarrow y \in w \wedge \varphi(y, \vec{z}))$, *for each wff* $\varphi(y, \vec{z})$ *of* $\mathscr{L}_{\in}^{+}$

Gloss: For every $\vec{z}$ and a *given set* w, there exists a set that contains all objects in w satisfying φ/that are φs/all φs in w.

Separation is a schema just like Naive Comprehension, except that the members of the new set x, the things that have the property expressed by φ, are drawn from a 'pre-existing' set w, to avoid contradiction: to derive the existence of a set x using this axiom, the members-to-be must already exist and belong to a (possibly larger) set w, none of them can be x itself or defined in terms of x in any way.

In what follows, let $\{v \in u : \varphi(v)\}$ be short for $\{v : v \in u \wedge \varphi(v)\}$. For any $\varphi(x)$ and any given set w, Separation entails the existence of the set $\{x \in$

$w : \varphi(x)\}$. For instance, there's a set of negative numbers, $\{x \in \mathbb{Z} : x \notin \mathbb{N}\}$ (which we sometimes denote by $\mathbb{Z}^-$; in general we use $+$ and $-$ as superscripts to indicate that we are only taking the positive and negative elements of the set, resp.), and a set of irrational numbers, $\{x \in \mathbb{R} : x \notin \mathbb{Q}\}$.

Proposition 3.5 (intersection). $\mathsf{ZF} \vdash \forall z \forall w \exists x \forall y\, (y \in x \leftrightarrow y \in w \wedge y \in z)$

Gloss: For any two sets z and w there is a set x of things that belong both to z and to w. By Proposition 3.2, we know that, for each z and w this set x is unique, so we can introduce the following definition (rearranging the variables):

Definition 3.4 (intersection). $\forall x \forall y\, x \cap y = \{z : z \in x \wedge z \in y\}$.

Example 3.3 (intersections).

- $\{0,1\} \cap \{1,2\} = \{1\}$
- $\mathbb{Z} \cap \mathbb{Q} = \mathbb{Z}$
- $\mathbb{N} \cap \mathbb{Z}^- = \varnothing$
- $\{\varnothing\} \cap \{\{\varnothing\}\} = \varnothing$

Exercise 3.13. *Indicate the value of the following operations:*

1. $\{0,1\} \cap \{1,2,3\}$
2. $\{1,2,3\} \cap \{2,\{3\},4\}$
3. $\{x \in \mathbb{Z} : x < 0\} \cap \mathbb{N}$
4. $\mathbb{Z} \cap \{0\}$

Exercise 3.14. *Show that* $\mathsf{ZF} \vdash$

1. $\forall x\, x \cap x = x$
2. $\forall x\, x \cap \varnothing = \varnothing$
3. $\forall x \forall y\, x \cap y = y \cap x$
4. $\forall x \forall y\, x \cap y \subseteq x$
5. $\forall x \forall y \forall z\, x \cap (y \cap z) = (x \cap y) \cap z$
6. $\forall x \forall y\, (x \subseteq y \leftrightarrow x \cap y = x)$
7. $\forall x \forall y \forall z\, (x \subseteq y \rightarrow x \cap z \subseteq y \cap z)$
8. $\forall x \forall y \forall z\, (z \subseteq x \wedge z \subseteq y \leftrightarrow z \subseteq x \cap y)$

Proposition 3.6 (subtraction). $\mathsf{ZF} \vdash \forall z \forall w \exists x \forall y\, (y \in x \leftrightarrow y \in w \wedge y \notin z)$

Gloss: For any two sets z and w there is a set x of things that belong to w but not to z.

Definition 3.5 (subtraction). $\forall x \forall y\, x - y = \{z : z \in x \wedge z \notin y\}$.

Example 3.4 (subtractions).

- $\{0, 1, 2\} - \{0, 1\} = \{2\}$
- $\mathbb{Z} - \mathbb{N} = \mathbb{Z}^-$
- $\mathbb{N} - \{x : x \text{ is even}\} = \{x \in \mathbb{N} : x \text{ is odd}\}$

Exercise 3.15. *Indicate the value of the following operations.*

1. $\{0, 1, 2\} - \{2, 3\}$
2. $\{\varnothing\} - \varnothing$
3. $\{\varnothing\} - \{\varnothing\}$
4. $\mathbb{N} - \mathbb{Z}$

Exercise 3.16. $\mathsf{ZF} \vdash$

1. $\forall x\, x - \varnothing = x$
2. $\forall x\, \varnothing - x = \varnothing$
3. $\forall x\, x - x = \varnothing$
4. $\forall x \forall y\, x - y \subseteq x$

Exercise 3.17. *Give a counterexample to the following claim:* $\forall x \forall y \forall z\, (x - y) - z = x - (y - z)$.

Theorem 3.1 (no set of all sets)**.** $\mathsf{ZF} \vdash \neg \exists y \forall x\, x \in y$

Proof. Assume u is the universal set, i.e. $\forall x\, x \in u$. By Separation,

$$\exists x \forall y\, (y \in x \leftrightarrow y \in u \wedge y \notin y)$$

which is equivalent to

$$\exists x \forall y\, (y \in x \leftrightarrow y \notin y)$$

for u is the universal set. The latter is the instance of Naive Comprehension involved in Russell's Paradox. □

Axiom (Pairing)**.** $\forall x \forall y \exists z \forall w\, (w \in z \leftrightarrow w = x \vee w = y)$

Gloss: For every x and y, there exists a set z whose elements are exactly x and y, which we denote by $\{x, y\}$.

Proposition 3.7 (singletons)**.** $\mathsf{ZF} \vdash \forall x \exists y \forall z\, (z \in y \leftrightarrow z = x)$

Gloss: For every x there is a set z that has x as its only member, which we denote by $\{x\}$ (the 'singleton' of x).

Proof.

1.	$\forall x \forall y \exists z \forall w\, (w \in z \leftrightarrow w = x \vee w = y)$	Pairing
2.	$\exists z \forall w\, (w \in z \leftrightarrow w = x \vee w = x)$	$\forall$E 1
3.	$\forall x \exists z \forall w\, (w \in z \leftrightarrow w = x)$	$\forall$I 2

□

Proposition 3.8 (ordered pairs)**.** $\mathsf{ZF} \vdash \forall x \forall y \exists z \forall w\, (w \in z \leftrightarrow w = \{x\} \vee w = \{x, y\})$.

Gloss: For every x and every y, there is a set whose only members are $\{x\}$ and $\{x, y\}$, i.e. $\{\{x\}, \{x, y\}\}$. We refer to this set as 'the ordered pair of x and y'.

Definition 3.6 (ordered pairs)**.** $\forall x \forall y\, \langle x, y\rangle = \{\{x\}, \{x, y\}\}$.

Definition 3.7 (ordered tuples)**.** $\forall x_1 \dots \forall x_n\, \langle x_1, \dots, x_n\rangle = \langle\langle x_1, \dots, x_{n-1}\rangle, x_n\rangle$, for every $n > 2$.

Example 3.5 (ordered tuples)**.**

- $\{0\} \neq \langle 0, 0\rangle$
- $\{0, 1\} \neq \langle 0, 1\rangle$
- $\langle 0, 1\rangle \neq \langle 1, 0\rangle$
- $\langle 1, 1\rangle \neq \langle 1, 1, 1\rangle$

Proposition 3.9. $\mathsf{ZF} \vdash$

1. $\forall x \forall y \forall z \forall w\, (\langle x, y\rangle = \langle z, w\rangle \leftrightarrow x = z \wedge y = w)$
2. *If* $n \geqslant 2$, $\forall x_1 \dots \forall x_n \forall y_1 \dots \forall y_n (\langle x_1, \dots, x_n\rangle = \langle y_1, \dots, y_n\rangle \leftrightarrow x_1 = y_1 \wedge \dots \wedge x_n = y_n)$
3. $\forall x\, \{x\} \neq \langle x, x\rangle$
4. $\forall x \forall y\, (x \neq y \rightarrow \langle x, y\rangle \neq \langle y, x\rangle)$

Proof. 1. For the left-to-right direction, assume $\langle x, y\rangle = \langle z, w\rangle$. By Def. of ordered pairs, $\{\{x\}, \{x, y\}\} = \{\{z\}, \{z, w\}\}$. Either $x = y$ or $x \neq y$. By Extensionality, if $x = y$ we have that $\{x, y\} = \{x\}$, so $\{x\} = \{z\}$ and $\{x\} = \{z, w\}$. This implies that $x = z$ and $x = w$, which entails $y = w$. If $x \neq y$ instead, Extensionality implies that $\{x\} = \{z\}$ and $\{x, y\} = \{z, w\}$. Then, $x = z$ and, thus, $y = w$. Therefore, $\langle x, y\rangle = \langle z, w\rangle \rightarrow x = z \wedge y = w$. The right-to-left direction is left as an exercise.

2. The proof is by *induction* on n.

   Base step: $n = 2$. The proof has been given in 1.

   Inductive step: Assume

   $$\forall x_1 \dots \forall x_n \forall y_1 \dots \forall y_n \\ (\langle x_1, \dots, x_n\rangle = \langle y_1, \dots, y_n\rangle \leftrightarrow x_1 = y_1 \wedge \dots \wedge x_n = y_n)$$

   We have to show that

   $$\forall x_1 \dots \forall x_{n+1} \forall y_1 \dots \forall y_{n+1} \\ (\langle x_1, \dots, x_{n+1}\rangle = \langle y_1, \dots, y_{n+1}\rangle \leftrightarrow x_1 = y_1 \wedge \dots \wedge x_{n+1} = y_{n+1})$$

For the left-to-right direction, assume $\langle x_1, \ldots, x_{n+1} \rangle = \langle y_1, \ldots, y_{n+1} \rangle$. By Def. of ordered tuples, $\langle \langle x_1, \ldots, x_n \rangle, x_{n+1} \rangle = \langle \langle y_1, \ldots, y_n \rangle, y_{n+1} \rangle$. By 1, we have that $\langle x_1, \ldots, x_n \rangle = \langle y_1, \ldots, y_n \rangle$ and $x_{n+1} = y_{n+1}$. And, by inductive hypothesis, $x_1 = y_1 \wedge \cdots \wedge x_n = y_n$. The right-to-left direction is left as an exercise.

$\square$

Axiom (Union)**.** $\forall x \exists y \forall z\, (z \in y \leftrightarrow \exists w\, (w \in x \wedge z \in w))$

Gloss: For every set of sets x there exists a set y whose elements are exactly the members each set in x. Since this set is unique, we can introduce the following definition.

Definition 3.8 (large union)**.** $\forall x \bigcup x = \{y : \exists z\, (z \in x \wedge y \in z)\}$.

Example 3.6 (large unions)**.**

- $\bigcup \{\{1\}, \{2\}, \{3\}\} = \{1, 2, 3\}$
- $\bigcup \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\} = \mathbb{N} \cup \mathbb{Z} \cup \mathbb{Q} \cup \mathbb{R} = \mathbb{R}$

Exercise 3.18. *Indicate the value of the following operations:*

1. $\bigcup \{\{0, 1\}, \{1, 2\}\}$
2. $\bigcup \{\{0, 1\}, \{1, 2\}, \{\varnothing\}\}$
3. $\bigcup \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$

Exercise 3.19. *Show that* $\mathsf{ZF} \vdash$

1. $\forall x \forall y\, (x \in y \rightarrow x \subseteq \bigcup y)$
2. $\bigcup \varnothing = \varnothing$
3. $\forall x \bigcup \{x\} = x$
4. $\forall x \forall y\, (x \subseteq y \rightarrow \bigcup x \subseteq \bigcup y)$

Proposition 3.10 (union)**.** $\mathsf{ZF} \vdash \forall x \forall y \exists z \forall w\, (w \in z \leftrightarrow w \in x \vee w \in y)$.

Gloss: For any two sets x and y there is a set z of things that belong either to x or to y.

Proof. Let x and y be arbitrary sets. By Pairing, we have $\{x, y\}$. By Union, we can construct $\bigcup \{x, y\}$, which, by Def. of $\bigcup$, is the set of things that belong to a member of $\{x, y\}$, i.e. either to x or to y. Thus, $\forall w\, (w \in \bigcup \{x, y\} \leftrightarrow w \in x \vee w \in y)$, so $\exists z \forall w\, (w \in z \leftrightarrow w \in x \vee w \in y)$. $\square$

Thus, we can introduce the following definition:

Definition 3.9 (union)**.** $\forall x \forall y\, x \cup y = \{z : z \in x \vee z \in y\}$.

Example 3.7 (unions)**.**

- $\{0\} \cup \{1\} = \{0, 1\}$
- $\{x \in \mathbb{N} : x \text{ is even}\} \cup \{x \in \mathbb{N} : x \text{ is odd}\} = \mathbb{N}$
- $\{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\}$
- $\mathbb{Q} \cup \mathbb{R} = \mathbb{R}$

Exercise 3.20. *Indicate the value of the following operations:*

1. $\{0, 1\} \cup \{1, 2, 3\}$
2. $\{1, 2, 3\} \cup \{2, \{3\}, 4\}$
3. $\{\varnothing, \{\varnothing\}\} \cup \{\{\varnothing\}\}$
4. $\varnothing \cup \{\varnothing\}$
5. $\varnothing \cup \varnothing$
6. $\mathbb{Z} \cup \mathbb{Q}$
7. $\{x \in \mathbb{Z} : x < 0\} \cup \mathbb{N}$
8. $\{x \in \mathbb{Z} : x > 0\} \cup \{0\}$

Exercise 3.21. *Show that* $\mathsf{ZF} \vdash$

1. $\forall x \forall y\, x \cup y = \bigcup \{x, y\}$
2. $\forall x\, x \cup x = x$
3. $\forall x\, x \cup \varnothing = x$
4. $\forall x \forall y\, x \cup y = y \cup x$
5. $\forall x \forall y\, x \subseteq x \cup y$
6. $\forall x \forall y \forall z\, x \cup (y \cup z) = (x \cup y) \cup z$
7. $\forall x \forall y\, (x \subseteq y \to x \cup y = y)$
8. $\forall x \forall y \forall z\, (x \subseteq y \to x \cup z \subseteq y \cup z)$
9. $\forall x \forall y \forall z\, (x \subseteq z \wedge y \subseteq z \leftrightarrow x \cup y \subseteq z)$

Given Exercises 3.14.5, and 3.21.5 we may drop the brackets and simply write, e.g. $x \cup y \cup z$ and $x \cap y \cap z$.

Exercise 3.22 (Absorption)**.** *Show that* $\mathsf{ZF} \vdash$

1. $\forall x \forall y\, x \cap (x \cup y) = x$
2. $\forall x \forall y\, x \cup (x \cap y) = x$

Exercise 3.23 (Distribution)**.** *Show that* $\mathsf{ZF} \vdash$

1. $\forall x \forall y \forall z\, x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$
2. $\forall x \forall y \forall z\, x \cup (y \cap z) = (x \cup y) \cap (x \cup z)$

Axiom (Power set)**.** $\forall x \exists y \forall z\,(z \in y \leftrightarrow z \subseteq x)$

Gloss: For every set x there exists a set y whose elements are exactly the subsets of x. Since this set is unique, we can define the following operation:

Definition 3.10 (power set)**.** $\forall x\, \mathscr{P}(x) = \{y : y \subseteq x\}$.

Example 3.8 (power sets)**.**

- $\mathscr{P}(\{0\}) = \{\varnothing, \{0\}\}$
- $\mathscr{P}(\{0,1\}) = \{\varnothing, \{0\}, \{1\}, \{0,1\}\}$
- $\mathscr{P}(\varnothing) = \{\varnothing\}$

Exercise 3.24. *Indicate the value of the following operations.*

1. $\mathscr{P}(\{0,1,2\})$
2. $\mathscr{P}(\mathscr{P}(\varnothing))$
3. $\mathscr{P}(\varnothing) - \varnothing$
4. $\mathscr{P}(\varnothing) - \{\varnothing\}$
5. $\{\varnothing\} \cap \mathscr{P}(\{1\})$

Exercise 3.25. *Show that* $\mathsf{ZF} \vdash$

1. $\forall x\, x \in \mathscr{P}(x)$
2. $\forall x\, \varnothing \in \mathscr{P}(x)$
3. $\forall x \forall y\,(x \subseteq y \leftrightarrow \mathscr{P}(x) \subseteq \mathscr{P}(y))$
4. $\forall x \forall y\,(\mathscr{P}(x) = \mathscr{P}(y) \to x = y)$
5. $\forall x \forall y\,(\mathscr{P}(x) \in \mathscr{P}(y) \to x \in y)$
6. $\forall x \forall y\,(\mathscr{P}(x) \cap \mathscr{P}(y) = \mathscr{P}(x \cap y))$
7. $\forall x \forall y\,(\mathscr{P}(x) \cup \mathscr{P}(y) \subseteq \mathscr{P}(x \cup y))$
8. $\exists x \exists y\,(\mathscr{P}(x) \cup \mathscr{P}(y) \neq \mathscr{P}(x \cup y))$
9. $\forall x \bigcup \mathscr{P}(x) = x$
10. $\forall x \forall y\,(x \subseteq \mathscr{P}(y) \to \bigcup x \subseteq y)$

Proposition 3.11. *If a set a has n members, then $\mathscr{P}(a)$ has 2^n members.*

Proof. Given a subset x of a, for each element y of a there are 2 possibilities: either $y \in x$ or $y \notin x$. Thus, there are $\underbrace{2 \times \cdots \times 2}_{n}$ possible subsets of a. □

Proposition 3.12 (cartesian product)**.** $\mathsf{ZF} \vdash \forall x \forall y \exists z \forall w\,(w \in z \leftrightarrow \exists w_1 \exists w_2\,(w_1 \in x \wedge w_2 \in y \wedge w = \langle w_1, w_2 \rangle))$.

Gloss: For any sets x and y there is a set z whose elements are exactly all the ordered pairs $\langle w_1, w_2\rangle$ that can be formed taking the first component w_1 from x and the second component w_2 from y.

Proof. Let x and y be arbitrary sets. By Proposition 3.10, we have $x \cup y$. By Power Set, we have $\mathscr{P}(\mathscr{P}(x \cup y))$. Then, Separation entails the existence of the following set:

$$\{w \in \mathscr{P}\mathscr{P}(x \cup y) : \exists w_1 \exists w_2\, (w = \langle w_1, w_2\rangle \wedge w_1 \in x \wedge w_2 \in y)\}$$

□

Since for every x and y such z is unique, we can introduce the following definition:

Definition 3.11 (cartesian product)**.** $\mathsf{ZF} \vdash \forall x \forall y\, x \times y = \{\langle z, w\rangle : z \in x \wedge w \in y\}$.

We often write $x \times y \times z$ as short for $(x \times y) \times z$. We also write x^2 for $x \times x$, and x^{n+1} for $x^n \times x$. Note that x^2 is then the set of al ordered pairs of members of x, x^3 the set of all ordered triples and, in general, x^n the set of all n-tuples of members of x.

Example 3.9 (cartesian products)**.**

- $\{P, Q, R\} \times \{\mathrm{T}, \mathrm{F}\} = \{\langle P, \mathrm{T}\rangle, \langle P, \mathrm{F}\rangle, \langle Q, \mathrm{T}\rangle, \langle Q, \mathrm{F}\rangle, \langle R, \mathrm{T}\rangle, \langle R, \mathrm{F}\rangle\}$
- $\{1, 2\}^2 = \{\langle 1, 1\rangle, \langle 1, 2\rangle, \langle 2, 1\rangle, \langle 2, 2\rangle\}$
- $\{\varnothing\}^3 = \{\langle \varnothing, \varnothing, \varnothing\rangle\}$

Exercise 3.26. *Show that* $\mathsf{ZF} \vdash$

1. $\forall x\, x \times \varnothing = \varnothing \times x = \varnothing$
2. $\neg \forall x \forall y\, x \times y = y \times x$
3. $\forall x \forall y \forall z\, x \times (y \cup z) = (x \times y) \cup (x \times z)$
4. $\forall x \forall y \forall z\, x \times (y \cap z) = (x \times y) \cap (x \times z)$

Proposition 3.13. *If a and b are sets with m and n elements respectively, $a \times b$ has $m \times n$ elements.*

Proof. For each $x \in a$, which are m in total, there are n possible $y \in b$ s.t. $\langle x, y\rangle \in a \times b$. Thus, $a \times b$ has $m \times n$ elements. □

Axiom (Infinity)**.** $\exists x\, (\varnothing \in x \wedge \forall y\, (y \in x \rightarrow y \cup \{y\} \in x))$

Gloss: There's a set that contains all the (infinitely many) following sets as elements:

$$\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\{\varnothing, \{\varnothing\}\}\}\} \dots$$

Infinity entails the existence of an infinite set, usually known as ω. Sometimes its members, known as the 'von Neumann ordinals', are identified with the natural numbers according to their cardinality (number of elements) and, thus, ω is identified with $\mathbb{N}$.

In what follows, let $\forall v \in t\, \varphi$ be short for $\forall v\, (v \in t \rightarrow \varphi)$, and $\exists v \in t\, \varphi$ be short for $\exists v\, (v \in t \wedge \varphi)$.

Axiom (Replacement)**.** $\forall x\, (\forall y \in x \exists! z\, \varphi(y, z) \rightarrow \exists y \forall z\, (z \in y \leftrightarrow \exists w\, (w \in x \wedge \varphi(w, z))))$*, for each wff* $\varphi(y, z)$ *of* $\mathscr{L}_\in$

Gloss: If φ defines a (total) function from a set x (in the sense that for every member y of x there is a unique z s.t. $\varphi(y, z)$), the values of this function form a set. Note that, just like Separation, Replacement is also not an axiom but an axiom-schema.

Proposition 3.14 (set of singletons)**.** $\mathsf{ZF} \vdash \forall x \exists y \forall z\, (z \in y \leftrightarrow \exists w\, (w \in x \wedge z = \{w\}))$

Gloss: For every set x there's a set y whose members are exactly the singletons $\{w\}$ of members of x.

Proof. By Proposition 3.7, each y has a singleton, $\{y\}$. Since we have that

$$\forall y \in x \exists! z\, z = \{y\}$$

by Replacement (where $\varphi(y, z)$ is instantiated by $z = \{y\}$), we can infer that

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

□

Proposition 3.15. *Empty Set, Replacement* $\vdash$ *Separation.*

Proof. We want to show that

$$\forall w \exists x \forall y\, (y \in x \leftrightarrow y \in w \wedge \varphi(y))$$

We omit parameters $\vec{z}$ for perspicuity, as they play no role in the proof. There are two possibilities: either $\neg\exists y\, (y \in w \wedge \varphi(y))$ or $\exists y\, (y \in w \wedge \varphi(y))$.

If $\neg\exists y\, (y \in w \wedge \varphi(y))$, then $x = \varnothing$ satisfies the condition, that is,

$$\forall y\, (y \in \varnothing \leftrightarrow y \in w \wedge \varphi(y))$$

Thus,

$$\exists x \forall y\, (y \in x \leftrightarrow y \in w \wedge \varphi(y))$$

which entails the corresponding instance of Separation.

Assume now that $\exists y\,(y \in w \wedge \varphi(y))$ and let $y_1 \in w \wedge \varphi(y_1)$. Consider the following formula:

$$((y \in w \wedge \varphi(y)) \wedge z = y) \vee (\neg(y \in w \wedge \varphi(y)) \wedge z = y_1) \tag{3.1}$$

(3.1) defines the function that maps every member of w satisfying φ to itself and all the others to y_1. It can be easily shown that (3.1) satisfies existence and uniqueness. Thus, the instance of Replacement for (3.1) entails the existence of the set x of members y of w s.t. $\varphi(y)$, as desired. □

Proposition 3.15 shows that Separation is not an *independent axiom*, it's redundant, so we can do away with it.

Axiom (Regularity/Foundation)**.** $\forall x\,(x \neq \varnothing \to \exists y \in x\, x \cap y = \varnothing)$

Gloss: Every non-empty set is disjoint from at least one of its elements.

Theorem 3.2 (No (finite) circles)**.** $\mathsf{ZF} \vdash$

1. $\forall x\, x \notin x$

2. $\neg\exists x_1 \ldots \exists x_n\,(x_1 \in x_2 \wedge x_2 \in x_3 \wedge \cdots \wedge x_{n-1} \in x_n \wedge x_n \in x_1)$

Proof. 1. Assume $x \in x$, so $x \in x \cap \{x\}$. Since $\{x\} \neq \varnothing$, Regularity implies

$$\exists y\,(y \in \{x\} \wedge y \cap \{x\} = \varnothing)$$

Since the only member of $\{x\}$ is x, we have that $x \cap \{x\} = \varnothing$, which contradicts our assumption.

2. Assume $x_1 \in x_2 \wedge x_2 \in x_3 \wedge \cdots \wedge x_{n-1} \in x_n \wedge x_n \in x_1$. By Regularity,

$$\exists y(y \in \{x_1, \ldots, x_n\} \wedge y \cap \{x_1, \ldots, x_n\} = \varnothing)$$

Again, this contradicts our assumption.

□

This completes the presentation of ZF. Usually, set theorists work with the following additional axiom (not derivable in ZF), and call the extended theory ZFC:

Axiom (Choice)**.** $\forall x\,(\forall y \in x\, y \neq \varnothing \to \exists z \forall y \in x \exists! w \in y\, w \in z)$

Gloss: Given any collection x of non-empty sets y, there exists a set z containing exactly one object w from each y.

Choice is not always included among the axioms of set theory, as it has some controversial consequences, such as the *Banach-Tarski Paradox*: every solid sphere can be decomposed into a finite number of disjoint 'pieces' (subsets), which can then be put back together in a different way to yield two identical copies of the original sphere.

Chapter 4

Relations and Functions

In this chapter we will introduce relations as a special kind of set and functions as a special kind of relation. We will often use constants (in the metalanguage) to name relations and functions. We will mostly use R for relations and $f, g, \ldots$ for functions, sometimes with subscripts.

4.1 Relations

Definition 4.1 (relation)**.** R is a relation from a set a to a set b iff $R \subseteq a \times b$. We call a the *set of departure* and b the *set of destination* of R.

We usually write Rxy or $R(x, y)$ for $\langle x, y\rangle \in R$.

Example 4.1 (relations)**.**

- $R_1 = \{\langle 0, 1\rangle, \langle 0, 2\rangle, \langle 1, 2\rangle\}$ is a relation from $\{0, 1, 2\}$ to itself, from $\{0, 1, 2\}$ to $\mathbb{N}$, from $\{0, 1, 2, 3\}$ to itself, etc.
- $R_2 = \{\langle P, \mathrm{F}\rangle, \langle Q, \mathrm{T}\rangle, \langle R, \mathrm{F}\rangle\}$ is a relation from $\{P, Q, R\}$ to $\{\mathrm{T}, \mathrm{F}\}$.
- $R_3 = \{\langle x, x\rangle : x \in \mathbb{N}\}$ is a relation *on* $\mathbb{N}$.
- $R_4 = \{1, 2\}^2$ is a relation on $\{1, 2\}$.
- $R_5 = \{\langle x, y\rangle : x$ and y are human beings and x is an ancestor of $y\}$ is a relation on the set of human beings.

Definition 4.2 (domain and range)**.** Let $R \subseteq a \times b$.
$Dom(R) = \{x \in a : \exists y \in b\, Rxy\}$
$Ran(R) = \{x \in b : \exists y \in a\, Ryx\}$

Example 4.2 (domains and ranges)**.**

- $Dom(R_1) = \{0, 1\}$, $Ran(R_1) = \{1, 2\}$.
- $Dom(R_2) = \{P, Q, R\}$, $Ran(R_2) = \{\mathrm{T}, \mathrm{F}\}$.

- $Dom(R_3) = \mathbb{N}$, $Ran(R_3) = \mathbb{N}$.
- $Dom(R_4) = \{1, 2\}$, $Ran(R_4) = \{1, 2\}$.

Exercise 4.1. *What is the domain and the range of R_5?*

Exercise 4.2. *Let $R = \{\langle 0,1\rangle, \langle 0,3\rangle, \langle 0,4\rangle, \langle 2,1\rangle, \langle 1,2\rangle, \langle 4,7\rangle\}$. Provide a set of departure and a set of destination, and determine $Dom(R)$, $Ran(R)$.*

Exercise 4.3. *Is $\varnothing$ a relation? If so, provide a set of departure and a set of destination?*

Definition 4.3 (n-place relation)**.** If $n > 2$, R is an n-place relation iff there are sets $a_1, \ldots, a_n$ s.t. $R \subseteq a_1 \times a_2 \times \cdots \times a_n$. R's set of departure is $a_1 \times a_2 \times \cdots \times a_{n-1}$ and its set of destination is a_n. If $R \subseteq a^n$ for some set a, we say that R is an n-place relation on a.

Exercise 4.4. *Let R be an n-place relation on the set a. Provide a set of departure and a set of destination for R.*

Example 4.3 (n-place relations)**.**

- $\{\langle x, y, z\rangle \in \mathbb{N}^3 : x < y < z\}$ is a ternary relation on $\mathbb{N}$.
- $\{\langle x, y, z, w\rangle \in \mathbb{R}^4 : (x \times y) + z = w\}$ is a 4-place relation on $\mathbb{R}$.

In what follows, let $\forall v_1, \ldots, v_n \in t\, \varphi$ abbreviate $\forall v_1 \in t \ldots \forall v_n \in t\, \varphi$.

Definition 4.4 (equivalence relation)**.** R is an equivalence relation on a set a iff $R \subseteq a^2$ and the following conditions hold:

- Reflexivity. $\forall x \in a\, Rxx$
- Symmetry. $\forall x, y \in a\, (Rxy \to Ryx)$
- Transitivity. $\forall x, y, z \in a\, (Rxy \wedge Ryz \to Rxz)$

Note that whether a set is an equivalence relation depends on what set it's a relation on.

Example 4.4 (equivalence relations)**.**

- $\{\langle x, x\rangle : x \in a\}$, on a^2
- $\{\langle x, y\rangle : x$ and y are human beings of the same age$\}$, on the set of human beings
- $\{\langle x, y\rangle : x$ and y are logically equivalent sentences of $\mathscr{L}_{\mathsf{PA}}\}$, on the set of sentences of $\mathscr{L}_{\mathsf{PA}}$
- $\{\langle x, y\rangle : x, y \subseteq \mathbb{N}$ and they have the same number of elements$\}$, on $\mathscr{P}(\mathbb{N})$

Example 4.5 (non-equivalence relations)**.**

- $\{\langle x, y\rangle \in \mathbb{Z}^2 : x \neq y\}$, on $\mathbb{Z}$
- $\{\langle x, y\rangle : x$ and y are human beings and are friends$\}$, on the set of human beings
- $\{\langle x, y\rangle \in \mathbb{N}^2 : x \leqslant y\}$, on $\mathbb{N}$

Exercise 4.5. *Show that the relations in Example 4.4 are equivalence relations.*

Exercise 4.6. *Show that the relations in Example 4.5 are not equivalence relations.*

Exercise 4.7. *Show that the relations in Example 4.4 are equivalence relations.*

Definition 4.5 (equivalence class)**.** Let R be an equivalence relation on a. For each $x \in a$, the equivalence class of x – which we note $[x]_R$ – is the set of things in a x relates to, i.e. $[x]_R = \{y \in a : Rxy\}$.

Definition 4.6 (partition)**.** A set b is a partition of a set a iff the following holds:

1. $b \subseteq \mathscr{P}(a)$
2. $\bigcup b = a$
3. $\forall x, y \in b\,(x = y \vee x \cap y = \varnothing)$

Gloss: b is a partition of a if it's a set of disjoint non-empty subsets of a, the union of which is a itself.

Proposition 4.1. *Let R be an equivalence relation on a. Then $\{[x]_R : x \in a\}$ is a partition of a.*

Proof. By Def. of equivalence class, all members of $\{[x]_R : x \in a\}$ are subsets of a. Thus, $\{[x]_R : x \in a\} \subseteq \mathscr{P}(a)$. By Exercise 3.25.10, we also have that

$$\bigcup\{[x]_R : x \in a\} \subseteq a \tag{4.1}$$

Assume now that $x \in a$. Then, $[x]_R \in \{[x]_R : x \in a\}$. Since $x \in [x]_R$ by reflexivity, by Def. of $\bigcup$, $x \in \bigcup\{[x]_R : x \in a\}$. Since x was arbitrary, we proved that $\forall x\,(x \in a \rightarrow x \in \bigcup\{[x]_R : x \in a\})$, i.e.

$$a \subseteq \bigcup\{[x]_R : x \in a\} \tag{4.2}$$

By Proposition 3.4.3, (4.1) and (4.2) imply that $\bigcup\{[x]_R : x \in a\} = a$, that is, condition 2 of the Def. of partition. Condition 3 is left as an exercise. $\square$

Exercise 4.8. *Finish the proof of Proposition 4.1.*

4.2 Functions

Definition 4.7 (total function). f is a (total) function from a set a to a set b iff f is a relation from a to b – i.e. if $f \subseteq a \times b$ – and satisfies the following conditions:

- Existence. $\forall x \in a \exists y \in b \langle x, y\rangle \in f$
- Uniqueness. $\forall x \in a \forall y, z \in b\,(\langle x, y\rangle \in f \wedge \langle x, z\rangle \in f \rightarrow y = z)$

Gloss: If f is a function from a to b then, by existence, every member of a is related to at least one member of b and, by uniqueness, every member of a is related to at most one member of b. Thus, every member x of a is related to exactly one member y of b. For each $x \in a$, the unique $y \in b$ s.t. $\langle x, y\rangle \in f$ is the *image* of x under f. We write $f(x) = y$ as short for $\langle x, y\rangle \in f$ and $f : a \rightarrow b$ to indicate that f is a function from a to b.

Note that the set of departure of a function is also its domain, by existence. The set of destination is also known as the *codomain* of the function. It only makes sense to talk about sets as functions if we have specified the departure and destination sets as well. Functions are not just sets of ordered pairs but come with a domain and a codomain.

Example 4.6 (functions).

- $f_1 : \{x : x \text{ is a human being}\} \rightarrow \mathbb{N}$ s.t. $f_1 = \{\langle x, y\rangle : x \text{ is a human being and } y \text{ is their age}\}$
- $f_2 : \mathbb{N} \rightarrow \mathbb{N}$ s.t. $f_2 = \{\langle x, y\rangle \in \mathbb{N}^2 : y = x + 1\}$. We write $f_2(x) = x + 1$ or $f_2 : x \mapsto x + 1$.
- $f_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ s.t. $f_3 = \{\langle x, y, z\rangle \in \mathbb{R}^3 : z = x \times y\}$. We write $f_3(x, y) = x \times y$ or $f_3 : x, y \mapsto x \times y$.[1]

Example 4.7 (non-functions).

- $\{\langle x, y\rangle : x, y \text{ are human beings and } y \text{ is a sister of } x\}$, on the set of human beings
- $\{\langle x, y\rangle \in \mathbb{N}^2 : y = x/2\}$, on $\mathbb{N}$
- $\{\langle x, y\rangle \in \mathbb{R}^2 : y = \sqrt{x}\}$, on $\mathbb{R}$

Exercise 4.9. *Show that the relations in Example 4.6 are functions.*

Exercise 4.10. *Show that the relations in Example 4.7 are not functions.*

Exercise 4.11. *Indicate which of the following relations can be functions if appropriate sets of departure are specified.*

[1] In general, for any n-place function f we omit the ordered n-tuple brackets and write $f(x_1, \ldots, x_n)$ instead of $f(\langle x_1, \ldots, x_n\rangle)$.

1. $\{\langle P, \mathrm{F}\rangle, \langle Q, \mathrm{F}\rangle\}$
2. $\{\langle a, b\rangle, \langle b, a\rangle\}$
3. $\{\langle a, a\rangle, \langle b, b\rangle\}$
4. $\{\langle a, a\rangle, \langle a, b\rangle, \langle b, a\rangle, \langle b, b\rangle\}$
5. $\{\langle 0,0\rangle, \langle 1,2\rangle, \langle 2,4\rangle, \langle 3,6\rangle, \langle 4,8\rangle, \langle 5,10\rangle\}$
6. $\{\langle 0,0\rangle, \langle 1,0\rangle, \langle 2,0\rangle, \langle 3,0\rangle, \langle 4,0\rangle, \langle 5,0\rangle\}$
7. $\{\langle 0,0\rangle, \langle 0,1\rangle, \langle 0,2\rangle, \langle 0,3\rangle, \langle 0,4\rangle, \langle 0,5\rangle\}$
8. $\{\langle x, y\rangle : x \text{ and } y \text{ are neighbouring European countries}\}$
9. $\{\langle x, y, z\rangle \in \mathbb{N}^3 : x < y < z < 5\}$
10. $\{\langle x, y, z\rangle : x, y, \text{ and } z \text{ are planets and } y \text{ orbits between } x \text{ and } z\}$
11. $\{\langle x, y, z\rangle \in \mathbb{N}^3 : x, y, z < 10 \wedge z = x/y\}$

Exercise 4.12. *Let a be a non-empty set. Is $\varnothing$ a function from $\varnothing$ to a? Is it a function from a to $\varnothing$?*

Definition 4.8 (partial function)**.** f is a partial function from a to b iff f is a relation from a to b and satisfies uniqueness.

Example 4.8 (partial functions)**.**

- $\{\langle x, y\rangle \in \mathbb{Z}^2 : y = x/2\}$, on $\mathbb{Z}$
- $\{\langle x, y\rangle \in \mathbb{N}^2 : y = \sqrt{x}\}$, on $\mathbb{N}$

Exercise 4.13. *Show that the relations in Example 4.8 are partial functions.*

Definition 4.9 (n-place function)**.** f is an n-place function iff there are sets $a_1, \ldots, a_n$ and b s.t. $f : a_1 \times \cdots \times a_n \to b$.

Example 4.9 (n-place functions)**.**

- f_3 is a 2-place function
- $f_4 : \mathbb{Z}^3 \to \mathbb{Z}$ s.t. $f_4(x, y, z) = x + y + z$ is a 3-place function

Definition 4.10 (1-1/injective function)**.** $f : a \to b$ is 1-1 iff $\forall x, y \in a\, (f(x) = f(y) \to x = y)$.

Gloss: No two members of a are mapped to the same member of b.

Definition 4.11 (onto/surjective function)**.** $f : a \to b$ is onto iff $\forall x \in b \exists y \in a\, f(y) = x$.

Gloss: Every member of b is the image of some member of a.

Definition 4.12 (bijective function)**.** A function is bijective iff it's both 1-1 and onto.

Example 4.10 (types of functions)**.**

- $\{\langle x, y\rangle : x$ is a human being and y is their age$\}$, from the set of human beings to $\mathbb{N}$ is neither injective nor surjective
- $f_5 : \mathbb{N} \to \mathbb{N}$ s.t. $f_5(x) = x + 1$ is injective but not surjective
- $f_6 : \mathbb{N}^2 \to \mathbb{N}$ s.t. $f_6(x, y) = x \times y$ is surjective but not injective
- $f_7 : \mathbb{N} \to \mathbb{N} - \{0\}$ s.t. $f_7(x) = x + 1$ is bijective
- $f_8 : \mathbb{R} \to \mathbb{R}$ s.t. $f_8(x) = 2x$ is bijective

Exercise 4.14. *Decide if the following functions are injective, surjective, or bijective:*

- $g_1 : \mathbb{Z} \to \mathbb{Z}$ *s.t.* $g_1(x) = x^2$
- $g_2 : \mathbb{Z} \to \mathbb{N}$ *s.t.* $g_2(x) = x^2$
- $g_3 : \mathbb{N} \to \mathbb{Z}$ *s.t.* $g_3(x) = x^2$
- $g_4 : \mathbb{N} \to \mathbb{N}$ *s.t.* $g_4(x) = x^2$

Exercise 4.15. *Let R be an equivalence relation on a and let $f : a \to \{[x]_R : x \in a\}$ be s.t. $f(x) = [x]_R$. Show that f is surjective. Under which conditions is it bijective?*

Exercise 4.16. *Show that, if $f : a \to b$ is injective, then $f' : a \to Ran(f)$ s.t. $\forall x \in a\, f'(x) = f(x)$ is bijective.*

Definition 4.13 (inverse function)**.** Let f be 1-1. The inverse function of f is $f^{-1} : Ran(f) \to Dom(f)$ s.t. $\forall x \in Ran(f)\forall y \in Dom(f)\, (f^{-1}(x) = y \leftrightarrow f(y) = x)$.

Example 4.11 (inverse functions.)**.**

- $f_5^{-1} : \mathbb{N} - \{0\} \to \mathbb{N}$ is s.t. $f_5^{-1}(x) = x - 1$
- $f_7^{-1} : \mathbb{N} - \{0\} \to \mathbb{N}$ is s.t. $f_7^{-1}(x) = x - 1$
- $f_8^{-1} : \mathbb{R} \to \mathbb{R}$ is s.t. $f_8^{-1}(x) = x/2$

Exercise 4.17. *Why does f have to be 1-1 for it to have an inverse function?*

Exercise 4.18. *Show that, if f is 1-1, then f^{-1} is bijective.*

Definition 4.14 (composition)**.** Let f and g be s.t. $Ran(f) \subseteq Dom(g)$. The composition function of g with f is $g \circ f : Dom(f) \to CoDom(g)$ s.t. $g \circ f(x) = g(f(x))$.[2]

[2] Rather obviously, $CoDom(g)$ is the codomain of g.

Example 4.12 (compositions)**.** Let $f : \mathbb{N} \to \mathbb{N}$ s.t. $f(x) = 2x$ and $g : \mathbb{N} \to \mathbb{N}$ s.t. $g(x) = x + 1$.

- $g \circ f : \mathbb{N} \to \mathbb{N}$ is s.t. $g \circ f(x) = 2x + 1$.
- $f \circ g : \mathbb{N} \to \mathbb{N}$ is s.t. $f \circ g(x) = 2(x + 1)$.

Exercise 4.19. *Why is it necessary that $Ran(f) \subseteq Dom(g)$ for $g \circ f$ to be a function?*

Exercise 4.20. *Show that the following claims hold:*

1. *If both f and g are injective and $Ran(f) \subseteq Dom(g)$, then $g \circ f$ is also injective.*
2. *If both f and g are surjective and $Ran(f) = Dom(g)$, then $g \circ f$ is also surjective.*
3. *If f is bijective, then $\forall x \in Dom(f)\, f^{-1} \circ f(x) = x$ and $\forall x \in Ran(f)\, f \circ f^{-1}(x) = x$.*
4. *If $Ran(f) \subseteq Dom(g)$, and $Ran(g) \subseteq Dom(h)$, then $h \circ (g \circ f) = (h \circ g) \circ f$.*

Chapter 5

The size of sets

The size or *cardinality* of a finite set is simply the number of elements in the set. What is the size of an infinite set? Is it just 'infinite' or are there infinities of different sizes? In this chapter we will see that the answer to the latter question is 'yes', there are infinitely many different sizes of infinities. How many?

5.1 Enumerablility

An *enumerable* (or *countable*) set is one whose members can be enumerated or counted, arranged in a single list with a first entry, a second entry, and so on, so that every member of the set appears sooner or later on the list as the n-th entry, for some number n. This doesn't mean that a member can't appear more than once in the list.

Example 5.1 (enumerations/lists)**.**

- $\mathbb{N}$ is enumerated by the list $0, 1, 2, 3, \ldots$
- $\mathbb{Z}^+$ is enumerated by the list $1, 2, 3, \ldots$
- $\mathbb{Z}^-$ is enumerated by the list $-1, -2, -3, \ldots$
- $\{x \in \mathbb{N} : x \text{ is odd }\}$ is enumerated by the list $1, 3, 5, \ldots$
- $\{x \in \mathbb{N} : x < 3\}$ is enumerated by the list $0, 1, 2, 0, 1, 2, \ldots$
- $\mathbb{N}$ is enumerated by the list $1, 0, 3, 4, 5, \ldots$
- $\{x \in \mathbb{N} : x \text{ is odd }\}$ is enumerated by the list $1, 3, 3, 5, \ldots$
- $\mathbb{Z}^+$ isn't enumerated by $2, 4, 6, \ldots 1, 3, 5, \ldots$, as e.g. there is no number n s.t. 1 appears on the list as the n-th entry
- $\mathbb{Z}^-$ isn't enumerated by $\ldots, -3, -2, -1$, as e.g. there is no number n s.t. -3 appears on the list as the n-th entry

- $\mathbb{N}$ isn't enumerated by $1, 2, 3, \ldots, 0$, as there is no number n s.t. 0 appears on the list as the n-th entry

Lists are too imprecise, though. For an infinite set, we can write a few entries and only hint at all the rest. Our list-talk can be replaced with the more precise function-talk.

Definition 5.1 (enumerability)**.** A set a is enumerable iff it is either $\varnothing$ or there is a surjective function $f : \mathbb{Z}^+ \to a$.

We include $\varnothing$ by courtesy. If a is non-empty, we say that f enumerates a.

Example 5.2 (enumerations/functions)**.**

- $\mathbb{N}$ is enumerated by $f : \mathbb{Z}^+ \to \mathbb{N}$ s.t. $f(x) = x - 1$
- $\mathbb{Z}^+$ is enumerated by $g : \mathbb{Z}^+ \to \mathbb{Z}^+$ s.t. $g(x) = x$
- $\mathbb{Z}^-$ is enumerated by $h : \mathbb{Z}^+ \to \mathbb{Z}^-$ s.t. $h(x) = -x$
- $\{x \in \mathbb{N} : x \text{ is odd }\}$ is enumerated by $j : \mathbb{Z}^+ \to \{x \in \mathbb{N} : x \text{ is odd }\}$ s.t. $j(x) = 2x - 1$

If a set is enumerable and infinite, we say it's *denumerable.*

Theorem 5.1 (Cantor's zig-zag)**.** $\mathbb{Z}^{+2}$ *(the set of ordered pairs of positive integers) is enumerable.*

Proof. We can enumerate the ordered pairs following Cantor's zig-zag diagram:

$$\begin{tikzcd}
\langle 1,1\rangle \arrow[r] & \langle 1,2\rangle \arrow[dl] & \langle 1,3\rangle \arrow[dl] & \langle 1,4\rangle \arrow[dl] & {} \\
\langle 2,1\rangle \arrow[urr] & \langle 2,2\rangle \arrow[dl] & \langle 2,3\rangle \arrow[dl] & & \\
\langle 3,1\rangle \arrow[uurrr] & \langle 3,2\rangle \arrow[dl] & & & \\
\langle 4,1\rangle \arrow[uuurrrr, dotted] & & & &
\end{tikzcd}$$

That is, we have the function $zz : \mathbb{Z}^+ \to \mathbb{Z}^{+2}$ s.t.

- $zz(1) = \langle 1, 1\rangle$ and
- $zz(x+1) = \begin{cases} \langle id_1^2(zz(x)) + 1, id_2^2(zz(x)) - 1\rangle & \text{if } id_2^2(zz(x)) \neq 1 \\ \langle 1, id_1^2(zz(x)) + 1\rangle & \text{otherwise} \end{cases}$

where $id_1^2 : \mathbb{Z}^2 \to \mathbb{Z}$ is s.t. $id_1^2(x, y) = x$ and $id_2^2 : \mathbb{Z}^2 \to \mathbb{Z}$ is s.t. $id_2^2(x, y) = y$.

We show that zz is surjective, that is, that every ordered pair of positive integers $\langle x, y\rangle$ is the image of some positive integer z. We do it by proving the following by total induction on w:

$$\forall w, x, y \in \mathbb{Z}^+ \, (w = x + y \to \exists z \in \mathbb{Z}^+ \, zz(z) = \langle x, y\rangle)$$

In plain words, for each positive integer $w \geqslant 2$, all ordered pairs $\langle x, y\rangle \in \mathbb{Z}^{+2}$ s.t. $x+y = w$ are in the range of zz. This is a form of total induction by proxy for ordered pairs of positive integers: since every ordered pair $\langle x, y\rangle$ is s.t. $x+y$ is a positive integer $\geqslant 2$, since every ordered pair is associated with a positive integer in this way, we are proving that every ordered pair of positive integers is in the range of zz.

Base step: $w = 2$. Then $x = 1$ and $y = 1$, so our z is 1, as $zz(1) = \langle x, y\rangle$.

Inductive step: Assume that, for every $w' < w$, if $x' + y' = w'$ then there's a z' s.t. $zz(z') = \langle x', y'\rangle$, and let $x + y = w$. We want to show that there's a z s.t. $zz(z) = \langle x, y\rangle$. We prove it by (a nested/sub-)induction on x:

- *(Sub-)Base step:* $x = 1$. Note that $(y-1)+1 < x+y = w$. By inductive hypothesis, there's a z' s.t. $zz(z') = \langle y-1, 1\rangle$, so $zz(z'+1) = \langle 1, y\rangle = \langle x, y\rangle$.
- *(Sub-)Inductive step:* Assume that, for every $x' < x$, if $x' + y' = w$ then there's a z' s.t. $zz(z') = \langle x', y'\rangle$. Thus, there's a z' s.t. $zz(z') = \langle x-1, y+1\rangle$, so $zz(z'+1) = \langle x, y\rangle$.

Another way to see that zz is surjective is to note that the function given by $1+2+\cdots+(x+y-2)+x$ is the inverse function of zz, i.e. $zz^{-1} : \mathbb{Z}^{+2} \to \mathbb{Z}^+$ is s.t.

$$zz^{-1}(x, y) = \frac{(x+y-2)(x+y-1)}{2+x}$$

□

Example 5.3. The set of positive rational numbers is enumerable.

$f : \mathbb{Z}^{+2} \to \mathbb{Q}^+$ s.t. $f(x, y) = \frac{x}{y}$ is surjective, as every positive rational number can be expressed as a ratio of positive integers. By Theorem 5.1, $zz : \mathbb{Z}^+ \to \mathbb{Z}^{+2}$ is surjective. By Exercise 4.20, we know that $f \circ zz : \mathbb{Z}^+ \to \mathbb{Q}^+$ is surjective, which means that $\mathbb{Q}^+$ is enumerable.

$$f \circ zz : x \mapsto \langle y, z\rangle \mapsto \frac{y}{z}$$

Exercise 5.1. *Show that the set of ordered pairs of natural numbers is enumerable.*

Proposition 5.1. *The union of two enumerable sets is enumerable.*

Proof. Let a and b be enumerable sets. If one of them is $\varnothing$, the result follows trivially. If none of them is, there are surjective functions $f_a : \mathbb{Z}^+ \to a$ and $f_b : \mathbb{Z}^+ \to b$ enumerating them. Thus, the following function $g : \mathbb{Z}^+ \to a \cup b$ is also surjective:

$$g(x) = \begin{cases} f_a(\frac{x+1}{2}) & \text{if } x \text{ is odd} \\ f_b(\frac{x}{2}) & \text{if } x \text{ is even} \end{cases}$$

Therefore, $a \cup b$ is enumerable. □

Proposition 5.2. *All finite sets are enumerable.*

Proof. Let a be finite and let $a_1, \ldots, a_n$ be its elements. Then, $f : \mathbb{Z}^+ \to a$ s.t.

$$f(x) = \begin{cases} a_i & \text{if } x = i \\ a_n & \text{otherwise} \end{cases}$$

is surjective, so a is enumerable. □

Example 5.4. The set of rational numbers is enumerable.
By Example 5.3 we know there's a surjective function $f : \mathbb{Z}^+ \to \mathbb{Q}^+$. Then, $f' : \mathbb{Z}^+ \to \mathbb{Q}^-$ s.t. $f'(x) = -f(x)$ is also surjective, which shows that $\mathbb{Q}^-$ is enumerable. By Proposition 5.1, $\mathbb{Q}^+ \cup \mathbb{Q}^-$ is enumerable too. Since $\{0\}$ is finite, by Proposition 5.2 it is enumerable, so $\mathbb{Q}^+ \cup \mathbb{Q}^- \cup \{0\}$, i.e. $\mathbb{Q}$, is also enumerable.

Example 5.5. The set of ordered triples of positive integers is enumerable.
Since zz is surjective, by Exercise 4.20 we then know that $f : \mathbb{Z}^+ \to \mathbb{Z}^{+3}$ s.t.

$$f(x) = \langle zz(id^2_1(zz(x))), id^2_2(zz(x)) \rangle$$

is surjective too.

$$f : x \mapsto \langle y_1, y_2 \rangle \mapsto \langle \langle z_1, z_2 \rangle, y_2 \rangle$$

Example 5.6. The set of ordered n-tuples of positive integers, for any fixed n, is enumerable.
This follows by induction on n.
Base step: $n = 2$. The result follows by Theorem 5.1.
Inductive step: Assume the set of ordered n-tuples of positive integers is enumerable. Then, there is a surjective $f : \mathbb{Z}^+ \to \mathbb{Z}^{+n}$. By Exercise 4.20, we then know that $g : \mathbb{Z}^+ \to \mathbb{Z}^{+n+1}$ s.t. $g(x) = \langle f(id^2_1(zz(x))), id^2_2(zz(x)) \rangle$ is also surjective.

$$f : x \mapsto \langle y_1, y_2 \rangle \mapsto \langle \langle z_1, \ldots, z_n \rangle, y_2 \rangle$$

Exercise 5.2. *Show that the set of ordered n-tuples of natural numbers is enumerable.*

Proposition 5.3. *Every subset of an enumerable set is enumerable.*

Proof. Let a be enumerable and let $b \subseteq a$. If $b = \varnothing$, b is trivially enumerable. If $b \neq \varnothing$, let $b_0 \in b$. Since $a \neq \varnothing$, there is a surjective function $f : \mathbb{Z}^+ \to a$. Thus, $g : \mathbb{Z}^+ \to b$ s.t.

$$g(x) = \begin{cases} f(x) & \text{if } f(x) \in b \\ b_0 & \text{otherise} \end{cases}$$

is also surjective, so b is enumerable. □

Proposition 5.4. *The union of enumerably many enumerable sets is enumerable.*

Proof. Let a be an enumerable set, all of whose elements are themselves enumerable sets. Thus, there is a surjective function $f : \mathbb{Z}^+ \to a$ s.t. for every $x \in \mathbb{Z}^+$, $f(x)$ is enumerable. Therefore, for every $x \in \mathbb{Z}^+$, if $f(x) \neq \varnothing$, there is a surjective function $g_x : \mathbb{Z}^+ \to f(x)$. By Exercise 4.20, $h : \mathbb{Z}^+ \to \bigcup a$ s.t. $h(x) = g_{id_1^2(zz(x))}(id_2^2(zz(x)))$ is also surjective.

$$h : x \mapsto \langle y_1, y_2 \rangle \mapsto g_{y_1}(y_2)$$

□

Example 5.7. The set of finite sequences of positive integers is enumerable.

To see it, note that the set of finite sequences of positive integers is $\bigcup\{\mathbb{Z}^{+n} : n \in \mathbb{Z}^+\}$. Since $\{\mathbb{Z}^{+n} : n \in \mathbb{Z}^+\}$ is enumerable and, by Example 5.6, we know that all of its members are too, by Proposition 5.4, $\bigcup\{\mathbb{Z}^{+n} : n \in \mathbb{Z}^+\}$ is enumerable too.

Exercise 5.3. *Show that the set of finite sequences of natural numbers is enumerable.*

Example 5.8. The set of finite strings from an enumerable alphabet of symbols is enumerable.

To see it, let a be an enumerable set of symbols. We assume $a \neq \varnothing$; otherwise the result is trivial. Thus, there's a surjective function $f : \mathbb{Z}^+ \to a$. By Example 5.7, there's also a surjective function $g : \mathbb{Z}^+ \to \bigcup\{\mathbb{Z}^{+n} : n \in \mathbb{Z}^+\}$. Let b be the set of finite strings of symbols in a. Then, $h : \mathbb{Z}^+ \to b$ s.t. $h(x) = a_0 \dots a_n$ iff $g(x) = \langle y_1, \dots, y_n \rangle$ and $f(y_i) = a_i$, for each $1 \leqslant i \leqslant n$, is also surjective.

$$h : x \mapsto \langle y_1, \dots, y_n \rangle \mapsto a_1 \dots a_n$$

Exercise 5.4. *Show that the set of all terms of a first-order language $\mathscr{L}$ is always denumerable.*

Exercise 5.5. *Show that the set of all wffs of a first-order language $\mathscr{L}$ is always denumerable.*

Example 5.9. The set of finite sets of positive integers is enumerable.

To see it, note that, by Example 5.7, there's a surjective function $f : \mathbb{Z}^+ \to \bigcup\{\mathbb{Z}^{+n} : n \in \mathbb{Z}^+\}$. Then, $g : \mathbb{Z}^+ \to \{x \in \mathscr{P}(\mathbb{Z}^+) : x \text{ is finite} \wedge x \neq \varnothing\}$ s.t. $g(x) = \{y_1, \ldots, y_n\}$ iff $f(x) = \langle y_1, \ldots, y_n\rangle$ is also surjective

$$g : x \mapsto \langle y_1, \ldots, y_n\rangle \mapsto \{y_1, \ldots, y_n\}$$

Thus, the set of all finite but non-empty subsets of $\mathbb{Z}^+$ is enumerable. Since $\{\varnothing\}$ is enumerable, by Proposition 5.1, $\{x \in \mathscr{P}(\mathbb{Z}^+) : x \text{ is finite}\}$ is enumerable.

Exercise 5.6. *Prove that a set a is enumerable iff there's an enumerable set b and a surjective function $f : b \to a$.*

5.2 Non-enumerability

Are all sets enumerable? Can the members of every set be arranged in a list? Cantor showed that the answer to this question is negative.

Theorem 5.2 (Cantor)**.** *The set of all sets of positive integers is not enumerable.*

Proof. Assume, for reductio, that $\mathscr{P}(\mathbb{Z}^+)$ is enumerable. Then, there's a surjective function $f : \mathbb{Z}^+ \to \mathscr{P}(\mathbb{Z}^+)$. By Separation, there's a set c whose members are the positive integers that don't belong to their image through f, i.e.

$$c = \{x \in \mathbb{Z}^+ : x \notin f(x)\}$$

Since $c \in \mathscr{P}(\mathbb{Z}^+)$ and f is surjective, there must be a $k \in \mathbb{Z}^+$ s.t. $f(k) = c$. Either $k \in c$ or $k \notin c$:

- If $k \in c$, then we must have that $k \notin f(k)$, which entails a contradiction, as $f(k) = c$.
- If $k \notin c$, since $k \in \mathbb{Z}^+$, we know that $k \in f(k)$, which entails a contradiction as well.

Thus, $\mathscr{P}(\mathbb{Z}^+)$ is not enumerable. □

Exercise 5.7. *Show that the set of all subsets of an infinite enumerable set is not enumerable.*

Exercise 5.8. *Let a subset a of $\mathbb{Z}^+$ be co-finite just in case $\mathbb{Z}^+ - a$ is finite. Is the set of co-finite subsets of $\mathbb{Z}^+$ enumerable? Justify your answers.*

Example 5.10. The set of all functions $f : \mathbb{Z}^+ \to \{0, 1\}$ is not enumerable.

To see it, let $f : \{x \subseteq \mathbb{Z}^{+2} : x : \mathbb{Z}^+ \to \{0, 1\}\} \to \mathscr{P}(\mathbb{Z}^+)$ be s.t. $f(x) = \{z \in \mathbb{Z}^+ : x(z) = 1\}$. In other words, f maps each function $x : \mathbb{Z}^+ \to \{0, 1\}$ to the subset of $\mathbb{Z}^+$ that has as members exactly those integers x assigns 1 to. We call x the 'characteristic function' of this subset. Note that f is surjective.

If $\{x \subseteq \mathbb{Z}^{+2} : x : \mathbb{Z}^+ \to \{0,1\}\}$ were enumerable, there would be a surjective function $g : \mathbb{Z}^+ \to \{x \subseteq \mathbb{Z}^{+2} : x : \mathbb{Z}^+ \to \{0,1\}\}$. But, then, $f \circ g : \mathbb{Z}^+ \to \mathscr{P}(\mathbb{Z}^+)$ would be surjective too, contradicting Cantor's Theorem.

Exercise 5.9. *Show that the set of all functions $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ is not enumerable.*

Example 5.11. The set of real numbers is not enumerable.

To see it, let $f : [0,1] \to \{x \subseteq \mathbb{Z}^{+2} : x : \mathbb{Z}^+ \to \{0,1,\ldots,9\}\}$ be s.t. $f(x) = y$ iff, for each $z \in \mathbb{Z}^+$, $y(z)$ is the z-th digit in the shortest decimal expansion of x,[1] i.e. if $k_i \in \{0,1,\ldots,9\}$

$$f : 0.k_0k_1\ldots k_n\ldots \;\mapsto\; x : \mathbb{Z}^+ \to \{0,1,\ldots,9\} \text{ s.t. } x(y) = k_y$$

Note that f is surjective. Thus, if $[0,1]$ were enumerable, by Exercise 5.6 so would $\{x \subseteq \mathbb{Z}^{+2} : x : \mathbb{Z}^+ \to \{0,1,\ldots,9\}\}$, contradicting Proposition 5.3 and Exercise 5.9.

Exercise 5.10. *Is the set of ordered pairs of real numbers enumerable? Justify your answers.*

Exercise 5.11. *Is the set of interpretations of the language of propositional logic – i.e. the functions that assign T or F to each sentence letter $P_1, \ldots, P_n, \ldots$ – enumerable? Justify your answers.*

Exercise 5.12. *What (if anything) is wrong with the following argument, known as 'Richard's Paradox'? Justify your answers.*

> *The set of all finite strings of symbols from the alphabet, including the space, capital letters, and punctuation marks, is enumerable; and for definiteness let us use the specific enumeration of finite strings based on prime decomposition. Some strings amount to definitions in English of sets of positive integers and others do not. Strike out the ones that do not, and we are left with an enumeration of all definitions in English of sets of positive integers, or, replacing each definition by the set it defines, an enumeration of all sets of positive integers that have definitions in English. Since some sets have more than one definition, there will be redundancies in this enumeration of sets. Strike them out to obtain an irredundant enumeration of all sets of positive integers that have definitions in English.*
>
> *Now consider the set of positive integers defined by the condition that a positive integer n is to belong to the set if and only if it does not belong to the n-th set in the irredundant enumeration just*

[1]Some reals have two decimal expansions. For instance, $0.2\widehat{9}$ and $0.3\widehat{0}$ denote the same number. The we take the latter to be the shortest decimal expansion, as it ends with 0s, which can be omitted.

> *described. This set does not appear in that enumeration. For it cannot appear at the n-th place for any n, since there is a positive integer, namely n itself, that belongs to this set if and only if it does not belong to the n-th set in the enumeration. Since this set does not appear in our enumeration, it cannot have a definition in English. And yet it does have a definition in English, and in fact we have just given such a definition in the preceding paragraph.*

5.3 Cardinality

We could say that $\mathscr{P}(\mathbb{Z}^+)$, the set of all functions on $\mathbb{Z}^+$, and $\mathbb{R}$ are all 'larger' than $\mathbb{Z}^+$. In general, if there's no surjective function $f : a \to b$ – if we can't 'fit' b in a – we say that the b is larger than a or that a is smaller than b. Moreover, if there are surjective functions $f : a \to b$ and $g : b \to a$, we say that a and b are *equinumerous*. Thus, if there is a surjective function $f : a \to b$, we say that b is smaller than or equinumerous with a.

Exercise 5.13. *Show that, if a, b, and c are sets, the following holds:*

1. *a is equinumerous with a.*
2. *If a is equinumerous with b, then b is equinumerous with a.*
3. *If a is equinumerous with b and b with c, then a is equinumerous with c.*

We can say that sets that are equinumerous to one another have the same size or *cardinality*. We write $|a|$ for the cardinality of a.

Theorem 5.3 (Cantor-Bernstein)**.** *If there are surjective functions $f : a \to b$ and $g : b \to a$, then there is a bijective function $h : a \to b$.*

Corollary 5.1. $|a| = |b|$ *iff there is a bijective function $f : a \to b$.*

Exercise 5.14. *Prove Corollary 5.1.*

Example 5.12. $|\mathbb{Z}^+| < |\mathscr{P}(\mathbb{Z}^+)|$

Thus, one can say that there are infinities of different sizes.

Theorem 5.4 (generalised Cantor)**.** $\mathsf{ZFC} \vdash \forall x \, |x| < |\mathscr{P}(x)|$

Gloss: The cardinality of a set is always smaller than the cardinality of its power set.

Proof. Since $f : \mathscr{P}(x) \to x$ s.t. f maps each singleton $\{y\}$ to y and every other subset of x to some $x_0 \in x$ is surjective (unless $x = \varnothing$, in which case the result follows trivially), we know that $|x| \leqslant |\mathscr{P}(x)|$.

To show that $|x| \neq |\mathscr{P}(x)|$, assume, for reductio, that $|x| = |\mathscr{P}(x)|$. Then, there's a bijective function $g : x \to \mathscr{P}(x)$. By Separation, there's a set c whose members are elements of x that don't belong to their image through g, i.e.

$$c = \{y \in x : y \notin g(y)\}$$

Since $c \in \mathscr{P}(x)$ and g is bijective, there must be an $x_d \in x$ s.t. $g(x_d) = c$. Either $x_d \in c$ or $x_d \notin c$:

- If $x_d \in c$, then we must have that $x_d \notin g(x_d)$, which entails a contradiction, as $g(x_d) = c$.
- If $x_d \notin c$, since $x_d \in x$, we know that $x_d \in g(x_d)$, which entails a contradiction as well.

Thus, $|x| \neq |\mathscr{P}(x)|$, so $|x| < |\mathscr{P}(x)|$. □

So one can say there are infinitely many infinities of different sizes!

Is there a set larger than $\mathbb{Z}^+$ but smaller than the power set of $\mathbb{Z}^+$, i.e. a set a s.t. $|\mathbb{Z}^+| < |a| < |\mathscr{P}(\mathbb{Z}^+)|$? The proposition that there is no such set is known as the *Continuum Hypothesis*. The proposition that there is no set whose cardinality is greater than the cardinality of a set x and smaller than the cardinality of $\mathscr{P}(x)$ is known as the *Generalised Continuum Hypothesis*.

Theorem 5.5 (Gödel-Cohen)**.** *The Continuum Hypothesis is independent of* ZFC*, that is, neither* $\mathsf{ZFC} \vdash \neg\exists x\, |\mathbb{Z}^+| < |x| < |\mathscr{P}(\mathbb{Z}^+)|$ *nor* $\mathsf{ZFC} \vdash \exists x\, |\mathbb{Z}^+| < |x| < |\mathscr{P}(\mathbb{Z}^+)|$.

Theorem 5.6 (Gödel-Cohen)**.** *The Generalised Continuum Hypothesis is independent of* ZFC*, that is,* $\mathsf{ZFC} \nvdash \forall x \neg\exists y\, |x| < |y| < |\mathscr{P}(x)|$ *and* $\mathsf{ZFC} \nvdash \neg\forall x \neg\exists y\, |x| < |y| < |\mathscr{P}(x)|$.

Chapter 6

Semantics for First-Order Languages

Let $\mathscr{L}$ be a first-order language.

Definition 6.1 (interpretation/model)**.** An interpretation/model $\mathcal{M}$ of $\mathscr{L}$ is an ordered pair of sets $\langle |\mathcal{M}|, .^{\mathcal{M}} \rangle$ – where $|\mathcal{M}|$ is the domain and $.^{\mathcal{M}}$ the interpretation function of $\mathcal{M}$ – s.t.

- $|\mathcal{M}| \neq \varnothing$
- for each individual constant $c \in V(\mathscr{L})$, $c^{\mathcal{M}} \in |\mathcal{M}|$
- for each n-place function symbol $f \in V(\mathscr{L})$, $f^{\mathcal{M}} : |\mathcal{M}|^n \to |\mathcal{M}|$
- for each n-place predicate symbol $P \in V(\mathscr{L})$, $P^{\mathcal{M}} \subseteq |\mathcal{M}|^n$

Some languages have *intended* or *standard* interpretations.

Example 6.1 (the standard interpretation of $\mathscr{L}_{\mathsf{PA}}$)**.** The standard/intended interpretation of $\mathscr{L}_{\mathsf{PA}}$, denoted by $\mathcal{N}$ (and sometimes $\mathbb{N}$) is s.t.

- $|\mathcal{N}| = \mathbb{N}$
- $0^{\mathcal{N}} = 0$
- $\mathrm{S}^{\mathcal{N}} : \mathbb{N} \to \mathbb{N}$ s.t. $\mathrm{S}^{\mathcal{N}}(x) = x + 1$
- $+^{\mathcal{N}} : \mathbb{N}^2 \to \mathbb{N}$ s.t. $+^{\mathcal{N}}(x, y) = x + y$
- $\times^{\mathcal{N}} : \mathbb{N}^2 \to \mathbb{N}$ s.t. $\times^{\mathcal{N}}(x, y) = x \times y$

Proposition 6.1 (no standard interpretation of $\mathscr{L}_{\in}$)**.** *$\mathscr{L}_{\in}$ has no standard interpretation.*

Proof. In their intended reading, the quantifiers in $\mathscr{L}_{\in}$ range over all sets. Thus, if $\mathscr{L}_{\in}$ had an intended interpretation $\mathcal{M}$, its domain would have to have all sets as members, i.e. $|\mathcal{M}|$ would have to be the set of all sets, which doesn't exist, by Theorem 3.1. $\square$

In what follows, let $\mathcal{M}$ be any interpretation of $\mathscr{L}$.

Definition 6.2 (assignment)**.** An assignment σ over $\mathcal{M}$ is a function from the set of terms of $\mathscr{L}$ to $|\mathcal{M}|$ s.t.

- for each individual constant $c \in V(\mathscr{L})$, $\sigma(c) = c^{\mathcal{M}}$
- for each n-place function symbol $f \in V(\mathscr{L})$ and terms $t_1, \ldots, t_n$ of $\mathscr{L}$,

$$\sigma(f(t_1, \ldots, t_n)) = f^{\mathcal{M}}(\sigma(t_1), \ldots, \sigma(t_n))$$

In what follows, let σ be an assignment over $\mathcal{M}$. If $d \in |\mathcal{M}|$, let $\sigma[v : d]$ be the assignment that is identical to σ except it maps v to d.

Definition 6.3 (satisfaction)**.** The notion of satisfaction of φ by $\mathcal{M}, \sigma$ ($\mathcal{M}, \sigma \vDash \varphi$) obeys the following conditions:

- $\mathcal{M}, \sigma \vDash t_1 = t_2$ iff $\sigma(t_1) = \sigma(t_2)$
- $\mathcal{M}, \sigma \vDash P(t_1, \ldots, t_n)$ iff $\langle \sigma(t_1), \ldots, \sigma(t_n) \rangle \in P^{\mathcal{M}}$
- $\mathcal{M}, \sigma \vDash \neg\varphi$ iff $\mathcal{M}, \sigma \nvDash \varphi$
- $\mathcal{M}, \sigma \vDash (\varphi \to \psi)$ iff $\mathcal{M}, \sigma \nvDash \varphi$ or $\mathcal{M}, \sigma \vDash \psi$
- $\mathcal{M}, \sigma \vDash \forall v\, \varphi$ iff $\mathcal{M}, \sigma[v : d] \vDash \varphi$ for each $d \in |\mathcal{M}|$

Definition 6.4 (truth)**.** Let φ be a sentence. $\mathcal{M} \vDash \varphi$ iff $\mathcal{M}, \sigma \vDash \varphi$ for some σ.

Proposition 6.2 (soundness of PA)**.** *The axioms of* PA *are all true in* $\mathcal{N}$*.*

*For **PA1**, note that*

$\forall y \in \mathbb{N}\, y + 1 \neq 0$	*so*	$\forall y \in \lvert\mathcal{N}\rvert\, \mathrm{S}^{\mathcal{N}}(y) \neq 0$	*def. of* $\mathcal{N}$
	so	$\mathrm{S}^{\mathcal{N}}(\sigma(x)) \neq 0$*, for every* σ *over* $\mathcal{N}$	*Def. of assign.*
	so	$\sigma(\mathrm{S}x) \neq \sigma(0)$*, for every* σ *over* $\mathcal{N}$	*Def. of assign.*
	so	$\mathcal{N}, \sigma \nvDash \mathrm{S}x = 0$*, for every* σ *over* $\mathcal{N}$	*Def. of sat.*
	so	$\mathcal{N}, \sigma \vDash \mathrm{S}x \neq 0$*, for every* σ *over* $\mathcal{N}$	*Def. of sat.*
	so	$\mathcal{N}, \sigma \vDash \forall x\, \mathrm{S}x \neq 0$*, for every* σ *over* $\mathcal{N}$	*Def. of sat.*
	so	$\mathcal{N} \vDash \forall x\, \mathrm{S}x \neq 0$	*Def. of truth*

Exercise 6.1. *Finish the proof of Proposition 6.2.*

Definition 6.5 (logical truth)**.** Let φ be a sentence. φ is a logical truth iff $\mathcal{M} \vDash \varphi$ for every $\mathcal{M}$.

Definition 6.6 (logical falsity)**.** Let φ be a sentence. φ is a logical falsity iff $\mathcal{M} \nvDash \varphi$ for every $\mathcal{M}$.

Let '$\mathcal{M} \vDash \Gamma$' be short for '$\mathcal{M} \vDash \psi$ for each $\psi \in \Gamma$'; e.g. $\mathcal{N} \vDash \mathsf{PA}$.

Definition 6.7 (logical consequence)**.** $\Gamma \vDash \varphi$ iff, for every $\mathcal{M}$ and every σ, if $\mathcal{M}, \sigma \vDash \Gamma$, $\mathcal{M}, \sigma \vDash \varphi$.

Exercise 6.2 (Modus Ponens)**.** *Show that, if* $\Gamma \vDash \varphi$ *and* $\Gamma \vDash \varphi \to \psi$*, then* $\Gamma \vDash \psi$*.*

Exercise 6.3 (Generalisation)**.** *Show that, if* $\Gamma \vDash \varphi$ *and* v *is not free in any member of* Γ*, then* $\Gamma \vDash \forall v\, \varphi$*.*

Exercise 6.4. *Show that, if* φ *is an instance of an axiom* **A1-A8***, then* $\Gamma \vDash \varphi$*.*

Exercise 6.5 (structural properties)**.** *Show that the following holds:*

- *Reflexivity.* $\Gamma, \varphi \vDash \varphi$*.*
- *Transitivity. If* $\Gamma \vDash \varphi$ *and* $\Delta, \varphi \vDash \psi$*, then* $\Gamma \cup \Delta \vDash \psi$*.*
- *Monotonicity. If* $\Gamma \vDash \varphi$*, then* $\Gamma \cup \Delta \vDash \varphi$*.*

Definition 6.8 (logical equivalence)**.** φ and ψ are logically equivalent iff, for every $\mathcal{M}$ and every σ, $\mathcal{M}, \sigma \vDash \varphi$ iff $\mathcal{M}, \sigma \vDash \psi$.

Definition 6.9 (satisfiability)**.** Γ is satisfiable iff there's an $\mathcal{M}$ and a σ s.t. $\mathcal{M}, \sigma \vDash \Gamma$.

If Γ is a set of sentences and $\mathcal{M} \vDash \Gamma$, we say that $\mathcal{M}$ *is a model of* Γ; e.g. $\mathcal{N}$ is a model of PA.

Exercise 6.6. *Show that, if* Γ *is satisfiable and* $\Delta \subseteq \Gamma$*, then* Δ *is satisfiable. Show that the converse is not true.*

Chapter 7

The size and number of models

7.1 The size of models

The size of a model is the *cardinality* of its domain. Models can be finite – i.e. have domains with one, two, three, etc. elements – or infinite. In the latter case, they can be *denumerable* or non-enumerable/*uncountable*. And if they are uncountable they can be equinumerous with $\mathscr{P}(\mathbb{Z}^+)$, or with $\mathscr{P}(\mathscr{P}(\mathbb{Z}^+))$, or $\mathscr{P}(\mathscr{P}(\mathscr{P}(\mathbb{Z}^+)))$, etc.

Example 7.1. $\forall x \forall y\, x = y$ has only models of cardinality 1.

Example 7.2. All models of $\forall x \forall y \forall z\, (x = y \vee y = z \vee x = z)$ are of cardinality at most 2.

Example 7.3. All models of $\exists x \exists y\, x \neq y$ and $\forall x \exists y\, x \neq y$ are of cardinality at least 2.

Example 7.4. For each $n > 1$, let $\gamma_{\geqslant n}$ be

$$\forall x_1 \ldots \forall x_{n-1} \exists y\, (y \neq x_1 = \wedge y \neq x_2 \wedge \cdots \wedge y \neq x_{n-1})$$

$\gamma_{\geqslant n}$ has only models of cardinality at least n.

Example 7.5. For each $n > 1$, let $\gamma_{\leqslant n}$ be $\neg\gamma_{\geqslant n+1}$. $\gamma_{\leqslant n}$ has only models of cardinality at most n.

Example 7.6. For each $n > 1$, let $\gamma_{=n}$ be $\gamma_{\leqslant n} \wedge \gamma_{\geqslant n}$. $\gamma_{=n}$ has only models of cardinality n.

Exercise 7.1. *Give a sentence that only has models of cardinality 2 or 8.*

Example 7.7. $\{\gamma_{\geqslant n} : n > 1\}$ has only infinite models.

Exercise 7.2. *What is the cardinality (finite, denumerable, or non-denumerable) of the smallest model of the following set?*

$$\{\gamma_{\geqslant n} : n \text{ is even}\}$$

Justify your answers.

Example 7.8. The following sentence has only infinite models:

$$\forall x \exists y\, Rxy \wedge \forall x \forall y\, \neg(Rxy \wedge Ryx) \wedge \forall x \forall y \forall z\, (Rxy \wedge Ryz \rightarrow Rxz)$$

Example 7.9. PA has only infinite models.
To see it, let $x < y$ be defined in $\mathscr{L}_{\mathsf{PA}}$ as $\exists z\, y = x + \mathrm{S}z$ (cf. Exercise 2.12). Moreover, take another look at the previous example and note that

$$\mathsf{PA} \vDash \forall x \exists y\, x < y \wedge \forall x \forall y\, \neg(x < y \wedge y < x) \wedge \forall x \forall y \forall z\, (x < y \wedge y < z \rightarrow x < z)$$

Alternatively, for each natural number n let us write $\overline{n}$ for $\underbrace{\mathrm{S} \ldots \mathrm{S}}_{n \text{ times}} 0$ – which we call the 'numeral' of n – and note that, by **PA1** and **PA2**, for every m and every n, if $m \neq n$, $\mathsf{PA} \vDash \overline{m} \neq \overline{n}$.

Exercise 7.3. *Let $<$ be defined as in Example 7.9. Prove that*

$$\mathsf{PA} \vDash \forall x \exists y\, x < y \wedge \forall x \forall y\, \neg(x < y \wedge y < x) \wedge \forall x \forall y \forall z\, (x < y \wedge y < z \rightarrow x < z)$$

7.2 Isomorphism, structures, and the indeterminacy of reference

Definition 7.1 (isomorphism)**.** An isomorphism between two interpretations $\mathcal{M}$ and $\mathcal{M}'$ of $\mathscr{L}$ is a bijection $h : |\mathcal{M}| \rightarrow |\mathcal{M}'|$ s.t.

- for each individual constant $c \in V(\mathscr{L})$, $h(c^{\mathcal{M}}) = c^{\mathcal{M}'}$
- for each n-place function symbol $f \in V(\mathscr{L})$ and every $x_1, \ldots x_n \in |\mathcal{M}|$,

$$h(f^{\mathcal{M}}(x_1, \ldots, x_n)) = f^{\mathcal{M}'}(h(x_1), \ldots, h(x_n))$$

- for each n-place predicate symbol $P \in V(\mathscr{L})$ and every $x_1, \ldots x_n \in |\mathcal{M}|$,

$$\langle x_1, \ldots, x_n \rangle \in P^{\mathcal{M}} \text{ iff } \langle h(x_1), \ldots, h(x_n) \rangle \in P^{\mathcal{M}'}$$

When there's an isomorphism between $\mathcal{M}$ and $\mathcal{M}'$ we say that they are isomorphic and write $\mathcal{M} \cong \mathcal{M}'$.

Example 7.10. Let $V(\mathscr{L}) = \{a, b, c, R\}$ and $\mathcal{M}_1$ and $\mathcal{M}_2$ be s.t.

- $|\mathcal{M}_1| = \{1, 2, 3\}$, $a^{\mathcal{M}_1} = 1$, $b^{\mathcal{M}_1} = 2$, $c^{\mathcal{M}_1} = 3$, and $R^{\mathcal{M}_1} = \{\langle 1, 1\rangle, \langle 2, 1\rangle, \langle 3, 1\rangle, \langle 3, 2\rangle\}$
- $|\mathcal{M}_2| = \{1, 2, 3\}$, $a^{\mathcal{M}_2} = 2$, $b^{\mathcal{M}_2} = 3$, $c^{\mathcal{M}_2} = 1$, and $R^{\mathcal{M}_2} = \{\langle 2, 2\rangle, \langle 3, 2\rangle, \langle 1, 2\rangle, \langle 1, 3\rangle\}$

$\mathcal{M}_1$ and $\mathcal{M}_2$ are isomorphic.
To see it, let $h : \{1, 2, 3\} \rightarrow \{1, 2, 3\}$ be s.t. $h(1) = 2, h(2) = 3$, and $h(3) = 1$, which is clearly bijective. Thus, we have that:

- $h(a^{\mathcal{M}_1}) = h(1) = 2 = a^{\mathcal{M}_2}$
- $h(b^{\mathcal{M}_1}) = h(2) = 3 = b^{\mathcal{M}_2}$
- $h(c^{\mathcal{M}_1}) = h(3) = 1 = c^{\mathcal{M}_2}$
- we can check that, for each $x_1, x_2 \in |\mathcal{M}_1|$ (because they are finitely many!), $\langle h(x_1), h(x_2)\rangle \in R^{\mathcal{M}_2}$ and that only the latter belong to $R^{\mathcal{M}_2}$.

Exercise 7.4. *Let $V(\mathscr{L})$ and $\mathcal{M}_1$ be as before and $\mathcal{M}_3$ be s.t. $|\mathcal{M}_3| = \{4, 5, 6\}$, $a^{\mathcal{M}_3} = 4$, $b^{\mathcal{M}_3} = 5$, $c^{\mathcal{M}_3} = 6$, and $R^{\mathcal{M}_3} = \{\langle 4,4\rangle, \langle 5,4\rangle, \langle 6,4\rangle, \langle 6,5\rangle\}$. Show that $\mathcal{M}_1 \cong \mathcal{M}_3$.*

Exercise 7.5. *Let $\mathcal{M}$ be an interpretation of $\mathscr{L}_{\mathsf{PA}}$ s.t. $|\mathcal{M}| = \{x \in \mathbb{N} : x$ is even$\}$, $0^{\mathcal{M}} = 0$, $\mathrm{S}^{\mathcal{M}}(x) = x+2$, $+^{\mathcal{M}}(x,y) = x+y$, and $\times^{\mathcal{M}}(x,y) = (x \times y) \div 2$. Show that $\mathcal{N} \cong \mathcal{M}$.*

Exercise 7.6. *Let $V(\mathscr{L}) = \{0, \mathrm{S}, <\}$ and let $\mathcal{M}$ and $\mathcal{M}'$ be interpretations of $\mathscr{L}$ such that*

- *$|\mathcal{M}| = \mathbb{N}$, $0^{\mathcal{M}} = 0$, $S^{\mathcal{M}}(x) = x+1$, and $<^{\mathcal{M}} = \{\langle x, y\rangle \in \mathbb{N}^2 : x < y\}$*
- *$|\mathcal{M}'| = \{x \in \mathbb{N} : x$ is odd$\}$, $0^{\mathcal{M}'} = 1$, $S^{\mathcal{M}'}(x) = x+2$, and $<^{\mathcal{M}'} = \{\langle x, y\rangle \in |\mathcal{M}'|^2 : x < y\}$*

Show that $\mathcal{M}$ and $\mathcal{M}'$ are isomorphic.

Exercise 7.7. *Let $\mathcal{M}$, $\mathcal{M}'$, and $\mathcal{M}''$ be interpretations of $\mathscr{L}$. Show that the following claims hold:*

- *$\mathcal{M} \cong \mathcal{M}$*
- *If $\mathcal{M} \cong \mathcal{M}'$, then $\mathcal{M}' \cong \mathcal{M}$*
- *If $\mathcal{M} \cong \mathcal{M}'$ and $\mathcal{M}' \cong \mathcal{M}''$, then $\mathcal{M} \cong \mathcal{M}''$*

Lemma 7.1. *If h is an isomorphism between two interpretations $\mathcal{M}$ and $\mathcal{M}'$ of $\mathscr{L}$ and σ an assignment over $\mathcal{M}$, then $h \circ \sigma$ is an assignment over $\mathcal{M}'$.*

Proof. Note first that $h \circ \sigma$ maps terms of $\mathscr{L}$ to members of $|\mathcal{M}'|$. Also:

- If c is an individual constant of $V(\mathscr{L})$,

$$\begin{aligned} h \circ \sigma(c) &= h(\sigma(c)) && \text{Def. of } \circ \\ &= h(c^{\mathcal{M}}) && \text{Def. of assign.} \\ &= c^{\mathcal{M}'} && \text{Def. of iso.} \end{aligned}$$

- If f is an n-place function of $V(\mathscr{L})$ and $t_1, \ldots, t_n$ are terms of $\mathscr{L}$,

$$\begin{aligned}
h \circ \sigma(f(t_1, \ldots, t_n)) &= h(\sigma(f(t_1, \ldots, t_n))) && \text{Def. of } \circ \\
&= h(f^{\mathcal{M}}(\sigma(t_1), \ldots, \sigma(t_n))) && \text{Def. of assign.} \\
&= f^{\mathcal{M}'}(h(\sigma(t_1)), \ldots, h(\sigma(t_n))) && \text{Def. of iso.} \\
&= f^{\mathcal{M}'}(h \circ \sigma(t_1), \ldots, h \circ \sigma(t_n)) && \text{Def. of } \circ
\end{aligned}$$

□

Lemma 7.2. *If h is an isomorphism between two interpretations $\mathcal{M}$ and $\mathcal{M}'$ of $\mathscr{L}$ and σ an assignment over $\mathcal{M}$, then for every wff φ of $\mathscr{L}$:*

$$\mathcal{M}, \sigma \vDash \varphi \quad \textit{iff} \quad \mathcal{M}', h \circ \sigma \vDash \varphi$$

Proof. We prove our result by induction on the complexity of φ.
Base step. If φ is atomic, it's either of the form $t_1 = t_2$ or $P(t_1, \ldots, t_n)$.

- If φ is of the form $t_1 = t_2$,

$$\begin{aligned}
\mathcal{M}, \sigma \vDash \varphi \quad &\text{iff} \quad \sigma(t_1) = \sigma(t_2) && \text{Def. of sat.} \\
&\text{iff} \quad h(\sigma(t_1)) = h(\sigma(t_2)) && h \text{ is bijective} \\
&\text{iff} \quad \mathcal{M}', h \circ \sigma \vDash t_1 = t_2 && \text{Lemma 7.1}
\end{aligned}$$

- If φ is of the form $P(t_1, \ldots, t_n)$,

$$\begin{aligned}
\mathcal{M}, \sigma \vDash \varphi \quad &\text{iff} \quad \langle \sigma(t_1), \ldots, \sigma(t_n) \rangle \in P^{\mathcal{M}} && \text{Def. of sat.} \\
&\text{iff} \quad \langle h(\sigma(t_1)), \ldots, h(\sigma(t_n)) \rangle \in P^{\mathcal{M}'} && \text{Def. of iso.} \\
&\text{iff} \quad \mathcal{M}', h \circ \sigma \vDash P(t_1, \ldots, t_n) && \text{Lemma 7.1}
\end{aligned}$$

Inductive step. Assume the claim holds of all wffs of complexity less than n and let φ be of complexity n.

- If φ is of the form $\neg\psi$,

$$\begin{aligned}
\mathcal{M}, \sigma \vDash \varphi \quad &\text{iff} \quad \mathcal{M}, \sigma \nvDash \psi && \text{Def. of sat.} \\
&\text{iff} \quad \mathcal{M}', h \circ \sigma \nvDash \psi && \text{IH} \\
&\text{iff} \quad \mathcal{M}', h \circ \sigma \vDash \neg\psi && \text{Def. of sat.}
\end{aligned}$$

- The case in which φ is of the form $\psi \to \chi$ is left as an exercise.
- If φ is of the form $\forall v\, \psi$,

$$\begin{aligned}
\mathcal{M}, \sigma \vDash \varphi \quad &\text{iff} \quad \mathcal{M}, \sigma[v : d] \vDash \psi, \text{ for each } d \in |\mathcal{M}| && \text{Def. of sat.} \\
&\text{iff} \quad \mathcal{M}', h \circ \sigma[v : d] \vDash \psi, \text{ for each } d \in |\mathcal{M}| && \text{IH} \\
&\text{iff} \quad \mathcal{M}', (h \circ \sigma)[v : d] \vDash \psi, \text{ for each } d \in |\mathcal{M}'| && \text{Def. of iso.} \\
&\text{iff} \quad \mathcal{M}', h \circ \sigma \vDash \forall v\, \psi && \text{Def. of sat.}
\end{aligned}$$

Thus, the result holds for every φ. $\square$

Exercise 7.8. *Finish the proof of Lemma 7.2.*

Theorem 7.1 (isomorphism)**.** *If there is an isomorphism between two interpretations $\mathcal{M}$ and $\mathcal{M}'$ of $\mathscr{L}$, then for every sentence φ of $\mathscr{L}$, $\mathcal{M} \vDash \varphi$ iff $\mathcal{M}' \vDash \varphi$.*

Gloss: Isomorphic interpretations make the same sentences true.

Exercise 7.9. *Show that the converse of Theorem 7.1 doesn't hold.*

Example 7.11. Let $\mathcal{M}$ be an interpretation of $\mathscr{L}_{\mathsf{PA}}$ s.t. $|\mathcal{M}| = \mathbb{Z}^+$, $0^{\mathcal{M}} = 1$, $\mathrm{S}^{\mathcal{M}}(x) = x + 1$, $+^{\mathcal{M}}(x, y) = x + y$, and $\times^{\mathcal{M}}(x, y) = x \times y$. Then, $\mathcal{N}$ and $\mathcal{M}$ are not isomorphic.
To see it, note that $\forall x\, x + 0 = x$ is true in $\mathcal{N}$ but not in $\mathcal{M}$.

Exercise 7.10. *Let $V(\mathscr{L}) = \{<\}$ and let $\mathcal{M}$ and $\mathcal{M}'$ be interpretations of $\mathscr{L}$ s.t.*

- $|\mathcal{M}| = \{0, 2, 4\}$ *and* $<^{\mathcal{M}} = \{\langle x, y\rangle \in |\mathcal{M}|^2 : x < y\}$
- $|\mathcal{M}'| = \{1, 3, 5\}$ *and* $<^{\mathcal{M}'} = \{\langle x, y\rangle \in |\mathcal{M}'|^2 : x \leqslant y\}$

Show that $\mathcal{M}$ and $\mathcal{M}'$ are not isomorphic.

An important consequence of Theorem 7.1 is the following. For the rest of this section, let Γ be a set of sentences.

Corollary 7.1 (push-through)**.** *If $\mathcal{M}$ is a model of Γ, then all interpretations that are isomorphic with $\mathcal{M}$ are models of Γ too.*

Proposition 7.1. *Let $\mathcal{M}$ be an interpretation of $\mathscr{L}$, a a set, and $h : |\mathcal{M}| \to a$ a bijection. Then, we can define an interpretation $\mathcal{M}'$ of $\mathscr{L}$ to be isomorphic to $\mathcal{M}$ as follows. Let:*

- $|\mathcal{M}'| = a$
- *for each individual constant $c \in V(\mathscr{L})$, $c^{\mathcal{M}'} = h(c^{\mathcal{M}})$*
- *for each n-place function symbol $f \in V(\mathscr{L})$ and each $x_1, \ldots, x_n \in |\mathcal{M}|$,*

$$f^{\mathcal{M}'}(h(x_1), \ldots, h(x_n)) = h(f^{\mathcal{M}}(x_1, \ldots, x_n))$$

- *for each n-place predicate symbol $P \in V(\mathscr{L})$ and every $x_1, \ldots x_n \in |\mathcal{M}|$,*

$$\langle h(x_1), \ldots, h(x_n)\rangle \in P^{\mathcal{M}'} \quad \textit{iff} \quad \langle x_1, \ldots, x_n\rangle \in P^{\mathcal{M}}$$

Gloss: Every bijection $h : |\mathcal{M}| \to a$ induces an isomorphic model $\mathcal{M}'$ with a as its domain.

Exercise 7.11. *Prove Proposition 7.1.*

Corollary 7.1 and Proposition 7.1 entail, for instance, that each permutation of the elements of a model of Γ delivers another model of Γ, e.g. there is a model of PA in which the constant '0' denotes 1 and 'S0' denotes 0. Moreover, they also entail that every set of sentences that has a model also has a numerical model.

Exercise 7.12. *Show that, if Γ has a model of cardinality n, it has a model whose domain is $\{1, 2, \ldots, n\}$ (i.e. a numerical model).*

Exercise 7.13. *Show that, if Γ has a denumerable model, it has a model whose domain is $\mathbb{Z}^+$ (i.e. a numerical model).*

In other words, first-order theories aren't strong enough to pin down *particular* interpretations but can only do so *up to isomorphism.* A set of sentences cannot have just one model, but might have just one *type of model* in the sense that all its models are isomorphic to each other or, as we will call it, one 'isomorphism type'.

This has important philosophical consequences, explored most saliently by Benacerraf and Putnam. According to Benacerraf, since, e.g. our arithmetical first-order theories – e.g. PA – cannot distinguish between 'the' natural numbers and any other isomorphic sequence of objects such as $\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \ldots,$ the only numerical properties that matter are those that are shared by all of these models. Thus, instead of focusing on particular models, mathematicians should focus on isomorphism types – i.e. structures.[1]

According to Putnam, if the meanings of the expressions in our theories were to be explained in terms of their models, our push-through result would force us to conclude that reference is radically indeterminate. Thus, he suggests that meaning must be fixed in an alternative manner, roughly by the principles that govern the expressions of the language, i.e. the axioms of our theories.[2]

Example 7.12. $\{\forall x\,(x = a \vee y = b), Rab, \forall x\,(Rxy \rightarrow \neg Ryx)\}$ has one isomorphism type.
Assume $\mathcal{M} \vDash \{\forall x\,(x = a \vee y = b), Rab, \forall x\,(Rxy \rightarrow \neg Ryx)\}$. Then, $\langle a^{\mathcal{M}}, b^{\mathcal{M}}\rangle \in R^{\mathcal{M}}$. But then, $\langle b^{\mathcal{M}}, a^{\mathcal{M}}\rangle \notin R^{\mathcal{M}}$, and also $\langle a^{\mathcal{M}}, a^{\mathcal{M}}\rangle \notin R^{\mathcal{M}}$ and $\langle b^{\mathcal{M}}, b^{\mathcal{M}}\rangle \notin R^{\mathcal{M}}$, so $a^{\mathcal{M}} \neq b^{\mathcal{M}}$. Moreover, $|\mathcal{M}| = \{a^{\mathcal{M}}, b^{\mathcal{M}}\}$, and so $R^{\mathcal{M}} = \{\langle a^{\mathcal{M}}, b^{\mathcal{M}}\rangle\}$.

Example 7.13. $\{Rab, \forall x\,(Rxy \rightarrow \neg Ryx)\}$ has more than one isomorphism type.
$\mathcal{M}_1$ and $\mathcal{M}_2$ are two non-isomorphic models of this set:

- $|\mathcal{M}_1| = \{1, 2\}$, $a^{\mathcal{M}_1} = 1$, $b^{\mathcal{M}_1} = 2$, and $R^{\mathcal{M}_1} = \{\langle 1, 2\rangle\}$.
- $|\mathcal{M}_2| = \{1, 2, 3\}$, $a^{\mathcal{M}_2} = 1$, $b^{\mathcal{M}_2} = 2$, and $R^{\mathcal{M}_1} = \{\langle 1, 2\rangle, \langle 2, 3\rangle\}$.

[1] Benacarraf, P. (1965) What numbers could not be, *Philosophical Review* 74, pp. 47-73.
[2] Putnam, H. (1980) Models and Reality, *Journal of Symbolic Logic* 45, pp. 464-482.

Exercise 7.14. *Let $V(\mathscr{L}_\equiv) = \{\equiv\}$ and let* **EQ** *be the following sentence of $\mathscr{L}_\equiv$:*

$$\forall x\, x \equiv x \ \wedge\ \forall x \forall y\, (x \equiv y \to y \equiv x) \ \wedge\ \forall x \forall y \forall z\, (x \equiv y \wedge y \equiv z \to x \equiv z) \quad (\mathbf{EQ})$$

Moreover, let $\mathcal{M}$ be a denumerable interpretation of $\mathscr{L}_\equiv$ s.t. $\mathcal{M} \vDash$ **EQ**.

1. *Show that $\equiv^{\mathcal{M}}$ is an equivalence relation on $|\mathcal{M}|$.*

2. *Let $\mathcal{M} \vDash \forall x \forall y\, x \equiv y$. In how many equivalence classes can $\equiv^{\mathcal{M}}$ partition $|\mathcal{M}|$? How many isomorphism types of denumerable model does $\{\mathbf{EQ}, \forall x \forall y\, x \equiv y\}$ have?*[3]

3. *Let $\mathcal{M} \vDash \forall x \forall y\, (x \equiv y \leftrightarrow x = y)$. In how many equivalence classes can $\equiv^{\mathcal{M}}$ partition $|\mathcal{M}|$? How many isomorphism types of denumerable model does $\{\mathbf{EQ}, \forall x \forall y\, (x \equiv y \leftrightarrow x = y)\}$ have?*

4. *Let $\mathcal{M} \vDash \forall x \forall y\, x \equiv y \vee \forall x \forall y\, (x \equiv y \leftrightarrow x = y)$. In how many equivalence classes can $\equiv^{\mathcal{M}}$ partition $|\mathcal{M}|$? How many isomorphism types of denumerable model does $\{\mathbf{EQ}, \forall x \forall y\, x \equiv y \vee \forall x \forall y\, (x \equiv y \leftrightarrow x = y)\}$ have?*

5. *For each n, indicate how to write down a sentence of $\mathscr{L}_\equiv$ that is true in a model of* **EQ** *iff there are at least n equivalence classes.*

6. *For each n, indicate how to write down a sentence of $\mathscr{L}_\equiv$ that is true in a model of* **EQ** *iff there are exactly n equivalence classes.*

7. *Let $\mathcal{M} \vDash \forall x \exists y\, (x \neq y \wedge x \equiv y \wedge \forall z\, (z \equiv x \to z = x \vee z = y))$. In how many equivalence classes can $\equiv^{\mathcal{M}}$ partition $|\mathcal{M}|$? How many isomorphism types of denumerable model does $\{\mathbf{EQ}, \forall x \exists y\, (x \neq y \wedge x \equiv y \wedge \forall z\, (z \equiv x \to z = x \vee z = y))\}$ have?*

8. *Let $\mathcal{M} \vDash \forall x \forall y\, (\exists z\, (z \neq x \wedge z \equiv x) \wedge \exists z\, (z \neq y \wedge z \equiv y) \to x \equiv y)$. In how many equivalence classes can $\equiv^{\mathcal{M}}$ partition $|\mathcal{M}|$? How many isomorphism types of denumerable model does $\{\mathbf{EQ}, \forall x \forall y\, (\exists z\, (z \neq x \wedge z \equiv x) \wedge \exists z\, (z \neq y \wedge z \equiv y) \to x \equiv y)\}$ have?*

9. *How many isomorphism types of denumerable model does* **EQ** *itself have?*

[3] That is, how many different denumerable models – in the sense of being non-isomorphic – does the set have?

Chapter 8

Soundness and Completeness of First-Order Logic

8.1 Soundness

Does our calculus allow us to infer false conclusions from true premises? In this section we will show that it doesn't, i.e. that, if φ is a proof theoretic consequence of Γ, then φ is a logical consequence of Γ. In other words, we will show that every inference that is admissible in our calculus is truth preserving, i.e. that our calculus is *sound*.

Theorem 8.1 (soundness)**.** *If* $\Gamma \vdash \varphi$*, then* $\Gamma \vDash \varphi$*.*

Proof. Assume $\Gamma \vdash \varphi$. Thus, there is a derivation $\mathcal{D}$ of φ from Γ. Let $\Delta \subseteq \Gamma$ be the set of members of Γ that occur in $\mathcal{D}$ as premises. We prove that $\Delta \vDash \varphi$ by induction on the length of $\mathcal{D}$. By monotonicity (cf. Exercise 6.5), we conclude that $\Gamma \vDash \varphi$.
Base step: $\mathcal{D}$ is one step long. Then, either φ is an instance of **A1-A8** or $\varphi \in \Delta$. If φ is an instance of **A1-A8**, by Exercise 6.4, $\Delta \vDash \varphi$. If $\varphi \in \Delta$, $\Delta \vDash \varphi$ by reflexivity (cf. Exercise 6.5).
Inductive step: Assume that if ψ is derivable from Δ in fewer than n steps, then $\Delta \vDash \psi$. Let $\mathcal{D}$ be exactly n steps long.

- If φ is a logical axiom or a member of Δ, we can reason as in the base step and conclude that $\Delta \vDash \varphi$.

- If φ was obtained by MP from two previous steps ψ and $\psi \to \varphi$ in $\mathcal{D}$, by inductive hypothesis we know that $\Delta \vDash \psi$ and that $\Delta \vDash \psi \to \varphi$. By Exercise 6.2, $\Delta \vDash \varphi$.

- If φ was obtained by Gen form a previous step ψ in $\mathcal{D}$, then φ must be of the form $\forall v\, \psi$ for some variable v not free in Δ. Moreover, by inductive hypothesis, $\Delta \vDash \psi$. By Exercise 6.3, we have that $\Delta \vDash \forall v\, \psi$.

□

Corollary 8.1. *If Γ is an inconsistent set of sentences, then Γ doesn't have a model.*

Exercise 8.1. *Prove Corollary 8.1.*

8.2 Completeness

We now know that our calculus derives some (and only) valid inferences. Is it sufficiently powerful to derive them *all*? In other words, is it true that, if $\Gamma \vDash \varphi$, then there is there a derivation of φ from Γ in our calculus? This question was originally posed as an open problem by David Hilbert in 1928. Kurt Gödel was the first to offer a positive answer together with a proof in his doctoral dissertation the following year. His result is known as the 'completeness theorem for first-order logic'. In this section we will offer a proof, but not Gödel's, which was rather complicated. Instead, we will roughly follow Leon Henkin's version, which appeared in the late 1940s and is now the standard way to prove completeness.

For simplicity, we will restrict our attention to the cases in which Γ is a set of sentences and φ a sentence. The proof can be modified to cover the cases in which free variables occur in members of Γ or in φ (let me know if you are interested!). To prove our result, we first need to introduce a few notions and establish a few lemmata and propositions.

In the rest of the book, let Γ be a set of sentences of $\mathscr{L}$.

Definition 8.1 (maximal consistency)**.** Γ is maximally consistent iff it is consistent and, for each sentence φ, if $\varphi \notin \Gamma$, then $\Gamma \cup \{\varphi\}$ is inconsistent.

Exercise 8.2. *Show that, if Γ is maximally consistent, for every sentence φ,*

$$\varphi \in \Gamma \quad \textit{iff} \quad \Gamma \vdash \varphi$$

Definition 8.2 (witness property)**.** Γ has the witness property iff for every φ s.t. $\neg\forall v\, \varphi \in \Gamma$, there is a closed term t s.t. $\neg\varphi[t/v] \in \Gamma$.

Proposition 8.1. *If Γ is consistent, $\neg\forall v\, \varphi \in \Gamma$, and c doesn't occur in Γ, then $\Gamma \cup \{\neg\varphi[c/v]\}$ is consistent too.*

Proof. Assume $\Gamma \cup \{\neg\varphi[c/v]\}$ is inconsistent. By $\neg$I, $\Gamma \vdash \varphi[c/v]$. By $\forall$I, we have that $\Gamma \vdash \forall v\, \varphi$. But we also have that $\Gamma \vdash \neg\forall v\, \varphi$, as $\neg\forall v\, \varphi \in \Gamma$. Thus, Γ is inconsistent, contradicting our initial assumptions. □

Exercise 8.3. *Show that, if Γ is maximally consistent, the following holds:*

1. *For every sentence φ, either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$.*
2. *If $(\varphi \to \psi) \in \Gamma$, then $\neg\varphi \in \Gamma$ or $\psi \in \Gamma$.*
3. *For every closed term t, if $\forall v\, \varphi \in \Gamma$, then $\varphi[t/v] \in \Gamma$.*

4. *If Γ has the witness property and, for every closed term t, $\varphi[t/v] \in \Gamma$, then $\forall v\, \varphi \in \Gamma$.*

Exercise 8.4. *Show that, if Γ is maximally consistent, then the following holds:*

1. *For every closed term t, $t = t \in \Gamma$.*
2. *If $t_1 = t_2 \in \Gamma$, then $t_2 = t_1 \in \Gamma$.*
3. *If $t_1 = t_2 \in \Gamma$ and $t_2 = t_3 \in \Gamma$, then $t_1 = t_3 \in \Gamma$.*

Gloss: The relation that holds between two closed terms t_1 and t_2 just in case $t_1 = t_2 \in \Gamma$ is an equivalence relation. As such, it generates a partition on the set of closed terms of the language of Γ into equivalence classes: for each closed term t, let

$$[t] = \{s : s \text{ is a closed term and } s = t \in \Gamma\}$$

Exercise 8.5. *Show that, if $\mathcal{M}$ is an interpretation of $\mathscr{L}$, the following set is maximally consistent:*

$$\{\varphi : \varphi \textit{ is a sentence of } \mathscr{L} \textit{ and } \mathcal{M} \vDash \varphi\}$$

In what follows, let $V(\mathscr{L}^+)$ obtain by adding denumerably many *new* individual constants to $V(\mathscr{L})$.

Lemma 8.1 (Lindenbaum)**.** *If Γ is consistent, then it can be extended to a maximally consistent set Γ^+ of sentences of $\mathscr{L}^+$ with the witness property.*

Proof. Let $\varphi_1, \varphi_2, \ldots$ be an enumeration of the set of all sentences of $\mathscr{L}^+$. We define a sequence of sets $\Gamma_1, \Gamma_2, \ldots, \Gamma_n, \ldots$ as follows:

- $\Gamma_1 = \Gamma$, and
- $$\Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{\varphi_n\} & \text{if } \Gamma_n \cup \{\varphi_n\} \text{ is consistent and} \\ & \varphi_n \text{ isn't of the form } \neg\forall v\, \psi \\ \Gamma_n \cup \{\neg\forall v\, \psi, \neg\psi[c/v]\} & \text{if } \Gamma_n \cup \{\varphi_n\} \text{ is consistent,} \\ & \varphi_n \text{ is of the form } \neg\forall v\, \psi \text{ and} \\ & c \text{ doesn't occur in } \Gamma_n \cup \{\varphi_n\} \\ \Gamma_n & \text{otherwise} \end{cases}$$

At each step $n > 1$ we add φ_{n-1} to our set just in case it doesn't generate inconsistencies. Moreover, when φ_{n-1} is of the form $\neg\forall v\, \psi$, we also add a 'fresh' witnessing expression. Note the following:

A) Each Γ_n in the sequence is consistent, by Proposition 8.1.

B) If $m \leqslant n$, $\Gamma_m \subseteq \Gamma_n$.

Let $\Gamma^+ = \bigcup\{\Gamma_n : n \in \mathbb{Z}^+\}$. Note the following:

- For each n, $\Gamma_n \subseteq \Gamma^+$ (cf. Proposition 12.1, Unit 2). In particular, $\Gamma_1 \subseteq \Gamma^+$: Γ^+ extends Γ.

- If $\neg\forall v\, \psi \in \Gamma^+$ then there must be an n s.t. $\Gamma_{n+1} = \Gamma_n \cup \{\neg\forall v\, \psi, \neg\varphi[c/v]\}$, for some individual constant c. By the previous item, $\neg\varphi[c/v] \in \Gamma^+$. Thus, Γ^+ has the witness property.

It remains to be shown that Γ^+ is maximally consistent. We do it in two steps:

1. First, we show that Γ^+ is consistent. Assume, for reductio, that it isn't. Thus, for some sentence φ there must be a derivation of φ and a derivation of $\neg\varphi$ from Γ^+. Let $\Delta \subseteq \Gamma^+$ be the set of members of Γ^+ that occur as premises in those derivations. Thus, Δ is finite and inconsistent. By construction of Γ^+, for each $\psi \in \Delta$ there must be an n s.t. $\psi \in \Gamma_n$. Let m be the maximum of those ns. By B, $\Delta \subseteq \Gamma_m$. But then, Γ_m would be inconsistent too, contradicting A.

2. Second, we show that, for every sentence φ of $\mathscr{L}^+$, if $\varphi \notin \Gamma^+$, then $\Gamma^+ \cup \{\varphi\}$ is inconsistent. Assume $\varphi \notin \Gamma^+$, so $\varphi \notin \Gamma_n$ for every n. In particular, if $\varphi = \varphi_m$, $\varphi \notin \Gamma_{m+1}$. By construction of Γ_{m+1}, $\Gamma_m \cup \{\varphi_m\}$ must be inconsistent, i.e. for every ψ, $\Gamma_m \cup \{\varphi_m\} \vdash \psi$. By monotonicity, $\Gamma^+ \cup \{\varphi_m\} \vdash \psi$ for every ψ too, that is, $\Gamma^+ \cup \{\varphi\}$ is inconsistent.

□

Definition 8.3 (term model)**.** A term model of a first-order language is a model in which every member of the domain is the denotation of a closed term of the language.

Note that, since there are only enumerably many closed terms in a first-order language, the domain of a term model is always enumerable.

Lemma 8.2 (term model)**.** *If Γ is maximally consistent and has the witness property, then it has a term model.*

Proof. For simplicity, we assume that $\mathscr{L}$ contains no function symbols, so the only closed terms occurring in members of Γ are individual constants.[1] We define $\mathcal{M}$ as follows:

1. $|\mathcal{M}| = \{[c] : c \in V(\mathscr{L})\}$

2. for each constant c of $\mathscr{L}$, $c^{\mathcal{M}} = [c]$

3. for each n-place predicate symbol $P \in V(\mathscr{L})$ and $c_1, \ldots, c_n \in V(\mathscr{L})$,

$$\langle c_1^{\mathcal{M}}, \ldots, c_n^{\mathcal{M}} \rangle \in P^{\mathcal{M}} \quad \text{iff} \quad P(c_1, \ldots, c_n) \in \Gamma$$

[1]For a proof that takes function symbols into account, see Boolos, G., Burgess, J. and Jeffrey, R. (2007) *Computability and Logic*, CUP, fifth edition, §13.3.

Note that this last clause is enough to fix the extension of P as, by 1, every member of $|\mathcal{M}|$ is the denotation of (at least) an individual constant of $\mathscr{L}$, i.e. $\mathcal{M}$ is a term model. We now show that, for every sentence φ of $\mathscr{L}$, $\mathcal{M} \models \varphi$ iff $\varphi \in \Gamma$, by induction on the complexity of φ.
Base step: φ is atomic. Thus, it's either of the form $c_1 = c_2$ or $P(c_1, \dots, c_n)$.

- If φ is of the form $c_1 = c_2$,

$$\begin{array}{rll} \mathcal{M} \models \varphi & \text{iff } c_1^{\mathcal{M}} = c_2^{\mathcal{M}} & \text{Def. of sat.} \\ & \text{iff } [c_1] = [c_2] & 2 \\ & \text{iff } \varphi \in \Gamma & \text{construction of [t]} \end{array}$$

- If φ is of the form $P(c_1, \dots, c_n)$, it follows from 3 that $\mathcal{M} \models P(c_1, \dots, c_n)$ iff $\varphi \in \Gamma$.

Inductive step: Assume every sentence of lower complexity than φ is s.t. $\mathcal{M} \models \psi$ iff $\psi \in \Gamma$.

- If φ is of the form $\neg\psi$,

$$\begin{array}{rll} \mathcal{M} \models \varphi & \text{iff } \mathcal{M} \not\models \psi & \text{Def. of sat.} \\ & \text{iff } \psi \notin \Gamma & \text{IH} \\ & \text{iff } \neg\psi \in \Gamma & \text{Exercise 8.3.1} \end{array}$$

- If φ is of the form $\forall v\, \psi$,

$$\begin{array}{rll} \mathcal{M} \models \varphi & \text{iff } \mathcal{M}, \sigma \models \psi, \text{ for each } \sigma & \text{Def. of sat.} \\ & \text{iff } \mathcal{M} \models \psi[c/v], \text{ for each } c & \text{1 and 2} \\ & \text{iff } \psi[c/v] \in \Gamma, \text{ for each } c & \text{IH} \\ & \text{iff } \forall v\, \psi \in \Gamma & \text{Exercise 8.3.3,4} \end{array}$$

The case of conditionals is left as an exercise. □

Exercise 8.6. *Finish the proof of Lemma 8.2.*

Definition 8.4 (expansion/reduct)**.** Let $V(\mathscr{L}) \subseteq V(\mathscr{L}')$, $\mathcal{M}$ be an interpretation of $\mathscr{L}$, and $\mathcal{M}'$ be an interpretation of $\mathscr{L}'$. $\mathcal{M}'$ is an expansion of $\mathcal{M}$ to $\mathscr{L}'$ (equivalently, $\mathcal{M}$ is a reduct of $\mathcal{M}'$ to $\mathscr{L}$) iff $|\mathcal{M}| = |\mathcal{M}'|$ and $\mathcal{M}$ and $\mathcal{M}'$ interpret every member of $V(\mathscr{L})$ in the same way.

Lemma 8.3 (Henkin)**.** *If* Γ *is consistent, then it has an enumerable model.*

Proof. Let Γ be a consistent set of sentences of $\mathscr{L}$. By Lemma 8.1, we can extend Γ to a maximally consistent set Γ^+ of sentences of $\mathscr{L}^+$ with the witness property. By Lemma 8.2, Γ^+ has a term model, $\mathcal{M}$. A fortiori, the reduct $\mathcal{M}'$ of $\mathcal{M}$ to $\mathscr{L}$ is a model of Γ. Moreover, since $\mathscr{L}^+$ has enumerably many terms, $|\mathcal{M}|$ is enumerable, and so is $|\mathcal{M}'|$. □

Let φ a sentence of $\mathscr{L}$.

Theorem 8.2 (completeness)**.** *If* $\Gamma \models \varphi$*, then* $\Gamma \vdash \varphi$*.*

Proof. We prove the contrapositive. Assume $\Gamma \nvdash \varphi$. Thus, $\Gamma \cup \{\neg\varphi\}$ is consistent, by Exercise 2.8. By Lemma 8.3, $\Gamma \cup \{\neg\varphi\}$ has a model. Thus, there is a model $\mathcal{M}$ s.t. $\mathcal{M} \models \Gamma$ and $\mathcal{M} \not\models \varphi$, i.e. $\Gamma \not\models \varphi$. $\square$

Chapter 9

Compactness and Löwenheim-Skolem

In this section we will establish some severely limitative results for first-order theories: the compactness and the Löwehnheim-Skolem theorems. They follow easily from the results proved in the previous section.

9.1 Compactness

Theorem 9.1 (compactness)**.** *If every finite subset of* Γ *has a model, then* Γ *has a model too.*

Proof. Assume, for contradiction, that every finite subset of Γ has a model but Γ doesn't. It follows that Γ doesn't have an enumerable model either. By Henkin's Lemma, Γ must be inconsistent, i.e. there are derivations in our calculus of sentences φ and $\neg\varphi$ from Γ. Let Δ be the set of sentences in Γ that occur in either of those derivations as premises. Thus, Δ is inconsistent so, by Corollary 8.1, it doesn't have a model. This contradicts our initial assumption, as $\Delta \subseteq \Gamma$ and Δ finite. □

Corollary 9.1. *If* $\Gamma \vDash \varphi$*, then for some finite* $\Delta \subseteq \Gamma$*,* $\Delta \vDash \varphi$*.*

Gloss: An infinite set of sentences doesn't imply more than what is implied by its finite subsets.

Exercise 9.1. *Prove Corollary 9.1.*

Proposition 9.1 (ω-rule)**.** *If* $\mathsf{PA} \cup \{\varphi(\overline{n}) : n \in \mathbb{N}\} \vDash \forall x\, \varphi(x)$*, then* $\mathsf{PA} \cup \{\varphi(\overline{n}) : n \in a\} \vDash \forall x\, \varphi(x)$*, for some finite* $a \subseteq \mathbb{N}$*.*

Gloss: If no set of finitely many numerical instances of $\varphi(x)$ – i.e. $\varphi(0), \varphi(\overline{1})$, $\varphi(\overline{2}), \ldots$ – entails $\forall x\, \varphi(x)$ against the background of PA, adding *all* the numerical instances won't do the job either.

Exercise 9.2. *Prove Proposition 9.1.*

Example 9.1 (non-standard models of arithmetic)**.** PA has a model that is not isomorphic to $\mathcal{N}$, i.e. a non-standard model.
Let $V(\mathscr{L}^c_{\mathsf{PA}}) = V(\mathscr{L}_{\mathsf{PA}}) \cup \{c\}$ and let $\Gamma = \{c \neq \overline{n} : n \in \mathbb{N}\}$. Note that, for every finite $\Delta \subseteq \Gamma$, there is an expansion $\mathcal{N}_\Delta$ of $\mathcal{N}$ to $\mathscr{L}^c_{\mathsf{PA}}$ s.t. $\mathcal{N}_\Delta \vDash \mathsf{PA} \cup \Delta$: let n be the maximum natural number such that $c \neq \overline{n} \in \Delta$ and let $c^{\mathcal{N}_\Delta} = n + 1$. By compactness, $\mathsf{PA} \cup \Gamma$ has a model too. The reduct of this model to $\mathscr{L}_{\mathsf{PA}}$ is not isomorphic with $\mathcal{N}$.

Exercise 9.3. *Show that the reduct of the model of* $\mathsf{PA} \cup \Gamma$ *to* $\mathscr{L}_{\mathsf{PA}}$ *in the previous example is not isomorphic with* $\mathcal{N}$*.*

Exercise 9.4. *Show that* $\Gamma \vdash \varphi$ *iff there is a finite subset* Δ *of* Γ *such that* $\Delta \cup \{\neg\varphi\}$ *is unsatisfiable.*

9.2 Löwehheim-Skolem

Theorem 9.2 (downwards Löwenheim-Skolem)**.** *If* Γ *has a model, it has an enumerable model.*

Proof. Assume Γ has a model. By Corollary 8.1, Γ must be consistent. By Henkin's Lemma, Γ has an enumerable model. □

Gloss: No set of sentences has only uncountable models. In other words, if expressibility of our first-order theories is to be cashed out in terms of models (*pace* Putnam), uncountability cannot be *expressed* in a first-order language. A fortiori, no first-order theory can correctly *describe* e.g. the real number structure.

Corollary 9.2 (overspill)**.** *If* Γ *has arbitrarily large finite models, then it has a denumerable model.*

Proof. Assume Γ has arbitrarily large finite models and let $\Gamma^* = \Gamma \cup \{\gamma_{\geqslant n} : n \in \mathbb{Z}^+\}$. Since every model of Γ of cardinality m is a model of $\Gamma \cup \{\gamma_{\geqslant 1}, \ldots, \gamma_{\geqslant m}\}$ and every finite subset of Γ^* is a subset of $\Gamma \cup \{\gamma_{\geqslant 1}, \ldots, \gamma_{\geqslant m}\}$ for some positive integer m, every finite subset of Γ^* has a model. By compactness, Γ^* has a model too. By downwards Löwenheim-Skolem, Γ^* has an enumerable model. Since all models of Γ^* are infinite, Γ^* must have a denumerable model. Since $\Gamma \subseteq \Gamma^*$, Γ has a denumerable model too. □

Gloss: No set of sentences has all and only finite models. In other words, the idea of finitude cannot be expressed in a first-order language.

Corollary 9.3. *If* Γ *has an infinite model, then it has a denumerable model.*

Exercise 9.5. *Prove Corollary 9.3.*

Exercise 9.6. *Show that* PA *has a denumerable model that is not isomorphic to* $\mathcal{N}$*.*

Exercise 9.7 (non-standard models of set theory)**.** *Show that, if* ZFC *is consistent, then it has a denumerable model.*

Exercise 9.8. *Show that if φ is true in all enumerable models, then $\vDash \varphi$.*

Exercise 9.9. *Show that if two sentences have the same enumerable models, then they are logically equivalent.*

Theorem 9.3 (upwards Löwenheim-Skolem)**.** *If Γ has a denumerable model, then it has an uncountable model too.*

Proof. We extend $\mathscr{L}$ to $\mathscr{L}^+$ by adding a *new* individual constant c_r for each $r \in \mathbb{R}$.[1] Let

$$\Delta = \{c_{r_1} \neq c_{r_2} : r_1 \neq r_2\}$$

and let Δ_0 be a finite subset of Δ. Clearly, each denumerable model of Γ can be expanded to a model of $\Gamma \cup \Delta_0$. Thus, every finite subset of $\Gamma \cup \Delta$ has a model. By compactness,[2] $\Gamma \cup \Delta$ must have a model too. Trivially, the domain of this model must be uncountable. □

As the downwards Löwenheim-Skolem result, its upwards version also has devastating consequences for our first-order theories.

Example 9.2. PA has uncountable models.

Example 9.3. Every consistent extension of PA has uncountable models.

Exercise 9.10. *Show that, if for every non-denumerable $\mathcal{M}$, $\mathcal{M} \nvDash \Gamma$, then Γ is unsatisfiable.*

[1]Note that $\mathscr{L}^+$ is not a first-order language strictly speaking, as it has uncountably many constants.

[2]Strictly speaking, we should first extend the proof of compactness to cover languages with uncountably many constants. This is possible alas non-trivial and will thus be omitted.

List of Symbols

Logical symbols

$=$	identity	'…is identical to …'
$\neq$	difference	'…is not identical to …'
$\neg$	negation	'it is not the case that …'
$\rightarrow$	conditional	'if …, then …'
$\wedge$	conjunction	'…and …'
$\vee$	disjunction	'…or …'
$\leftrightarrow$	biconditional	'…if and only if …'
$\forall$	universal quantifier	'for all …'
$\exists$	existential quantifier	'exists …'
$\exists!$		'exists a unique …'

Metavariables (possibly with subindexes)

$\mathscr{L}$	first-order language
c	individual constant
u, v	individual variable
f	function symbol
s, t	term
P	predicate φ, ψ, χ
wff of a first-order language	
Γ, Δ, Σ	sets of wff of a first-order language

Logical notions

$\varphi[t/v]$	the result of replacing all free occurrences of v in φ with t, if v is free for t
$\vdash$	proof-theoretic consequence/logical theoremhood

Proper names

$\mathscr{L}_{\mathsf{PA}}$	the language of first-order arithmetic
PA	first-order Peano arithmetic
$\mathscr{L}_{\in}$	the language of first-order set theory
$\mathbb{N}$	the set of natural numbers
$\mathbb{Z}$	the set of integers
$\mathbb{Z}^+$	the set of positive integers
$\mathbb{Q}$	the set of rational numbers
$\mathbb{R}$	the set of real numbers

Abbreviations

Set theoretic vocabulary and abbreviations

$\in$	membership	‘
	‘. . . is not a member of/does not belong to . . . ’ $\varnothing$	t
$\{v : \varphi\}$	the set of φ $\{t_1, \ldots, t_n\}$	t
inclusion	‘. . . is a subset of/is included in . . . ’ $\subsetneq$	p
‘. . . is a proper subset of/is properly included in . . . ’		