DAOUD SINIORA

Mathematical Logic

Lecture Notes

February 20, 2024

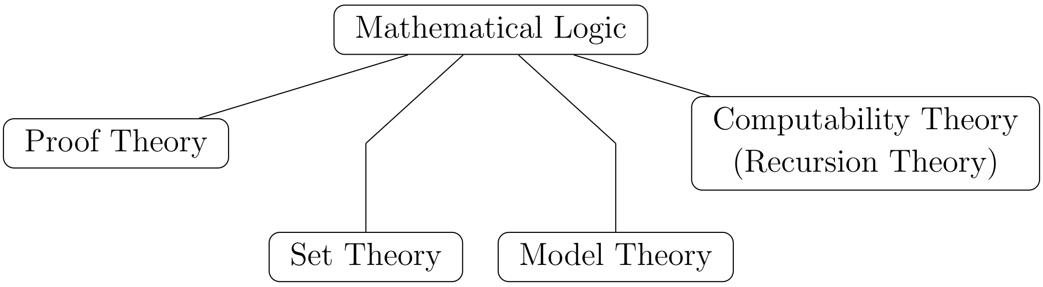
Contents

Introduction

The word ‘logic’ originates from the Greek word ‘logos’ which has a variety of meanings such as reasoning or language. Logic is the study of the laws of thought. Logic is a precise formal language with exact rules on how to combine symbols to form sentences. Aristotle (384 - 322 BC) studied rules and patterns that describe human thinking and reasoning. He focused on a particular type of deductive argument called *syllogism*. A syllogism is an argument consisting of three successive assertions: the first two are the premises and the third is the conclusion. For example:

- Every river contains water.
- The Nile is a river.
- Therefore, the Nile contains water.

Gottfried Leibniz (1646 - 1716) had the dream to reduce human thinking to machine calculation. In the nineteenth century, logic started to develop as a mathematical subject by the work of Boole, Cantor, Frege, Peano, and others. In the first half of the twentieth century the development of mathematical logic progressed by important new ideas from Russel, Zermalo, Hilbert, Löwenheim, Ramsey, Skolem, Lusin, Post, Herbrand, Gödel, Tarski, Church, Kleene, Turing, and Gentzen. The strong motivation for developing mathematical logic was the urge to provide solid foundations for mathematics. Mathematical logic is now an independent subject of its own and interacts with other areas of mathematics and computer science. In the second half of the twentieth century, mathematical logic branched into four main areas.



It is worth mentioning that one distinctive feature of mathematical logic is that *statements about mathematical objects* are studied as mathematical objects in their own right. In the same way we prove theorems about prime numbers, vector spaces, groups, rings, fields, partial orders, graphs, et cetera, we prove theorems about statements we construct by the exact rules of the logic, such statements are called *formulas.* Another distinctive practice in mathematical logic is the formalisation (formulation in a precise mathematical way) of notions used informally by mathematicians such as: *property*, *statement* in a given natural language, mathematical *structure*, *truth* of a statement in a structure, *proof* from a given set of axioms, and *algorithm*. After defining these notions in a precise mathematical way, we can start proving theorems about these formalised notions.

For example, let us express some properties and statements about the natural numbers. The set of natural numbers is $\mathbb{N} = \{0, 1, 2, \ldots\}$, the symbols $0, 1, +, \cdot, <$ denote the usual arithmetic constants, operations, and relations on $\mathbb{N}$, and the variables range over $\mathbb{N}$. With this has been said, here are examples of formal statements. To state that a natural number x is even we write $\exists y(x = y + y)$. And to express that x is prime we write

$$(1 < x) \land \forall y \forall z(x = y \cdot z \to (y = 1 \lor z = 1)).$$

Recall that the Goldbach Conjecture states that every even integer greater than 2 is a sum of two prime numbers. Let $E(x)$ abbreviates "x is even" and let $P(x)$ abbreviates "x is prime". Then we can express the Goldbach Conjecture as follows:

$$\forall x \left((E(x) \land 1 + 1 < x) \to \exists y \exists z (P(x) \land P(y) \land x = y + z) \right).$$

A formal system consists of a list of axioms (sentences in a given formal language) together with a list of deduction rules. Using formal systems we express the notion of a *formal proof.* Let Σ be a set of sentences and let σ be a sentence. We say Σ proves σ, or σ is *provable* from Σ, and write $\Sigma \vdash \sigma$, if there is a proof in the formal system which uses the axioms, deduction rules, and statements from Σ that leads to σ. We say a mathematical structure M is a *model* of σ, and write $M \models \sigma$, if the sentence σ is true in M. Moreover, M is a model of Σ, written as $M \models \Sigma$, if every sentence in Σ is true in M. Finally, we say that Σ *logically implies* σ, and write $\Sigma \models \sigma$, if σ is true in every model of Σ. Here we have two notions: "provability from a given set of axioms" and "truth in a given structure". These notions are related as described by the results below. Such results are among the most famous theorems in mathematical logic.

Completeness Theorem (Gödel, 1930). Given a formal language, let Σ be a set of sentences and let σ be a sentence. Then σ is provable from Σ if and only if σ is true in all models of Σ. That is,

$$\Sigma \vdash \sigma \text{ if and only if } \Sigma \models \sigma.$$

Compactness Theorem (Gödel, Mal'cev). Let Σ be a set of sentences in some language. Then Σ has a model if and only if every finite subset of Σ has a model.

In the beginning of the twentieth century, the German mathematician David Hilbert proposed a solution to what is known as the *foundational crises of mathematics* where paradoxes and inconsistencies arose in the face of earlier attempts to put mathematics on strong foundations. Hilbert's proposal, known as *Hilbert's Program*, was to formulate a finite, complete set of axioms for all theories of mathematics, and to provide a proof that these axioms are consistent (do not lead to a contradiction). Hilbert's famous words that he spoke in a speech in 1930 were:

Wir müssen wissen. Wir werden wissen.
We must know. We will know.

Kurt Gödel played a dramatic role in the progress of Hilbert's program. In 1931, he showed that Hilbert's program was unattainable for vital areas of mathematics. Gödel's first incompleteness theorem states that any consistent formal system with a computable set of axioms which is capable of expressing arithmetic can never be complete (there will be a sentence which is neither provable from the axioms nor its negation is provable from the axioms). Gödel's second incompleteness theorem states that such a system cannot prove its own consistency.

Incompleteness Theorem (Gödel, 1931). There is a sentence in the language of arithmetic that is true in the structure $(\mathbb{N}, 0, 1, +, \cdot, <)$, but not provable from the Peano axioms.

References

[1] René Cori and Daniel Lascar (2000), Translated by Donald Pelletier, *Mathematical Logic: A Course With Exercises*, Oxford University Press.

[2] Derek Goldrei (2005), *Propositional and Predicate Calculus: A Model of Argument*, Springer-Verlag London.

[3] Jochen Koenigsmann, *Logic Course (B1.1) Slides*, University of Oxford.

[4] Anand Pillay, *Models and Sets (MATH3120) Lecture Notes*, University of Leeds.

[5] Lou van den Dries (2019), *Mathematical Logic Lecture Notes*, University of Illinois Urbana-Champaign.

[6] A. G. Hamilton (1988), *Logic for Mathematicians*, Cambridge University Press.

[7] Herbert Enderton (2001), *A Mathematical Introduction to Logic*, Harcourt/ Academic Press.

[8] Martin Hils and François Loeser (2019), *A First Journey through Logic*, American Mathematical Society.

Chapter 1

Propositional Logic

In propositional logic we construct formal strings of symbols called propositional formulas. The building blocks of formulas are propositional variables, which are then assembled together according to precise rules using logical connectives such as negation, conjunction, disjunction, implication, and equivalence. This construction process of formulas constitute the *syntax* of propositional logic. On the other part, the *semantics* of the logic, we will give meaning to these formulas. We need to decide whether a formula is true or false based on the truth values of the propositional variables appearing in the formula. We will see that the truth value of a formula depends on the formal construction of the formula. We then discuss a semantic and a syntactic approach to study the notion of a formula being implied from a set of formulas. The semantic notion is called *logical consequence*. To define the syntactic notion we need to introduce *proof systems* to be able to define a formal *derivation* of a formula from a set of formulas. We show that these two notions: logical consequence and derivability, are equivalent by proving the soundness theorem and the completeness theorem for propositional logic.

1.1 Syntax of Propositional Logic

1.1.1 Words over an Alphabet

Let $\mathcal{A}$ be a nonempty set of symbols, which we will call the *alphabet*. A *word* over the alphabet $\mathcal{A}$ is a finite sequence of symbols from $\mathcal{A}$. So a word w has the form

$$a_1\, a_2\, a_3\, \ldots\, a_n$$

where each a_i belongs to $\mathcal{A}$ and n is a positive integer. The integer n is called the length of the word w, and we write $l[w] = n$. Notice that the length of w is the number of symbols in w counting repetitions. Strictly speaking, a word $w = a_1 a_2 \ldots a_n$ over $\mathcal{A}$ is a function $w : \{1, 2, \ldots, n\} \to \mathcal{A}$ given by $w(i) = a_i$.

There is the *empty word* which is denoted by λ and has length 0. Two words $w_1 = a_1a_2 \ldots a_n$ and $w_2 = b_1b_2 \ldots b_m$ are *equal* if $n = m$ and $a_i = b_i$ for every $1 \leq i \leq n$.

The set of all words is denoted by $\mathcal{W}(\mathcal{A})$ or by $\mathcal{A}^*$ (Kleene star). So $\mathcal{W}(\mathcal{A})$ is the set of all finite sequences of symbols from $\mathcal{A}$. Let $w_1 = a_1a_2 \ldots a_n$ and $w_2 = b_1b_2 \ldots b_m$ be two words over $\mathcal{A}$. The *concatenation* of w_1 and w_2 is the word

$$w_1w_2 = a_1\, a_2\, \ldots\, a_n\, b_1\, b_2\, \ldots\, b_m.$$

In other words, the concatenation of w_1 and w_2 is the word $w = w_1w_2$ given by

$$w(i) = \begin{cases} a_i & \text{if } 1 \leq i \leq n; \\ b_{i-n} & \text{if } n+1 \leq i \leq n+m. \end{cases}$$

A word w' is called an *initial segment* of a word w if there exists a word w'' such that $w = w'w''$. Thus, an initial segment of $w = a_1a_2 \ldots a_n$ is either the empty word or a word of the form $a_1a_2 \ldots a_k$ where $1 \leq k \leq n$. Clearly, the empty word and w itself are both initial segments of w. An initial segment of w is called *proper* if it is different from w. When a letter α from the alphabet appears in a word $w = a_1\, a_2\, a_3\, \ldots\, a_n$, we say that α has an occurrence in w and the positions where it appears are called the *occurrences* of α in w. For example, take the alphabet $\mathcal{A} = \{a, b, c\}$, and take the word $w = bcabacbaccac$. Then the letter a has 4 occurrences in w at positions $3, 5, 8, 11$, and the letter b has 3 occurrences in w at positions $1, 4, 7$.

Lemma 1.1.1. *For all words w, w_1, w_2, w_3, and w_4, the following hold.*

- $l[w_1w_2] = l[w_1] + l[w_2]$.
- $(w_1w_2)w_3 = w_1(w_2w_3)$. *(concatenation is associative)*
- *If $ww_1 = ww_2$, then $w_1 = w_2$.* *(left cancellation)*
- *If $w_1w = w_2w$, then $w_1 = w_2$.* *(right cancellation)*
- *If $w_1w_2 = w_3w_4$, then either w_1 is an initial segment of w_3 or else w_3 is an initial segment of w_1.*
- *If w_1 is an initial segment of w_2 and w_2 is an initial segment of w_1, then $w_1 = w_2$.*
- *If $\mathcal{A}$ is finite or countably infinite, then $\mathcal{W}(\mathcal{A})$ is countably infinite.*

Remark. The set $\mathcal{W}(\mathcal{A})$ of all words over $\mathcal{A}$ together with the concatenation operation form a *monoid* where the identity element is the empty word.

1.1.2 Propositional Formulas

In this section we introduce the syntax of propositional logic, we describe precisely how sentences in propositional logic are constructed. Sentences in propositional logic will be strings of symbols of a certain kind, and they are called *propositional formulas*. The symbols we start with are called the *alphabet* of propositional logic and they consist of three types of symbols: *propositional variables*, *propositional connectives*, and parentheses. The set $\mathbf{P}$ of propositional variables is a nonempty, finite or infinite, set of symbols. The propositional variables will usually be denoted by

$$p,\ q,\ r,\ p_0,\ p_1,\ p_2,\ \dots\ ,\ q_0,\ q_1,\ \dots$$

Next, there are five symbols for propositional connectives.

Symbol	$\neg$	$\wedge$	$\vee$	$\rightarrow$	$\leftrightarrow$
Read as	'not'	'and'	'or'	'implies'	'is equivalent to'
Called	negation	conjunction	disjunction	implication	equivalence
Arity	unary	binary	binary	binary	binary

Finally, there are two parentheses symbols.

Symbol	)	(
Called	closing parenthesis	opening parenthesis

We assume that the three sets: $\mathbf{P}$, $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\}$, and $\{\,),\,(\,\}$ are pairwise disjoint. The alphabet of propositional logic is thus the set

$$\mathcal{A} = \mathbf{P} \cup \{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\} \cup \{\,),\,(\,\}\,.$$

Propositional formulas are special words over the alphabet $\mathcal{A}$. The syntax of constructing these sentences is as follows.

We define *propositional formulas* inductively as follows.

(i) Any propositional variable is a propositional formula.

(ii) If φ is a propositional formula, then so is $\neg\varphi$.

(iii) If φ and ψ are propositional formulas, then so are $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, and $(\varphi \leftrightarrow \psi)$.

(iv) Nothing else is a propositional formula.

More precisely we define, by induction, a sequence of subsets of $\mathcal{W}(\mathcal{A})$ and then define the set of all propositional formulas to be the union of this chain of subsets as shown below.

Definition.

- We set $\mathcal{F}_0 = \mathbf{P}$.
- For each natural number n, we define

$$\begin{aligned}\mathcal{F}_{n+1} = \mathcal{F}_n \ \cup\ \{\neg\varphi \mid \varphi \in \mathcal{F}_n\} \ \cup\ \{(\varphi \wedge \psi) \mid \varphi, \psi \in \mathcal{F}_n\} \cup \{(\varphi \vee \psi) \mid \varphi, \psi \in \mathcal{F}_n\} \\ \cup\ \{(\varphi \to \psi) \mid \varphi, \psi \in \mathcal{F}_n\} \ \cup\ \{(\varphi \leftrightarrow \psi) \mid \varphi, \psi \in \mathcal{F}_n\}.\end{aligned}$$

- We define the set $\mathcal{F}$ of all propositional formulas to be

$$\mathcal{F} = \bigcup_{n \in \mathbb{N}} \mathcal{F}_n.$$

Remark. Observe that $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$.

The logical connectives induce operations (functions) on the set $\mathcal{W}(\mathcal{A})$ of all words over the alphabet of propositional logic. The operation associated with the negation symbol is $\neg : \mathcal{W}(\mathcal{A}) \to \mathcal{W}(\mathcal{A})$ given by

$$w \mapsto \neg w$$

for any word $w \in \mathcal{W}(\mathcal{A})$.

Moreover, there are four more operations from the set $\mathcal{W}(\mathcal{A}) \times \mathcal{W}(\mathcal{A})$ to the set $\mathcal{W}(\mathcal{A})$; one for each binary logical connective. These functions are given by:

$$\begin{aligned}[w, v] &\mapsto (w \wedge v) \\ [w, v] &\mapsto (w \vee v) \\ [w, v] &\mapsto (w \to v) \\ [w, v] &\mapsto (w \leftrightarrow v)\end{aligned}$$

for every pair $[w, v]$ of words over $\mathcal{A}$. A subset $\mathcal{V} \subseteq \mathcal{W}(\mathcal{A})$ is said to be *closed* under taking a logical symbol if it is closed under the operation associated with that symbol. For instance, $\mathcal{V}$ is closed under taking negations if whenever $w \in \mathcal{V}$, then the word $\neg w$ belongs to $\mathcal{V}$ as well. And $\mathcal{V}$ is closed under taking conjunctions if whenever $w, v \in \mathcal{V}$, then the word $(w \wedge v)$ belongs to $\mathcal{V}$ as well.

Theorem 1.1.2. *The set $\mathcal{F}$ of all propositional formulas is the smallest subset of $\mathcal{W}(\mathcal{A})$ which contains $\mathbf{P}$ and is closed under taking negations, conjunctions, disjunctions, implications, and equivalences.*

Proof. Let $\mathcal{S}$ be the smallest subset of $\mathcal{W}(\mathcal{A})$ which contains $\mathbf{P}$ and is closed under taking the propositional connectives. This means that if X is a subset of $\mathcal{W}(\mathcal{A})$ which contains $\mathbf{P}$ and is closed under taking the propositional connectives, then $\mathcal{S} \subseteq X$.

Clearly, $\mathcal{F}$ contains $\mathbf{P}$. Suppose $\varphi, \psi \in \mathcal{F}$, then there are $m, n \in \mathbb{N}$ such that $\varphi \in \mathcal{F}_m$ and $\psi \in \mathcal{F}_n$. By definition, $\neg\varphi \in \mathcal{F}_{m+1}$, and so $\neg\varphi \in \mathcal{F}$. Thus $\mathcal{F}$ is closed under taking negations. Let $k = \max(m, n)$. Then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \to \psi)$, and $(\varphi \leftrightarrow \psi)$ all belong to $\mathcal{F}_{k+1}$, and so they belong to $\mathcal{F}$. Therefore, $\mathcal{F}$ contains $\mathbf{P}$ and is closed under taking the propositional connectives. It follows that $\mathcal{S} \subseteq \mathcal{F}$.

Next, we show by induction that $\mathcal{F} \subseteq \mathcal{S}$. As $\mathcal{S}$ contains $\mathbf{P}$, it follows that $\mathcal{F}_0 \subseteq \mathcal{S}$. Now suppose that $\mathcal{F}_n \subseteq \mathcal{S}$. Since $\mathcal{S}$ is closed under taking the propositional connectives, it follows that every formula in $\mathcal{F}_{n+1}$ belongs to $\mathcal{S}$. Thus, $\mathcal{F}_{n+1} \subseteq \mathcal{S}$. Therefore, we have shown that $\mathcal{F}_n \subseteq \mathcal{S}$ for every $n \in \mathbb{N}$. It follows that $\mathcal{F} \subseteq \mathcal{S}$ as desired. Therefore, $\mathcal{F} = \mathcal{S}$. ■

Definition. The *height* of a formula $\varphi \in \mathcal{F}$ is the least integer n such that $\varphi \in \mathcal{F}_n$. The height of φ is denoted by $h[\varphi]$.

Example. Let p and q be propositional variables.

- $h[p] = 0$.
- $h[((p \wedge q) \wedge p)] = 2$.
- $h[\neg\neg\neg\neg p] = 4$.
- Let $\varphi := ((p \vee q) \wedge (q \to p))$. Then $h[\varphi] = 2$. ♠

Remark.

- The set $\mathcal{F}_n$ is the set of propositional formulas of height at most n.
- The set $\mathcal{F}_{n+1} \setminus \mathcal{F}_n$ is the set of all formulas of height exactly $n + 1$.
- For any formula φ, we have $h[\neg\varphi] \leq h[\varphi] + 1$.
- For any formulas φ and ψ, we have $h[(\varphi \diamond \psi)] \leq \max(h[\varphi], h[\psi]) + 1$, where $\diamond$ denotes any of the binary propositional connectives.

1.1.3 Proofs by Induction on Formulas

Suppose we want to show that every propositional formula φ has property $\mathcal{Y}$. To achieve this we use mathematical induction on the height of the formula φ.

- **Base case.** Show that every formula in $\mathcal{F}_0$ has property $\mathcal{Y}$.
- **Induction step.** Fix some natural number n. Suppose that every formula belonging to $\mathcal{F}_n$ has property $\mathcal{Y}$. Then show that property $\mathcal{Y}$ holds for every formula belonging to $\mathcal{F}_{n+1}$.

Alternatively, we may proceed as follows.

Lemma 1.1.3 (Induction on formulas)**.** *To show that every propositional formula has property $\mathcal{Y}$, it is sufficient to show the following.*

(i) Every propositional variable in $\mathbf{P}$ *has property* $\mathcal{Y}$*.*

(ii) Whenever a formula φ *satisfies property* $\mathcal{Y}$*, then so does the formula* $\neg\varphi$*.*

(iii) Whenever formulas φ *and* ψ *satisfy property* $\mathcal{Y}$*, then so do the formulas* $(\varphi\wedge\psi)$*,* $(\varphi \vee \psi)$*,* $(\varphi \to \psi)$*, and* $(\varphi \leftrightarrow \psi)$*.*

Proof. Suppose that the three conditions in the statement are satisfied. Let X be the set of formulas which have property $\mathcal{Y}$. In other words,

$$X = \{\, \varphi \in \mathcal{F} \mid \varphi \text{ has property } \mathcal{Y}\,\}\,.$$

So $X \subseteq \mathcal{F}$. Moreover, by condition (i) we get that $\mathbf{P} \subseteq X$. By condition (ii) , the set X is closed under taking negations. By condition (iii), the set X is closed under taking conjunctions, disjunctions, implications, and equivalences. By Theorem 1.1.2, $\mathcal{F} \subseteq X$. Therefore, $X = \mathcal{F}$, that is, every formula has property $\mathcal{Y}$. ■

Theorem 1.1.4. *The height of a formula is always strictly less than its length. That is, for any propositional formula* φ*, we have that*

$$h[\varphi] < l[\varphi].$$

Proof. Let us say that a formula φ has property $\mathcal{Y}$ when its height is strictly less than its length. We need to show that every formula φ has property $\mathcal{Y}$. We proceed by induction on the formulas. We first show that this property holds for propositional variables. Let $p \in \mathbf{P}$. As $p \in \mathcal{F}_0$, we get that $h[p] = 0$. Clearly, $l[p] = 1$. So $h[p] < l[p]$; the inequality is verified for p.

Now suppose that φ is a formula that satisfies $h[\varphi] < l[\varphi]$. Then,

$$h[\neg\varphi] \le h[\varphi] + 1 < l[\varphi] + 1 = l[\neg\varphi].$$

This shows that $\neg\varphi$ has property $\mathcal{Y}$.

Next, suppose that φ and ψ are formulas that satisfy $h[\varphi] < l[\varphi]$ and $h[\psi] < l[\psi]$. And let $\diamond$ be one of the binary propositional connectives. Then,

$$h[(\varphi \diamond \psi)] \le \max(h[\varphi], h[\psi]) + 1 \le h[\varphi] + h[\psi] + 1 < l[\varphi] + l[\psi] + 3 = l[(\varphi \diamond \psi)].$$

Thus, the formula $(\varphi \diamond \psi)$ satisfies property $\mathcal{Y}$ and the proof is finished. ■

Corollary 1.1.5. *The empty word is not a formula. The only formulas of length 1 are the propositional variables.*

1.1.4 Unique Decomposition Theorem

Lemma 1.1.6. *For any propositional formula $\theta \in \mathcal{F}$, exactly one of the following three cases arises:*

(i) θ is a propositional variable.

(ii) There exists a formula φ such that $\theta = \neg\varphi$.

(iii) There are formulas φ and ψ and a binary connective $\diamond$ such that $\theta = (\varphi \diamond \psi)$.

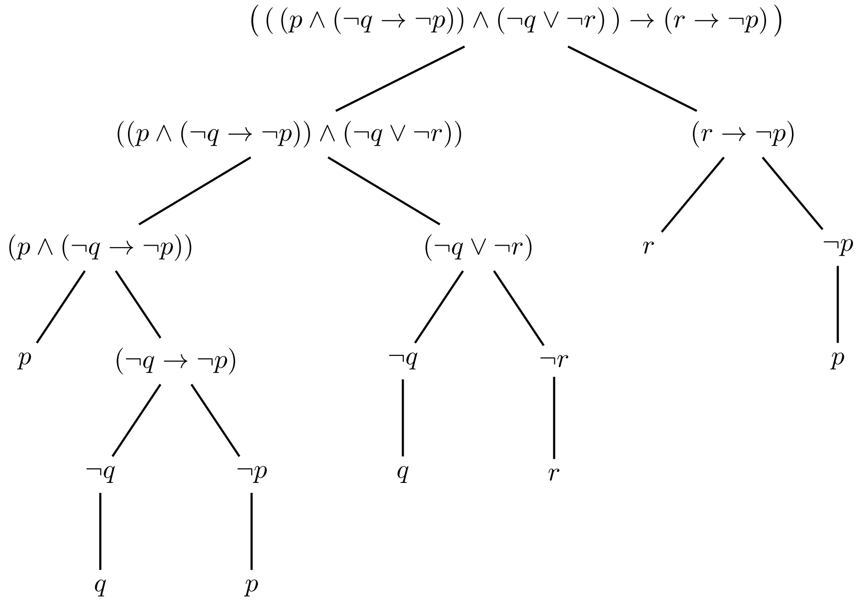
Proof. First observe that no two cases can arise at the same time because the first symbol of case 1 is a propositional variable, the first symbol of case 2 is the negation symbol $\neg$, and the first symbol of case 3 is the opening parenthesis (.

Now let θ be a formula and so $\theta \in \mathcal{F}$. If $\theta \in \mathcal{F}_0$, then θ is a propositional variable, and this is the first case. Otherwise, if $\theta \notin \mathcal{F}_0$, there exists a natural number n such that $\theta \in \mathcal{F}_{n+1} \setminus \mathcal{F}_n$. By definition of $\mathcal{F}_{n+1}$ and since $\theta \notin \mathcal{F}_n$, either there is a formula $\varphi \in \mathcal{F}_n$ such that $\theta = \neg\varphi$ and so the second case arises, or else there are formulas $\varphi, \psi \in \mathcal{F}_n$ and a binary connective $\diamond$ such that $\theta = (\varphi \diamond \psi)$ and so the third case occurs. ■

Lemma 1.1.7. *For any $w \in \mathcal{W}(\mathcal{A})$, if w is not a formula, then neither is $\neg w$.*

Proof. We will show the contrapositive. Let w be a word over the alphabet $\mathcal{A}$, and suppose that $\neg w$ is a formula. Thus, one of the three cases of Lemma 1.1.6 must occur. As the first symbol of $\neg w$ is the negation symbol $\neg$, it follows that $\neg w$ is neither a propositional variable nor a formula of the form $(\varphi \diamond \psi)$. Therefore, it must be the case that there is a formula $\theta \in \mathcal{F}$ such that $\neg w = \neg\theta$. By left cancellation, it follows that $w = \theta$, which shows that w is a formula. ■

The tree below is called a *decomposition tree* of the formula at the very top (the root of the tree). The tree shows the steps of constructing that formula, and we can see that the height of that formula is at most 5. Here, 5 is the length of a longest branch from a leaf to the root.


$(((p \wedge (\neg q \to \neg p)) \wedge (\neg q \vee \neg r)) \to (r \to \neg p))$
$((p \wedge (\neg q \to \neg p)) \wedge (\neg q \vee \neg r))$
$(r \to \neg p)$
$(p \wedge (\neg q \to \neg p))$
$(\neg q \vee \neg r)$
r
$\neg p$
p
$(\neg q \to \neg p)$
$\neg q$
$\neg r$
p
$\neg q$
$\neg p$
q
r
q
p

The decomposition tree above shows that the word w given by the sequence

$$(((p \wedge (\neg q \to \neg p)) \wedge (\neg q \vee \neg r)) \to (r \to \neg p))$$

is a formula because it was obtained by starting from the propositional variables p, q, and r (see the leaves of the tree), and then by applying propositional connectives (as in the definition of $\mathcal{F}$) finitely many times. The tree shows that $w \in \mathcal{F}_5$ and so w is a formula and $h[w] \leq 5$. At this point, we feel hesitant to claim that $h[w] = 5$, as there may be another decomposition which constructs the formula w in $\mathcal{F}_4$ or even before.

One can prove by induction on formulas that any formula has a decomposition tree. Here is an informal discussion. By definition of a formula, a word θ is a formula if it belongs to the set $\mathcal{F}$. Since $\mathcal{F} = \bigcup_{n \in \mathbb{N}} \mathcal{F}_n$, we can find the least natural number n such that $\theta \in \mathcal{F}_n$. If $n = 0$, then θ is a propositional variable (a leaf of the tree). Otherwise, $n \geq 1$ and so by definition of $\mathcal{F}_n$, either there is a formula $\varphi \in \mathcal{F}_{n-1}$ such that $\theta = \neg\varphi$ (in such case there will be one branch leaving θ to φ in the level below) or else there are formulas $\varphi, \psi \in \mathcal{F}_{n-1}$ and a binary connective $\diamond$ such that $\theta = (\varphi \diamond \psi)$ (here there are two branches leaving θ, one to φ and the other to ψ in the level below). Then we repeat this to the new shorter formulas until all the leaves of the tree are propositional variables. Soon we will show that such a decomposition tree is unique.

Let w be a word in $\mathcal{W}(\mathcal{A})$. We denote by $\mathrm{o}[w]$ the number of opening parentheses that occur in w. Similarly, we denote by $\mathrm{c}[w]$ the number of closing parentheses that occur in w.

Theorem 1.1.8. *In any propositional formula, the number of opening parentheses is equal to the number of closing parentheses.*

Proof. We proceed by induction on the formulas. Let p be any propositional variable. Then $\mathrm{o}[p] = 0 = \mathrm{c}[p]$.

Suppose that φ is a formula such that $\mathrm{o}[\varphi] = \mathrm{c}[\varphi]$. Then,

$$\mathrm{o}[\neg\varphi] = \mathrm{o}[\varphi] = \mathrm{c}[\varphi] = \mathrm{c}[\neg\varphi].$$

Suppose that φ and ψ are formulas such that $\mathrm{o}[\varphi] = \mathrm{c}[\varphi]$ and $\mathrm{o}[\psi] = \mathrm{c}[\psi]$. Let $\diamond$ be any symbol of the binary propositional connectives. Then,

$$\mathrm{o}[(\varphi \diamond \psi)] = 1 + \mathrm{o}[\varphi] + \mathrm{o}[\psi] = 1 + \mathrm{c}[\varphi] + \mathrm{c}[\psi] = \mathrm{c}[(\varphi \diamond \psi)].$$

Therefore, we have shown that $\mathrm{o}[\varphi] = \mathrm{c}[\varphi]$ for any propositional formula φ. ■

Theorem 1.1.9. *For any formula φ in $\mathcal{F}$ and any word w in $\mathcal{W}(\mathcal{A})$, if w is an initial segment of φ, then $\mathrm{o}[w] \geq \mathrm{c}[w]$.*

Proof. We proceed by induction on the formulas. Let p be any propositional variable and let w be an initial segment of p. It follows that w is either the empty word or p itself. In each case, we have $\mathrm{o}[w] = 0$ and $\mathrm{c}[w] = 0$. So $\mathrm{o}[w] \geq \mathrm{c}[w]$.

Let φ be a formula such that every initial segment w of φ satisfies $\mathrm{o}[w] \geq \mathrm{c}[w]$. Now let v be an initial segment of $\neg\varphi$.

- If $v = \lambda$ (the empty word), then $\mathrm{o}[\lambda] = 0$ and $\mathrm{c}[\lambda] = 0$. So $\mathrm{o}[v] \geq \mathrm{c}[v]$.
- If v is not the empty word, then $v = \neg w$ for some initial segment w of φ. It follows that
$$\mathrm{o}[v] = \mathrm{o}[\neg w] = \mathrm{o}[w] \geq \mathrm{c}[w] = \mathrm{c}[\neg w] = \mathrm{c}[v].$$

Suppose that φ and ψ are formulas such that all of their initial segments have at least as many opening parentheses as closing parentheses. Let $\diamond$ be any of the binary propositional connectives. Now choose any initial segment v of $(\varphi \diamond \psi)$. Four cases can arise.

- If $v = \lambda$, then $\mathrm{o}[\lambda] = 0$ and $\mathrm{c}[\lambda] = 0$. So $\mathrm{o}[v] \geq \mathrm{c}[v]$.
- If $v = (w$ for some initial segment w of φ. It follows that
$$\mathrm{o}[v] = \mathrm{o}[(w] = 1 + \mathrm{o}[w] > \mathrm{o}[w] \geq \mathrm{c}[w] = \mathrm{c}[(w] = \mathrm{c}[v].$$
- If $v = (\,\varphi\,\diamond\,w'$ for some initial segment w' of ψ. It follows that
$$\mathrm{o}[v] = \mathrm{o}[(\varphi\,\diamond\,w'] = 1 + \mathrm{o}[\varphi] + \mathrm{o}[w'] > \mathrm{o}[\varphi] + \mathrm{o}[w'] \geq \mathrm{c}[\varphi] + \mathrm{c}[w'] = \mathrm{c}[v].$$
- If $v = (\varphi \diamond \psi)$, then $\mathrm{o}[v] = \mathrm{c}[v]$ because v is a formula in this case.

Thus in all cases we got that $\mathrm{o}[v] \geq \mathrm{c}[v]$. ■

Lemma 1.1.10. *Suppose that θ is a formula of the form $(\varphi \diamond \psi)$ where φ and ψ are arbitrary formulas and $\diamond$ is a symbol from the binary propositional connectives. Then, for every word $w \in \mathcal{W}(\mathcal{A})$ which is a nonempty proper initial segment of θ, we have that*

$$o[w] > c[w].$$

Theorem 1.1.11. *For any formula $\varphi \in \mathcal{F}$ and for any word $w \in \mathcal{W}(\mathcal{A})$, if w is an initial segment of φ different from φ, then w is not a formula.*

Proof. We have seen previously that the empty word is not a formula. It remains to show that any nonempty proper initial segment of a formula is not a formula. We proceed by induction on the formulas.

Propositional variables have no nonempty proper initial segments.

Suppose that φ is a formula such that every nonempty proper initial segment of φ is not a formula. Let w be a nonempty proper initial segment of $\neg\varphi$. It follows that either $w = \neg$ which is not a formula because the only formulas of length 1 are the propositional variables or $w = \neg v$ where v is a nonempty proper initial segment of φ. By induction hypothesis, v is not a formula, and so by Lemma 1.1.7, $\neg v$ is not a formula as well. So w is not a formula.

Lastly, suppose that φ and ψ are formulas. Let w be a nonempty proper initial segment of $(\varphi \diamond \psi)$ where $\diamond$ is a binary propositional connective. By Lemma 1.1.10, we get that $\mathrm{o}[w] > \mathrm{c}[w]$. We conclude that w is not a formula because if it were we must have $\mathrm{o}[w] = \mathrm{c}[w]$ by Theorem 1.1.8. ■

Theorem 1.1.12 (Unique decomposition theorem)**.** *For any formula $\theta \in \mathcal{F}$, exactly one of the following three cases arises:*

(i) θ is a propositional variable.

(ii) There is a unique formula φ such that $\theta = \neg\varphi$.

(iii) There are unique formulas φ and ψ and a unique binary connective $\diamond$ such that $\theta = (\varphi \diamond \psi)$.

Proof. By Lemma 1.1.6, it remains to show the uniqueness of the decomposition in case 2 and case 3.

For case 2, suppose that $\theta = \neg\varphi$ and $\theta = \neg\varphi'$ for some formulas φ and φ'. Thus, we have that $\neg\varphi = \neg\varphi'$. By left cancellation, it follows that $\varphi = \varphi'$.

For case 3, suppose that $\theta = (\varphi \diamond \psi)$ and $\theta = (\varphi' \vartriangle \psi')$ where φ, ψ, φ', and ψ' are formulas and $\diamond$ and $\vartriangle$ are binary connectives. Therefore, we get $(\varphi \diamond \psi) = (\varphi' \vartriangle \psi')$. By left and right cancellation, we get $\varphi \diamond \psi = \varphi' \vartriangle \psi'$. We conclude that one of the formulas φ and φ' is an initial segment of the other. Since both φ and φ' are formulas, by Theorem 1.1.11, it must be that $\varphi = \varphi'$. Since $\varphi \diamond \psi = \varphi' \vartriangle \psi'$ and $\varphi = \varphi'$, by left cancellation, it follows that $\diamond \psi = \vartriangle \psi'$. We conclude that the first symbol of the word $\diamond \psi$ is identical to the first symbol of the word $\vartriangle \psi'$, so $\diamond$ and $\vartriangle$ are identical. This forces ψ and ψ' to be identical as well. ■

Corollary 1.1.13. *The decomposition tree of any formula is unique.*

Corollary 1.1.14. *Let φ and ψ be propositional formulas. Then*

- $h[\neg\varphi] = h[\varphi] + 1$.
- $h[(\varphi \diamond \psi)] = \max(h[\varphi],\, h[\psi]) + 1$.

Proof. We prove the second equality. Let $\theta = (\varphi \diamond \psi)$. Since θ is not a propositional variable, it follows that $h[\theta] = n + 1$ for some natural number n. This means that $\theta \in \mathcal{F}_{n+1}$ but $\theta \notin \mathcal{F}_n$. By definition of $\mathcal{F}_{n+1}$ and since θ begins with an opening parenthesis, there are formulas φ' and ψ' in $\mathcal{F}_n$ and a binary connective $\vartriangle$ such that $\theta = (\varphi' \vartriangle \psi')$. The unique decomposition theorem implies that $\varphi = \varphi'$, $\psi = \psi'$, and $\diamond = \vartriangle$. Thus, both φ and ψ belong to $\mathcal{F}_n$.

Claim. $h[\varphi] = n$ or $h[\psi] = n$.
Proof of the claim. Since $\varphi, \psi \in \mathcal{F}_n$, it follows that $h[\varphi] \leq n$ and $h[\psi] \leq n$. If $n = 0$, then all formulas in $\mathcal{F}_0$ have height 0, and we are done. Otherwise, suppose $n \geq 1$, and for the sake of contradiction assume that $h[\varphi] < n$ and $h[\psi] < n$. It follows that φ and ψ belong to $\mathcal{F}_{n-1}$, and consequently θ would belong to $\mathcal{F}_n$, a contradiction! It follows that $h[\varphi] = n$ or $h[\psi] = n$ establishing the claim. □

By the claim we conclude that $\max(h[\varphi], h[\psi]) = n$. Therefore, we have that $h[(\varphi \diamond \psi)] = h[\theta] = n + 1 = \max(h[\varphi], h[\psi]) + 1$, as desired. ■

The following definition is an example of "definition by induction".

Definition. For each propositional formula $\theta \in \mathcal{F}$ we associate a set $\mathrm{sf}[\theta]$, called the set of *subformulas* of θ, as follows.

- If θ is a propositional variable, then $\mathrm{sf}[\theta] = \{\theta\}$.
- If $\theta = \neg\varphi$ for some formula φ, then $\mathrm{sf}[\theta] = \mathrm{sf}[\varphi] \cup \{\theta\}$.
- If $\theta = (\varphi \diamond \psi)$ for some formulas φ and ψ, then $\mathrm{sf}[\theta] = \mathrm{sf}[\varphi] \cup \mathrm{sf}[\psi] \cup \{\theta\}$.

Remark. Every subformula is itself a formula. The subformulas of a formula θ are exactly the nodes that appear in the decomposition tree of θ.

Example. Let $\theta = ((p \wedge q) \leftrightarrow (q \to r))$. Then θ has 6 subformulas, here are they:

$$\begin{aligned} \mathrm{sf}[\theta] &= \mathrm{sf}[((p \wedge q) \leftrightarrow (q \to r))] \\ &= \mathrm{sf}[(p \wedge q)] \cup \mathrm{sf}[(q \to r)] \cup \{((p \wedge q) \leftrightarrow (q \to r))\} \\ &= \mathrm{sf}[p] \cup \mathrm{sf}[q] \cup \{(p \wedge q)\} \cup \mathrm{sf}[q] \cup \mathrm{sf}[r] \cup \{(q \to r)\} \cup \{((p \wedge q) \leftrightarrow (q \to r))\} \\ &= \{p\} \cup \{q\} \cup \{(p \wedge q)\} \cup \{q\} \cup \{r\} \cup \{(q \to r)\} \cup \{((p \wedge q) \leftrightarrow (q \to r))\} \\ &= \{p,\, q,\, r,\, (p \wedge q),\, (q \to r),\, ((p \wedge q) \leftrightarrow (q \to r))\}. \end{aligned}$$

♠

1.1.5 Substitutions in Propositional Formulas

Let φ be a propositional formula and let $p_1, p_2, \ldots, p_n$ be propositional variables that are pairwise distinct. We will use the notation $\varphi(p_1, p_2, \ldots, p_n)$ to say that the propositional variables that occur in φ are *among* $p_1, p_2, \ldots, p_n$. In other words, the formula φ has no propositional variables other than $p_1, p_2, \ldots, p_n$. For example, for the formula $\varphi = (p \to (q \vee p))$ we could write $\varphi(p, q)$, but also we may write $\varphi(p, q, r, s)$ or $\varphi(t, v, p, q, r, s)$.

Suppose we are given a formula $\varphi(p_1, p_2, \ldots, p_n)$ together with formulas $\psi_1, \psi_2, \ldots, \psi_n$. We form a new word by substituting the formula ψ_i for every occurrence of the variable p_i in the formula φ. We substitute the formulas $\psi_1, \psi_2, \ldots, \psi_n$ for the variables $p_1, p_2, \ldots, p_n$ in the formula φ *simultaneously.* The resultant word is denoted by

$$\varphi\big(\psi_1/p_1,\ \psi_2/p_2,\ \ldots,\ \psi_n/p_n\big)$$

and it is read as "φ where ψ_1 replaces p_1, ψ_2 replaces p_2, and so on, ψ_n replaces p_n". We may also denote the new word by $\varphi(\psi_1, \psi_2, \ldots, \psi_n)$.

Example.

- Let $\theta = (p \wedge q)$, $\psi_1 = (p \vee q)$, and $\psi_2 = (p \to q)$. Then $\theta(\psi_1/p,\ \psi_2/q)$ or $\theta(\psi_1,\ \psi_2)$ is the word
$$((p \vee q) \wedge (p \to q)).$$
- Consider the formula $\varphi = \varphi(p, q) = (p \to (q \vee p))$ and the formula $\psi = (q \to p)$. Then $\varphi(\psi/p, q)$, or simply $\varphi(\psi, q)$, is the word
$$((q \to p) \to (q \vee (q \to p))).$$
For a propositional variable r, the word $\varphi(r/p, q)$, or $\varphi(r, q)$, is $(r \to (q \vee r))$. Finally, the word $\varphi(\psi/p, r/q)$, or $\varphi(\psi, r)$, is
$$((q \to p) \to (r \vee (q \to p))).$$

♠

By the above example one can feel that the word obtained after substituting formulas for propositional variables in a formula is itself a formula.

Theorem 1.1.15. *Given a propositional formula $\varphi(p_1, p_2, \ldots, p_n)$ and propositional formulas $\psi_1, \psi_2, \ldots, \psi_n$, the word $\varphi\big(\psi_1/p_1,\ \psi_2/p_2,\ \ldots,\ \psi_n/p_n\big)$ is a formula as well.*

Proof. Fix formulas $\psi_1, \psi_2, \ldots, \psi_n$ and propositional variables $p_1, p_2, \ldots, p_n$. We prove the theorem by induction on the formula $\varphi(p_1, p_2, \ldots, p_n)$.

- If $\varphi \in \mathbf{P}$, then either $\varphi = p_i$ for some $1 \leq i \leq n$ or $\varphi \notin \{p_1, p_2, \ldots, p_n\}$. In the former case $\varphi(\psi_1/p_1,\ \psi_2/p_2,\ \ldots,\ \psi_n/p_n) = \psi_i$, and in the latter case $\varphi(\psi_1/p_1,\ \psi_2/p_2,\ \ldots,\ \psi_n/p_n) = \varphi$. In both cases this is a formula.

- If $\varphi = \neg\theta$ and $\theta(\psi_1/p_1,\ \psi_2/p_2,\ \ldots,\ \psi_n/p_n)$ is a formula, then
$$\varphi(\psi_1/p_1,\ \psi_2/p_2,\ \ldots,\ \psi_n/p_n) = \neg\theta(\psi_1/p_1,\ \psi_2/p_2,\ \ldots,\ \psi_n/p_n)$$
is again a formula.

- If $\varphi = (\theta \diamond \chi)$ and both $\theta(\psi_1/p_1,\ \ldots,\ \psi_n/p_n)$ and $\chi(\psi_1/p_1,\ \ldots,\ \psi_n/p_n)$ are formulas, then
$$\varphi(\psi_1/p_1,\ \ldots,\ \psi_n/p_n) = (\theta(\psi_1/p_1,\ \ldots,\ \psi_n/p_n) \diamond \chi(\psi_1/p_1,\ \ldots,\ \psi_n/p_n))$$
is again a formula.

■

1.2 Semantics of Propositional Logic

1.2.1 Truth Assignments

Definition. A *truth assignment* is a function from the set of propositional variables $\mathbf{P}$ to the set $\{0,1\}$.

A truth assignment is also called an 'assignment of truth values', 'distribution of truth values', 'valuation', or 'evaluation'. Intuitively, one may think of 0 as false, and of 1 as true. We call 0 and 1 truth values. It is possible, as we shall see, to extend a truth assignment to the set $\mathcal{F}$ of all propositional formulas in one and only one way while agreeing with our mathematical intuition of the names we have given to the symbols for propositional connectives.

In particular, for any formulas φ and ψ we intend to satisfy the following. For the negation connective: we want $\neg\varphi$ to be true exactly when φ is false. For the conjunction connective: $(\varphi \wedge \psi)$ is true exactly when both φ and ψ are true. For the disjunction connective: $(\varphi \vee \psi)$ is false exactly when both φ and ψ are false. For the implication connective: $(\varphi \to \psi)$ is false exactly when φ is true and ψ is false. For the equivalence connective: $(\varphi \leftrightarrow \psi)$ is true exactly when both φ and ψ have the same truth value.

Given functions f and g, we say that g is an *extension* of f if $\mathrm{dom}(f) \subseteq \mathrm{dom}(g)$ and $f(a) = g(a)$ for all $a \in \mathrm{dom}(f)$, that is, g agrees with f on its domain.

Theorem 1.2.1. *Any truth assignment $\delta : \mathbf{P} \to \{0,1\}$ admits a unique extension $\hat{\delta} : \mathcal{F} \to \{0,1\}$ satisfying the following conditions for all formulas φ and ψ:*

(i) $\hat{\delta}[\neg\varphi] = 1$ *if and only if* $\hat{\delta}[\varphi] = 0$*;*

(ii) $\hat{\delta}[(\varphi \wedge \psi)] = 1$ *if and only if* $\hat{\delta}[\varphi] = 1$ *and* $\hat{\delta}[\psi] = 1$*;*

(iii) $\hat{\delta}[(\varphi \vee \psi)] = 0$ *if and only if* $\hat{\delta}[\varphi] = 0$ *and* $\hat{\delta}[\psi] = 0$*;*

(iv) $\hat{\delta}[(\varphi \to \psi)] = 0$ *if and only if* $\hat{\delta}[\varphi] = 1$ *and* $\hat{\delta}[\psi] = 0$*;*

(v) $\hat{\delta}[(\varphi \leftrightarrow \psi)] = 1$ *if and only if* $\hat{\delta}[\varphi] = \hat{\delta}[\psi]$*.*

Proof. Given a truth assignment $\delta : \mathbf{P} \to \{0,1\}$, we will define an extension $\hat{\delta}$ inductively on formulas as described below. The set $\{0,1\}$ is equipped with addition and multiplication modulo 2. We will use these arithmetic operations to express the conditions stated in the theorem.

We start by defining the value of $\hat{\delta}$ on propositional variables. For every propositional variable $p \in \mathbf{P}$ we define $\hat{\delta}[p] = \delta[p]$. So $\hat{\delta}$ agrees with δ on $\mathbf{P}$.

Next suppose that φ is formula where the value $\hat{\delta}[\varphi]$ was defined. We define

$$\hat{\delta}[\neg\varphi] = 1 + \hat{\delta}[\varphi].$$

This ensures that condition (i) is satisfied by $\hat{\delta}$.

Now suppose that φ and ψ are formulas where the values $\hat{\delta}[\varphi]$ and $\hat{\delta}[\psi]$ were defined. We define the following.

- $\hat{\delta}[(\varphi \wedge \psi)] = \hat{\delta}[\varphi] \cdot \hat{\delta}[\psi]$;
- $\hat{\delta}[(\varphi \vee \psi)] = \hat{\delta}[\varphi] + \hat{\delta}[\psi] + \hat{\delta}[\varphi] \cdot \hat{\delta}[\psi]$;
- $\hat{\delta}[(\varphi \to \psi)] = 1 + \hat{\delta}[\varphi] + \hat{\delta}[\varphi] \cdot \hat{\delta}[\psi]$;
- $\hat{\delta}[(\varphi \leftrightarrow \psi)] = 1 + \hat{\delta}[\varphi] + \hat{\delta}[\psi]$.

This ensures that conditions (ii), (iii), (iv), and (v) are satisfied by $\hat{\delta}$. Thus, $\hat{\delta}[\varphi]$ is now defined for every formula $\varphi \in \mathcal{F}$. This proves that $\hat{\delta} : \mathcal{F} \to \{0, 1\}$ is an extension of $\delta : \mathbf{P} \to \{0, 1\}$ satisfying the five conditions.

Next we prove that $\hat{\delta}$ is unique. Towards this end, suppose that $\lambda : \mathcal{F} \to \{0, 1\}$ is a function which extends $\delta : \mathbf{P} \to \{0, 1\}$ and satisfies the five conditions. We will show that $\hat{\delta} = \lambda$, that is, we need to show that $\hat{\delta}[\varphi] = \lambda[\varphi]$ for every formula $\varphi \in \mathcal{F}$. We will proceed by induction on formulas.

Suppose that $p \in \mathbf{P}$ is a propositional variable. Then $\hat{\delta}[p] = \delta[p] = \lambda[p]$ since both $\hat{\delta}$ and λ agree with δ on $\mathbf{P}$.

Next suppose that φ is formula where $\hat{\delta}[\varphi] = \lambda[\varphi]$. Since both $\hat{\delta}$ and λ satisfy condition (i) we get that

$$\hat{\delta}[\neg\varphi] = 1 + \hat{\delta}[\varphi] = 1 + \lambda[\varphi] = \lambda[\neg\varphi].$$

Next suppose that φ and ψ are formula where $\hat{\delta}[\varphi] = \lambda[\varphi]$ and $\hat{\delta}[\psi] = \lambda[\psi]$. Since both $\hat{\delta}$ and λ satisfy condition (ii) we get that

$$\hat{\delta}[(\varphi \wedge \psi)] = \hat{\delta}[\varphi] \cdot \hat{\delta}[\psi] = \lambda[\varphi] \cdot \lambda[\psi] = \lambda[(\varphi \wedge \psi)].$$

Since both $\hat{\delta}$ and λ satisfy condition (iii) we get that

$$\hat{\delta}[(\varphi \vee \psi)] = \hat{\delta}[\varphi] + \hat{\delta}[\psi] + \hat{\delta}[\varphi] \cdot \hat{\delta}[\psi] = \lambda[\varphi] + \lambda[\psi] + \lambda[\varphi] \cdot \lambda[\psi] = \lambda[(\varphi \vee \psi)].$$

Since both $\hat{\delta}$ and λ satisfy condition (iv) we get that

$$\hat{\delta}[(\varphi \to \psi)] = 1 + \hat{\delta}[\varphi] + \hat{\delta}[\varphi] \cdot \hat{\delta}[\psi] = 1 + \lambda[\varphi] + \lambda[\varphi] \cdot \lambda[\psi] = \lambda[(\varphi \to \psi)].$$

Since both $\hat{\delta}$ and λ satisfy condition (v) we get that

$$\hat{\delta}[(\varphi \leftrightarrow \psi)] = 1 + \hat{\delta}[\varphi] + \hat{\delta}[\psi] = 1 + \lambda[\varphi] + \lambda[\psi] = \lambda[(\varphi \leftrightarrow \psi)].$$

Therefore, by induction on formulas, we have shown that $\hat{\delta}[\varphi] = \lambda[\varphi]$ for every formula $\varphi \in \mathcal{F}$. That is, $\hat{\delta} = \lambda$, and so $\hat{\delta}$ is unique. ■

Another way to express the above five conditions is by tables called the *truth tables.*

Negation Truth Table

φ	$\neg\varphi$
1	0
0	1

Conjunction Truth Table

φ	ψ	$(\varphi \wedge \psi)$
1	1	1
1	0	0
0	1	0
0	0	0

Disjunction Truth Table

φ	ψ	$(\varphi \vee \psi)$
1	1	1
1	0	1
0	1	1
0	0	0

Implication Truth Table

φ	ψ	$(\varphi \to \psi)$
1	1	1
1	0	0
0	1	1
0	0	1

Equivalence Truth Table

φ	ψ	$(\varphi \leftrightarrow \psi)$
1	1	1
1	0	0
0	1	0
0	0	1

From now on we will not make a distinction between a truth assignment and its unique extension to the set of all formulas.

Definition. Let $\delta : \mathcal{F} \to \{0, 1\}$ be a truth assignment and let φ be a formula.

- We say δ *satisfies* φ if $\delta[\varphi] = 1$. We also say φ is *satisfied* by δ for this.
- We say φ is *satisfiable* if there is some truth assignment which satisfies φ.

Given a truth assignment $\delta : \mathbf{P} \to \{0, 1\}$ and a formula $\varphi \in \mathcal{F}$, how to find the truth value $\delta[\varphi]$ using the unique extension of δ? Well, we first find all subformulas of φ. The truth assignment δ tells us the truth value of all propositional variables in φ, and these are the subformulas of height 0. Then we use the conditions in the definition of the unique extension of δ to compute the truth values of all subformulas of height 1. Then we use the conditions again to find the truth value of all subformulas of height 2, and so on, until we get the truth value of φ, which is denoted by $\delta[\varphi]$.

Example. Let $\delta : \{p, q, r\} \to \{0, 1\}$ be a truth assignment given by $\delta[p] = 0$, $\delta[q] = 0$, and $\delta[r] = 1$. Find the truth value of

$$\varphi = ((p \to q) \to (q \vee (p \leftrightarrow r)))$$

under the unique extension of δ. First observe that

$$\mathrm{sf}[\varphi] = \{p,\ q,\ r,\ (p \to q),\ (p \leftrightarrow r),\ (q \vee (p \leftrightarrow r)),\ ((p \to q) \to (q \vee (p \leftrightarrow r)))\}.$$

- Truth values of subformulas of height 0.

$$\delta[p] = 0, \quad \delta[q] = 0, \quad \delta[r] = 1.$$

- Truth values of subformulas of height 1.

$$\delta[(p \to q)] = 1, \quad \delta[(p \leftrightarrow r)] = 0.$$

- Truth values of subformulas of height 2.

$$\delta[(q \vee (p \leftrightarrow r))] = 0.$$

- Truth values of subformulas of height 3.

$$\delta[\varphi] = \delta[((p \to q) \to (q \vee (p \leftrightarrow r)))] = 0.$$

Thus, δ does not satisfy φ. Is φ satisfiable? ♠

Example. Let $\delta : \{p, q, r\} \to \{0, 1\}$ be a truth assignment given by $\delta[p] = 0$, $\delta[q] = 0$, and $\delta[r] = 1$. Find the truth value of

$$\theta = (p \to (((q \wedge \neg p) \vee (\neg r \wedge p)) \leftrightarrow (p \vee (p \to \neg q)))).$$

Observe that θ is of the form $(p \to \psi)$ where $\psi = (((q \wedge \neg p) \vee (\neg r \wedge p)) \leftrightarrow (p \vee (p \to \neg q)))$. Since $\delta[p] = 0$, it follows that $\delta[\theta] = 1$, without the necessity of computing the truth values of all subformulas of θ. Thus, θ is satisfiable. ♠

Lemma 1.2.2. *For any truth assignments δ and λ and any formula $\varphi(p_1, p_2, \ldots, p_n)$, if δ and λ agree on the set $\{p_1, p_2, \ldots, p_n\}$, then $\delta[\varphi] = \lambda[\varphi]$.*

Proof. By induction on the formulas. ■

The lemma above tells us that when we want to investigate the truth value of a particular formula $\varphi(p_1, p_2, \ldots, p_n)$ under all truth assignments, it is sufficient to assume that the domain of the truth assignment in hand is the set $\{p_1, p_2, \ldots, p_n\}$ instead of the set $\mathbf{P}$ of all propositional variables. There are then finitely many truth assignments to consider, namely, all functions from $\{p_1, p_2, \ldots, p_n\}$ to $\{0, 1\}$. There are 2^n many such functions. Each truth assignment $\delta : \{p_1, p_2, \ldots, p_n\} \to \{0, 1\}$ can be expressed as an n-tuple of the form:

$$\big(\delta[p_1],\ \delta[p_2],\ \ldots,\ \delta[p_n]\big).$$

We then can make a table, called a *truth table*, that presents the truth value of φ under each of the 2^n different truth assignments. For each one of the 2^n truth assignments there will be a row corresponding to it that also contains the corresponding truth value of φ and possibly other subformulas of φ.

Example. Find the truth table of the formula

$$\psi = \psi(p, q, r) = ((q \vee (\neg r \wedge p)) \leftrightarrow (p \to \neg q)).$$

p	q	r	$\neg q$	$\neg r$	$(\neg r \wedge p)$	$(p \to \neg q)$	$(q \vee (\neg r \wedge p))$	ψ
1	1	1	0	0	0	0	1	0
1	1	0	0	1	1	0	1	0
1	0	1	1	0	0	1	0	0
1	0	0	1	1	1	1	1	1
0	1	1	0	0	0	1	1	1
0	1	0	0	1	0	1	1	1
0	0	1	1	0	0	1	0	0
0	0	0	1	1	0	1	0	0

The second row in the table above says that the truth assignment $\delta : \{p, q, r\} \to \{0, 1\}$ given by $\delta[p] = 1$, $\delta[q] = 1$, and $\delta[r] = 0$ would assign the value 0 to the formula ψ. So ψ is not satisfied by δ. Note that we can express δ as the tuple $(1, 1, 0)$. However, ψ is satisfiable because it is satisfied by the truth assignment λ represented by the tuple $(1, 0, 0)$. ♠

In general, there are different truth tables for the same formula, for example, we may consider truth tables which involve propositional variables that do not occur in the formula or we may change the order of the rows in the table (there are $(2^n)!$ many ways to put the 2^n rows one after the other). We will restrict ourselves to truth tables which only contain propositional variables that occur at least once in the formula. Additionally, we will include a column for every subformula of the formula. Furthermore, we will arrange the rows in decreasing order according to the lexicographical order (the 'dictionary' order): for binary strings we declare $(a_1, a_2, \ldots, a_n) < (b_1, b_2, \ldots, b_n)$ if $a_k < b_k$ where k is the least such that $a_k \neq b_k$. Consequently, we can speak of "the truth table" of a formula.

Observe that the truth table of a formula $\varphi(p_1, p_2, \ldots, p_n)$ is a function

$$f_\varphi : \{0, 1\}^n \to \{0, 1\}.$$

The domain $\{0, 1\}^n$ of f_φ is the set of all truth assignments $\delta : \{p_1, p_2, \ldots, p_n\} \to \{0, 1\}$ since we can think of δ as an n-tuple of 0s and 1s. For every δ in $\{0, 1\}^n$ we have that $f_\varphi[\delta] = \delta[\varphi]$. So every formula $\varphi(p_1, p_2, \ldots, p_n)$ gives rise to a function $f_\varphi : \{0, 1\}^n \to \{0, 1\}$. Conversely, given any function $f : \{0, 1\}^n \to \{0, 1\}$, one can show that there exists a formula φ (in disjuntive normal form: disjunctions of conjunctions of propositional variables and their negations) such that $f = f_\varphi$. In other words, any function $f : \{0, 1\}^n \to \{0, 1\}$ is the truth table of some formula $\varphi(p_1, p_2, \ldots, p_n)$.

Boolean Satisfiability Problem (SAT)
Given a formula $\varphi(p_1, p_2, \dots, p_n)$, the problem is to decide whether φ is satisfiable or not. Well, we can check if any of the 2^n different truth assignments satisfies φ. The question is: how long it takes to find that truth assignment in terms of the length of the formula?
If a truth assignment δ is known, then it is faster to decide whether δ satisfies φ or not; so SAT is an NP (nondeterministic polynomial time) problem. However, to search for a suitable truth assignment is seemingly an exponential search. Is there a polynomial-time (P) algorithm to find a truth assignment which satisfies φ? That's the big question!
SAT is a very important problem in computer science, it is NP-complete, and so resolving it would solve the famous "P versus NP" problem in the field of complexity theory. The P vs. NP problem is one of the Clay Millennium Problems and it carries a prize worth of one million US dollars.

1.2.2 Tautologies and Logical Equivalence

Definition.

- A propositional formula φ is called a *tautology* if for every truth assignment $\delta : \mathbf{P} \to \{0, 1\}$ we have $\delta[\varphi] = 1$.
- The notation for φ being a tautology is $\models \varphi$.
- A propositional formula φ is called a *contradiction* if for every truth assignment $\delta : \mathbf{P} \to \{0, 1\}$ we have $\delta[\varphi] = 0$.
- Given two propositional formulas φ and ψ, we say that φ is *logically equivalent* to ψ if for every truth assignment $\delta : \mathbf{P} \to \{0, 1\}$ we have $\delta[\varphi] = \delta[\psi]$.
- We write $\varphi \equiv \psi$ when φ is logically equivalent to ψ.

A tautology is a formula which is always true. A tautology is therefore a formula whose truth table contains only 1s in the column corresponding to the formula. Similarly, a contradiction is always false; the column corresponding to a contradiction contains only 0s. Two logically equivalent formulas have identical truth tables. Any formula logically equivalent to a tautology is itself a tautology. Similarly, any formula logically equivalent to a contradiction is itself a contradiction.

Lemma 1.2.3. *Let φ and ψ be propositional formulas. Then the following hold.*

- *If $\varphi \equiv \psi$, then $\neg\varphi \equiv \neg\psi$.*
- *$\varphi \equiv \psi$ if and only if $\models (\varphi \leftrightarrow \psi)$.*

Lemma 1.2.4. *The logical equivalence relation is an equivalence relation on the set $\mathcal{F}$ of all propositional formulas. In other words,*

- *(Reflexivity). For any $\varphi \in \mathcal{F}$ we have $\varphi \equiv \varphi$.*
- *(Symmetry). For any $\varphi, \psi \in \mathcal{F}$ we have if $\varphi \equiv \psi$, then $\psi \equiv \varphi$.*
- *(Transitivity). For any $\varphi, \psi, \theta \in \mathcal{F}$ we have if $\varphi \equiv \psi$ and $\psi \equiv \theta$, then $\varphi \equiv \theta$.*

Recall that the equivalence class of a formula φ under the logical equivalence relation is $[\varphi] = \{\psi \in \mathcal{F} \mid \psi \equiv \varphi\}$ and the set $\mathcal{F}/_{\equiv}$ is the set of all equivalence classes. That is, $\mathcal{F}/_{\equiv} = \big\{[\varphi] \mid \varphi \in \mathcal{F}\big\}$. Suppose that $\mathbf{P} = \{p_1, p_2, \dots, p_n\}$. We then can define a function T from the set $\mathcal{F}/_{\equiv}$ to the set $\{f \text{ is a function} \mid f : \{0,1\}^n \to \{0,1\}\}$ by setting $T([\varphi]) = f_\varphi$; check above to see the definition of f_φ. The reader is encouraged to show that T is indeed well-defined, and moreover it is injective, and surjective. (See the discussion above on f_φ. We will discuss this again in further detail in the next section.) Therefore, the two sets have equal cardinality. It follows that there are 2^{2^n} distinct equivalence classes. In other words, there are 2^{2^n} formulas in propositional variables $p_1, p_2, \dots, p_n$ up to logical equivalence.

We now examine the effect of substitutions on the truth values of formulas. In particular, we will show that substitutions preserve tautologies.

Theorem 1.2.5. *Let $p_1, p_2, \dots, p_n$ be distinct propositional variables, and consider propositional formulas $\varphi, \psi_1, \psi_2, \dots, \psi_n$. Let δ be an arbitrary truth assignment. Define a truth assignment λ as follows: for all $x \in \mathbf{P}$ we set*

$$\lambda[x] = \begin{cases} \delta[x] & \text{if } x \notin \{p_1, p_2, \dots, p_n\}; \\ \delta[\psi_i] & \text{if } x = p_i \text{ for some } 1 \le i \le n. \end{cases}$$

Then we have $\delta[\varphi(\psi_1/p_1,\ \psi_2/p_2,\ \dots, \psi_n/p_n)] = \lambda[\varphi]$.

Proof. Fix some distinct propositional variables $p_1, p_2, \dots, p_n$ and some formulas $\psi_1, \psi_2, \dots, \psi_n$. We now argue by induction on formulas.

Suppose that $\varphi = p$ for some propositional variable p. If $p \notin \{p_1, p_2, \dots, p_n\}$, then in this case $\varphi(\psi_1/p_1, \dots \psi_n/p_n) = p$. Then by definition of λ we get

$$\delta[\varphi(\psi_1/p_1,\ \psi_2/p_2,\ \dots, \psi_n/p_n)] = \delta[p] = \lambda[p] = \lambda[\varphi].$$

Otherwise, $\varphi = p_i$ for some $1 \le i \le n$. In this case, $\varphi(\psi_1/p_1, \dots \psi_n/p_n) = \psi_i$. Then by definition of λ we get

$$\delta[\varphi(\psi_1/p_1,\ \psi_2/p_2,\ \dots, \psi_n/p_n)] = \delta[\psi_i] = \lambda[p_i] = \lambda[\varphi].$$

Suppose that φ and θ are formulas such that $\delta[\varphi(\psi_1/p_1, \ldots, \psi_n/p_n)] = \lambda[\varphi]$ and $\delta[\theta(\psi_1/p_1, \ldots, \psi_n/p_n)] = \lambda[\theta]$. We now show that the equality holds for $\eta = \neg\varphi$.

$$\begin{aligned}\delta[\eta(\psi_1/p_1, \ldots, \psi_n/p_n)] &= \delta[\neg\varphi(\psi_1/p_1, \ldots, \psi_n/p_n)]\\ &= 1 + \delta[\varphi(\psi_1/p_1, \ldots, \psi_n/p_n)]\\ &= 1 + \lambda[\varphi]\\ &= \lambda[\neg\varphi] = \lambda[\eta].\end{aligned}$$

Next we show that the equality holds for $\chi = (\varphi \wedge \theta)$.

$$\begin{aligned}\delta[\chi(\psi_1/p_1, \psi_2/p_2, \ldots, \psi_n/p_n)] &= \delta[(\varphi(\psi_1/p_1, \ldots, \psi_n/p_n) \wedge \theta(\psi_1/p_1, \ldots, \psi_n/p_n))]\\ &= \delta[(\varphi(\psi_1/p_1, \ldots, \psi_n/p_n)] \cdot \delta[\theta(\psi_1/p_1, \ldots, \psi_n/p_n)]\\ &= \lambda[\varphi] \cdot \lambda[\theta]\\ &= \lambda[(\varphi \wedge \theta)]\\ &= \lambda[\chi].\end{aligned}$$

Showing that the equality holds for $(\varphi \vee \theta)$, $(\varphi \to \theta)$, and $(\varphi \leftrightarrow \theta)$ is shown in a similar fashion. ■

Corollary 1.2.6. *For all distinct propositional variables $p_1, p_2, \ldots, p_n$ and propositional formulas $\varphi, \psi_1, \psi_2, \ldots, \psi_n$, if φ is a tautology, then the formula $\varphi(\psi_1/p_1, \ldots, \psi_n/p_n)$ is a tautology as well.*

Proof. Suppose that φ is a tautology. Let δ be any truth assignment, and let λ be the corresponding truth assignment as defined in the previous theorem. Then

$$\delta[\varphi(\psi_1/p_1, \psi_2/p_2, \ldots, \psi_n/p_n)] = \lambda[\varphi] = 1,$$

since φ is assigned the value 1 by every truth assignment. ■

Corollary 1.2.7. *For any formulas $\varphi(p_1, \ldots, p_n)$, $\theta(p_1, \ldots, p_n)$, $\psi_1, \psi_2, \ldots, \psi_n$, we have if $\varphi \equiv \theta$, then*

$$\varphi(\psi_1/p_1, \psi_2/p_2, \ldots, \psi_n/p_n) \equiv \theta(\psi_1/p_1, \psi_2/p_2, \ldots, \psi_n/p_n).$$

Theorem 1.2.8. *Let φ, ψ, and θ be propositional formulas. Suppose that ψ is a subformula of φ and $\theta \equiv \psi$. Then the formula $\hat{\varphi}$ obtained from φ by substituting θ for the subformula ψ is logically equivalent to φ.*

Proof. We argue by induction on the formula φ.

Suppose that φ is a propositional variable, say $\varphi = p$ for some $p \in \mathbf{P}$, and let ψ be a subformula of φ. Then it must be that $\psi = p$. Now let θ be a formula such that $\theta \equiv \psi$ and observe that $\hat{\varphi} = \varphi(\theta/p) = \theta$. Since $\varphi = \psi$, $\psi \equiv \theta$, and $\theta = \hat{\varphi}$, we conclude that $\varphi \equiv \hat{\varphi}$.

Let $\varphi = \neg\eta$ where η is a formula satisfying the statement of the theorem. Let ψ be a subformula of φ and let $\theta \equiv \psi$. Either $\psi = \varphi$, and in this case $\hat{\varphi} = \theta$ and so $\varphi \equiv \hat{\varphi}$, or else ψ is a subformula of η, and by induction hypothesis, the formula $\hat{\eta}$ which results from substituting θ for ψ in η is logically equivalent to η. Observe that $\hat{\varphi} = \neg\hat{\eta}$. Since $\eta \equiv \hat{\eta}$ we must have that $\neg\eta \equiv \neg\hat{\eta}$ by Lemma 1.3.1. Therefore, $\varphi \equiv \hat{\varphi}$.

Next suppose that $\varphi = (\eta \wedge \beta)$ where η and β are formulas which satisfy the statement of the theorem. Let ψ is a subformula of φ and $\theta \equiv \psi$. There are three possibilities: $\psi = \varphi$, $\psi \in \mathrm{sf}[\eta]$, or $\psi \in \mathrm{sf}[\beta]$. If $\psi = \varphi$, then $\hat{\varphi} = \theta$ and so $\varphi \equiv \hat{\varphi}$. If ψ is a subformula of η, then $\hat{\varphi} = (\hat{\eta} \wedge \beta)$ and by induction hypothesis we know that $\eta \equiv \hat{\eta}$. Now let δ be a truth assignment, then

$$\delta[\varphi] = \delta[(\eta \wedge \beta)] = \delta[\eta] \cdot \delta[\beta] = \delta[\hat{\eta}] \cdot \delta[\beta] = \delta[(\hat{\eta} \wedge \beta)] = \delta[\hat{\varphi}].$$

Therefore, $\varphi \equiv \hat{\varphi}$ as desired. The last possibility where ψ is a subformula of β is analogous.

The remaining cases where $\varphi = (\eta \vee \beta)$, $\varphi = (\eta \rightarrow \beta)$, and $\varphi = (\eta \leftrightarrow \beta)$ are dealt with in a similar fashion. ■

Example. Let χ, ψ, θ, and φ be formulas as given below.

- Show that $\chi = \chi(p) = (p \vee \neg p)$ is a tautology.
  Let $\delta : \mathbf{P} \rightarrow \{0,1\}$ be any truth assignment. If $\delta[p] = 0$, then

  $$\delta[\chi] = \delta[(p \vee \neg p)] = \delta[p] + \delta[\neg p] + \delta[p] \cdot \delta[\neg p] = 0 + 1 + 0 \cdot 1 = 1.$$

  Otherwise, $\delta[p] = 1$, and so

  $$\delta[\chi] = \delta[(p \vee \neg p)] = \delta[p] + \delta[\neg p] + \delta[p] \cdot \delta[\neg p] = 1 + 0 + 1 \cdot 0 = 1.$$

  Therefore, χ takes the value 1 by any truth assignment, and so χ is a tautology.

- Let $\psi = (\neg p \vee q)$ and $\theta = (p \rightarrow q)$. Show that $\psi \equiv \theta$.
  Let $\delta : \mathbf{P} \rightarrow \{0,1\}$ be any truth assignment. Then

  $$\begin{aligned} \delta[\psi] &= \delta[(\neg p \vee q)] \\ &= \delta[\neg p] + \delta[q] + \delta[\neg p] \cdot \delta[q] \\ &= 1 + \delta[p] + \delta[q] + (1 + \delta[p]) \cdot \delta[q] \\ &= 1 + \delta[p] + \delta[q] + \delta[q] + \delta[p] \cdot \delta[q] \\ &= 1 + \delta[p] + \delta[p] \cdot \delta[q] \\ &= \delta[(p \rightarrow q)] \\ &= \delta[\theta]. \end{aligned}$$

- Show that $\varphi = ((\neg p \vee q) \vee \neg(p \to q))$ is a tautology.
  Observe that ψ is a subformula of φ and as we have shown $\psi \equiv \theta$. Thus, by Theorem 1.2.8, the formula $\hat{\varphi}$ obtained from φ by substituting θ for the subformula ψ is logically equivalent to φ. Here,

  $$\hat{\varphi} = ((p \to q) \vee \neg(p \to q)).$$

  Next observe that $\hat{\varphi} = \chi(\theta/p)$. Since χ is a tautology, by Corollary 1.2.6, we get that $\chi(\theta/p)$ is a tautology as well, and so $\hat{\varphi}$ is a tautology. Finally, since $\hat{\varphi} \equiv \varphi$, we conclude that φ is a tautology as desired. ♠

We use the symbol $\top$ to denote an arbitrary tautology and the symbol $\bot$ to denote an arbitrary contradiction.

Lemma 1.2.9. *Let p, q, and r be propositional variables. Then the following hold.*

1. $(p \wedge p) \equiv p$ *(Idempotence of $\wedge$)*
2. $(p \vee p) \equiv p$ *(Idempotence of $\vee$)*
3. $(p \wedge q) \equiv (q \wedge p)$ *(Commutativity of $\wedge$)*
4. $(p \vee q) \equiv (q \vee p)$ *(Commutativity of $\vee$)*
5. $(p \wedge (q \wedge r)) \equiv ((p \wedge q) \wedge r)$ *(Associativity of $\wedge$)*
6. $(p \vee (q \vee r)) \equiv ((p \vee q) \vee r)$ *(Associativity of $\vee$)*
7. $(p \wedge (q \vee r)) \equiv ((p \wedge q) \vee (p \wedge r))$ *(Distributivity of $\wedge$ over $\vee$)*
8. $(p \vee (q \wedge r)) \equiv ((p \vee q) \wedge (p \vee r))$ *(Distributivity of $\vee$ over $\wedge$)*
9. $(p \wedge (p \vee q)) \equiv p$ *(Absorption law)*
10. $(p \vee (p \wedge q)) \equiv p$ *(Absorption law)*
11. $\neg(p \wedge q) \equiv (\neg p \vee \neg q)$ *(De Morgan's law)*
12. $\neg(p \vee q) \equiv (\neg p \wedge \neg q)$ *(De Morgan's law)*
13. $(p \wedge \top) \equiv p$ *(Identity element for $\wedge$)*
14. $(p \vee \bot) \equiv p$ *(Identity element for $\vee$)*
15. $(p \wedge \bot) \equiv \bot$ *(Zero element for $\wedge$)*
16. $(p \vee \top) \equiv \top$ *(Zero element for $\vee$)*
17. $(p \to q) \equiv (\neg q \to \neg p)$ *(Contrapositive)*
18. $(p \to q) \equiv (\neg p \vee q)$
19. $(p \to q) \equiv ((p \wedge q) \leftrightarrow p)$
20. $(p \to q) \equiv ((p \vee q) \leftrightarrow q)$
21. $(p \leftrightarrow q) \equiv ((p \to q) \wedge (q \to p))$

22. $(p \leftrightarrow q) \equiv (\neg p \leftrightarrow \neg q)$
23. $p \equiv (\top \to p)$
24. $\neg p \equiv (p \to \bot)$

Remark. By Corollary 1.2.7, we may substitute arbitrary formulas for the propositional variables p, q, r and still maintain the equivalences above.

Lemma 1.2.10. *Let p, q, r be propositional variables. The following are tautologies.*

1. $(p \vee \neg p)$
2. $(p \to p)$
3. $(p \leftrightarrow p)$
4. $(\neg\neg p \to p)$
5. $(p \to (p \vee q))$
6. $((p \wedge q) \to p)$
7. $(((p \to q) \wedge p) \to q)$
8. $(((p \to q) \wedge \neg q) \to \neg p)$
9. $((\neg p \to p) \to p)$
10. $((\neg p \to p) \leftrightarrow p)$
11. $(\neg p \to (p \to q))$
12. $(p \vee (p \to q))$
13. $(p \to (q \to p))$
14. $(((p \to q) \wedge (q \to r)) \to (p \to r))$
15. $((p \to q) \vee (r \to p))$
16. $((p \to q) \to ((q \to r) \to (p \to r)))$
17. $(\neg p \to (\neg q \leftrightarrow (q \to p)))$
18. $((p \to q) \to (((p \to r) \to q) \to q))$

Remark. By Corollary 1.2.6, we may substitute arbitrary formulas for the propositional variables p, q, r and still maintain the tautologies above.

1.2.3 Disjunctive Normal Form

Abuse of notation. We will drop the inner parentheses in formulas constructed by taking successive conjunctions of formulas and keep just the two outermost parentheses. We will write $(\varphi_1 \wedge \varphi_2 \wedge \varphi_3)$ for the formula $((\varphi_1 \wedge \varphi_2) \wedge \varphi_3)$. Alternatively, we could choose $(\varphi_1 \wedge \varphi_2 \wedge \varphi_3)$ to represent the formula $(\varphi_1 \wedge (\varphi_2 \wedge \varphi_3))$ instead, but both are logically equivalent as $\wedge$ is associative and we are interested here in the truth value of formulas rather than their syntactic form. Similarly we will write $(\varphi_1 \wedge \varphi_2 \wedge \varphi_3 \wedge \varphi_4)$ for the formula $(((\varphi_1 \wedge \varphi_2) \wedge \varphi_3) \wedge \varphi_4)$. In general, taking successive conjunctions of n formulas will be written as

$$(\varphi_1 \wedge \varphi_2 \wedge \varphi_3 \wedge \cdots \wedge \varphi_n) \qquad \text{or} \qquad \bigwedge_{i=1}^{n} \varphi_i \qquad \text{or} \qquad \bigwedge_{i \in I} \varphi_i \,.$$

Similarly, as $\vee$ is associative, we shall write a formula constructed by taking successive disjunctions of formulas as

$$(\varphi_1 \vee \varphi_2 \vee \varphi_3 \vee \cdots \vee \varphi_n) \qquad \text{or} \qquad \bigvee_{i=1}^{n} \varphi_i \qquad \text{or} \qquad \bigvee_{i \in I} \varphi_i \,.$$

Lemma 1.2.11. *The following hold for any integer $n \geq 1$.*

- $\neg \bigwedge_{i=1}^{n} \varphi_i \equiv \bigvee_{i=1}^{n} \neg\varphi_i$.
- $\neg \bigvee_{i=1}^{n} \varphi_i \equiv \bigwedge_{i=1}^{n} \neg\varphi_i$.

Proof. We prove the first equivalence by induction on n. For the base case,

$$\neg \bigwedge_{i=1}^{1} \varphi_i = \neg\varphi_1 = \bigvee_{i=1}^{1} \neg\varphi_i.$$

Now suppose the equivalence holds for some $n \geq 1$. It follows that

$$\begin{aligned}
\neg \bigwedge_{i=1}^{n+1} \varphi_i &= \neg \left(\left(\bigwedge_{i=1}^{n} \varphi_i \right) \wedge \varphi_{n+1} \right) \\
&\equiv \left(\neg \left(\bigwedge_{i=1}^{n} \varphi_i \right) \vee \neg\varphi_{n+1} \right) \\
&\equiv \left(\left(\bigvee_{i=1}^{n} \neg\varphi_i \right) \vee \neg\varphi_{n+1} \right) \\
&\equiv \bigvee_{i=1}^{n+1} \neg\varphi_i.
\end{aligned}$$

To get the equivalence before the last one we used the induction hypothesis and Theorem 1.2.8. ■

Set $\neg\mathbf{P} = \{\neg p \mid p \in \mathbf{P}\}$. Any word in $\mathbf{P} \cup \neg\mathbf{P}$ is called a *literal.* So a literal is either a propositional variable or a negation of a propositional variable.

Definition.

- A propositional formula is in *disjunctive normal formal* (DNF) if it is a disjunction of formulas which are conjunctions of propositional variables and their negations. In other words, it is a formula of the form
$$\bigvee_{i=1}^{n}\left(\bigwedge_{j=1}^{k_i} x_{ij}\right) \text{ where } x_{ij} \in \mathbf{P} \cup \neg\mathbf{P}.$$

- A propositional formula is in *conjunctive normal formal* (CNF) if it is a conjunction of formulas which are disjunctions of propositional variables and their negations. In other words, it is a formula of the form
$$\bigwedge_{i=1}^{n}\left(\bigvee_{j=1}^{k_i} x_{ij}\right) \text{ where } x_{ij} \in \mathbf{P} \cup \neg\mathbf{P}.$$

Example.

- A propositional variable p is both in DNF and CNF.
- $(p \wedge q)$, $(p \vee \neg p)$, $(p \wedge \neg q \wedge \neg r)$, and $(\neg p \vee q \vee r)$ are in DNF and CNF.
- $((p \wedge \neg q \wedge r) \vee (\neg r \wedge s) \vee (q \wedge q \wedge \neg s) \vee r)$ is in DNF.
- $((s \vee \neg q \vee \neg r) \wedge \neg r \wedge (q \vee r \vee \neg s \vee p) \wedge (r \vee s))$ is in CNF.

In the remaining part of this section we assume that $\mathbf{P} = \{p_1, p_2, \ldots, p_n\}$ for some $n \geq 1$.

Notation.

- Let $p \in \mathbf{P}$ and $\epsilon \in \{0,1\}$ we define
$$p^\epsilon = \begin{cases} p & \text{if } \epsilon = 1; \\ \neg p & \text{if } \epsilon = 0. \end{cases}$$
So $p^1 = p$ and $p^0 = \neg p$.

- Consider an n-tuple $\bar{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n) \in \{0,1\}^n$. Then by $\delta_{\bar{\epsilon}}$ we denote the truth assignment $\delta_{\bar{\epsilon}} : \{p_1, p_2, \ldots, p_n\} \to \{0,1\}$ given by $\delta_{\bar{\epsilon}}[p_i] = \epsilon_i$.

  For example, let $\bar{\epsilon} = (1,0,1,0) \in \{0,1\}^4$. Then $\delta_{\bar{\epsilon}} : \{p_1, p_2, p_3, p_4\} \to \{0,1\}$ is given by $\delta_{\bar{\epsilon}}[p_1] = 1$, $\delta_{\bar{\epsilon}}[p_2] = 0$, $\delta_{\bar{\epsilon}}[p_3] = 1$, and $\delta_{\bar{\epsilon}}[p_4] = 0$.

- For each formula $\varphi(p_1, \ldots, p_n) \in \mathcal{F}$ we define a function $f_\varphi : \{0,1\}^n \to \{0,1\}$ by setting
$$f_\varphi[\bar{\epsilon}] = \delta_{\bar{\epsilon}}[\varphi]$$
for every $\bar{\epsilon} \in \{0,1\}^n$. Observe that f_φ is the function given by the truth table of the formula φ. It follows that two formulas φ and ψ are logically equivalent if and only if $f_\varphi = f_\psi$.

Our aim now is to figure out whether the mapping from the set of formulas $\mathcal{F}$ to the set $\{f \text{ is a function} \mid f : \{0,1\}^n \to \{0,1\}\}$ given by $\varphi \mapsto f_\varphi$ is surjective. In other words, can every function $f : \{0,1\}^n \to \{0,1\}$ be viewed as the truth table of some formula?

Lemma 1.2.12. *Fix some n-tuple $\bar{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n) \in \{0,1\}^n$, and consider the formula*

$$\chi = \chi_{\bar{\epsilon}} = \bigwedge_{i=1}^{n} p_i^{\epsilon_i} = (p_1^{\epsilon_1} \wedge p_2^{\epsilon_2} \wedge \cdots \wedge p_n^{\epsilon_n}) .$$

Then $\delta_{\bar{\epsilon}}[\chi] = 1$ and no other truth assignment satisfies χ.

Proof. Fix some n-tuple $\bar{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n) \in \{0,1\}^n$. Now if $\epsilon_i = 1$, then

$$\delta_{\bar{\epsilon}}[p_i^{\epsilon_i}] = \delta_{\bar{\epsilon}}[p_i^1] = \delta_{\bar{\epsilon}}[p_i] = \epsilon_i = 1.$$

If $\epsilon_i = 0$, then

$$\delta_{\bar{\epsilon}}[p_i^{\epsilon_i}] = \delta_{\bar{\epsilon}}[p_i^0] = \delta_{\bar{\epsilon}}[\neg p_i] = 1 + \delta_{\bar{\epsilon}}[p_i] = 1 + \epsilon_i = 1 + 0 = 1.$$

Thus, $\delta_{\bar{\epsilon}}[p_i^{\epsilon_i}] = 1$ for all $1 \leq i \leq n$.

$$\delta_{\bar{\epsilon}}[\chi] = \delta_{\bar{\epsilon}}\left[\bigwedge_{i=1}^{n} p_i^{\epsilon_i}\right] = \delta_{\bar{\epsilon}}[p_1^{\epsilon_1}] \cdot \delta_{\bar{\epsilon}}[p_2^{\epsilon_2}] \cdot \ldots \cdot \delta_{\bar{\epsilon}}[p_n^{\epsilon_n}] = 1 \cdot 1 \cdot \ldots \cdot 1 = 1.$$

Next, suppose that λ is a truth assignment such that $\lambda \neq \delta_{\bar{\epsilon}}$. Thus there is some j such that $\lambda[p_j] \neq \delta_{\bar{\epsilon}}[p_j]$ and as $\delta_{\bar{\epsilon}}[p_j] = \epsilon_j$, it must be $\lambda[p_j] \neq \epsilon_j$. If $\epsilon_j = 1$, we get $\lambda[p_j] = 0$, and thus $\lambda[p_j^{\epsilon_j}] = \lambda[p_j^1] = \lambda[p_j] = 0$. If $\epsilon_j = 0$, it must be $\lambda[p_j] = 1$, and thus $\lambda[p_j^{\epsilon_j}] = \lambda[p_j^0] = \lambda[\neg p_j] = 1 + \lambda[p_j] = 1 + 1 = 0$. Thus $\lambda[p_j^{\epsilon_j}] = 0$ in both cases. It follows that $\lambda[\chi] = 0$, and so χ is not satisfied by λ. ■

Theorem 1.2.13. *Let $f : \{0,1\}^n \to \{0,1\}$ be a function. Then there exists a formula φ in disjuntive normal form such that $f = f_\varphi$.*

Proof. Fix some function $f : \{0,1\}^n \to \{0,1\}$. If f assigns 0 to all tuples in its domain, then take $\varphi = (p_1 \wedge \neg p_1)$. Clearly, $f = f_\varphi$ since φ is a contradiction. Otherwise, f assigns 1 at least once. Let

$$\Delta = \{\bar{e} \in \{0,1\}^n \mid f(\bar{e}) = 1\}.$$

Clearly, $\Delta \neq \emptyset$. Consider the formula

$$\varphi = \bigvee_{\bar{e} \in \Delta} \chi_{\bar{e}}.$$

Clearly, φ is in disjunctive normal form since each $\chi_{\bar{e}}$ is a conjunction of literals (see the form of $\chi_{\bar{e}}$ from Lemma 1.2.12). It remains to show that $f = f_\varphi$. By

Lemma 1.2.12, we know that $\chi_{\bar{\epsilon}}$ is only satisfied by the truth assignment $\delta_{\bar{\epsilon}}$ for any $\bar{\epsilon} \in \{0,1\}^n$. Let $\bar{\epsilon}$ be an n-tuple from the domain $\{0,1\}^n$. If $\bar{\epsilon} \in \Delta$, then

$$f_\varphi[\bar{\epsilon}] = \delta_{\bar{\epsilon}}[\varphi] = \delta_{\bar{\epsilon}}\left[\bigvee_{\bar{e}\in\Delta} \chi_{\bar{e}}\right] = \delta_{\bar{\epsilon}}\left[\left(\bigvee_{\bar{e}\in\Delta\setminus\{\bar{\epsilon}\}} \chi_{\bar{e}}\right) \vee \chi_{\bar{\epsilon}}\right] = 1$$

because $\delta_{\bar{\epsilon}}[\chi_{\bar{\epsilon}}] = 1$ and $\delta_{\bar{\epsilon}}$ satisfies the truth table of $\vee$. As $\bar{\epsilon} \in \Delta$, we get that $f[\bar{\epsilon}] = 1$ as well. So $f[\bar{\epsilon}] = f_\varphi[\bar{\epsilon}]$ when $\bar{\epsilon} \in \Delta$.

For the other case, suppose $\bar{\epsilon} \notin \Delta$. Then $f[\bar{\epsilon}] = 0$. Moreover, $\delta_{\bar{\epsilon}}[\chi_{\bar{e}}] = 0$ for every $\bar{e} \in \Delta$ since $\chi_{\bar{e}}$ is only satisfied by $\delta_{\bar{e}}$ and $\bar{\epsilon} \neq \bar{e}$ for every $\bar{e} \in \Delta$. It follows that

$$f_\varphi[\bar{\epsilon}] = \delta_{\bar{\epsilon}}[\varphi] = \delta_{\bar{\epsilon}}\left[\bigvee_{\bar{e}\in\Delta} \chi_{\bar{e}}\right] = 0.$$

We have shown that if $\bar{\epsilon} \in \Delta$, then $f[\bar{\epsilon}] = 1 = f_\varphi[\bar{\epsilon}]$, and if $\bar{\epsilon} \notin \Delta$, then $f[\bar{\epsilon}] = 0 = f_\varphi[\bar{\epsilon}]$. Therefore, $f = f_\varphi$ as desired. ■

Corollary 1.2.14. *Every formula is logically equivalent to one in disjunctive normal form and one in conjunctive normal form.*

Proof. Fix an arbitrary formula $\psi(p_1, \dots, p_n)$ and let $f_\psi : \{0,1\}^n \to \{0,1\}$ be the function given by its truth table. By Theorem 1.2.13, there exists a formula φ in disjunctive normal form such that $f_\psi = f_\varphi$. This implies that φ is logically equivalent to ψ as desired. So every formula is logically equivalent to a formula in disjunctive normal form.

It remains to find a formula in conjunctive normal form logically equivalent to ψ. By the first paragraph, we know that $\neg\psi$ is logically equivalent to some formula θ in disjunctive normal form. Say

$$\theta = \bigvee_{i=1}^{n}\left(\bigwedge_{j=1}^{k_i} x_{ij}\right) \text{ where } x_{ij} \in \mathbf{P} \cup \neg\mathbf{P}.$$

By Lemma 1.2.11 we now obtain that

$$\psi \equiv \neg\neg\psi \equiv \neg\bigvee_{i=1}^{n}\left(\bigwedge_{j=1}^{k_i} x_{ij}\right) \equiv \bigwedge_{i=1}^{n}\neg\left(\bigwedge_{j=1}^{k_i} x_{ij}\right) \equiv \bigwedge_{i=1}^{n}\left(\bigvee_{j=1}^{k_i} \neg x_{ij}\right).$$

Since the negation of a literal is logically equivalent to a literal, it follows that ψ is logically equivalent to a formula in conjuntive normal form. ■

Example. Follow the proofs of Theorem 1.2.13 and Corollary 1.2.14 to find a formula in DNF and a formula in CNF logically equivalent to the formula

$$\psi = \psi(p,q,r) = ((q \vee (\neg r \wedge p)) \leftrightarrow (p \to \neg q)).$$

We have previously constructed the truth table of the formula ψ.

p	q	r	ψ	$\neg\psi$
1	1	1	0	1
1	1	0	0	1
1	0	1	0	1
1	0	0	**1**	0
0	1	1	**1**	0
0	1	0	**1**	0
0	0	1	0	1
0	0	0	0	1

Now let $f_\psi : \{0,1\}^3 \to \{0,1\}$ be the function given by its truth table. Let

$$\Delta = \{\bar{e} \in \{0,1\}^3 \mid f_\psi(\bar{e}) = 1\} = \{(1,0,0),\ (0,1,1),\ (0,1,0)\}.$$

The formula in DNF logically equivalent to ψ is

$$\begin{aligned}\varphi = \bigvee_{\bar{e}\in\Delta} \chi_{\bar{e}} &= \chi_{(1,0,0)} \vee \chi_{(0,1,1)} \vee \chi_{(0,1,0)} \\ &= (p^1 \wedge q^0 \wedge r^0) \vee (p^0 \wedge q^1 \wedge r^1) \vee (p^0 \wedge q^1 \wedge r^0) \\ &= (p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge q \wedge r) \vee (\neg p \wedge q \wedge \neg r).\end{aligned}$$

Next, to find a formula in CNF logically equivalent to ψ, we first find a formula θ in DNF logically equivalent to $\neg\psi$.

$$\begin{aligned}\neg\psi \equiv \theta = \bigvee_{\bar{e}\notin\Delta} \chi_{\bar{e}} &= \chi_{(1,1,1)} \vee \chi_{(1,1,0)} \vee \chi_{(1,0,1)} \vee \chi_{(0,0,1)} \vee \chi_{(0,0,0)} \\ &= (p^1 \wedge q^1 \wedge r^1) \vee (p^1 \wedge q^1 \wedge r^0) \vee (p^1 \wedge q^0 \wedge r^1) \vee (p^0 \wedge q^0 \wedge r^1) \vee (p^0 \wedge q^0 \wedge r^0) \\ &= (\neg p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge \neg q \wedge r) \vee (\neg p \wedge q \wedge \neg r) \vee (p \wedge q \wedge \neg r) \vee (p \wedge q \wedge r).\end{aligned}$$

Finally,

$$\begin{aligned}\psi &\equiv \neg\theta \\ &\equiv \neg\Big((\neg p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge \neg q \wedge r) \vee (\neg p \wedge q \wedge \neg r) \vee (p \wedge q \wedge \neg r) \vee (p \wedge q \wedge r)\Big) \\ &\equiv \neg(\neg p \wedge \neg q \wedge \neg r) \wedge \neg(\neg p \wedge \neg q \wedge r) \wedge \neg(\neg p \wedge q \wedge \neg r) \wedge \neg(p \wedge q \wedge \neg r) \wedge \neg(p \wedge q \wedge r) \\ &\equiv (p \vee q \vee r) \wedge (p \vee q \vee \neg r) \wedge (p \vee \neg q \vee r) \wedge (\neg p \vee \neg q \vee r) \wedge (\neg p \vee \neg q \vee \neg r).\end{aligned}$$

Clearly, the last formula is in CNF.

We have shown that

$$\begin{aligned}\psi &= ((q \vee (\neg r \wedge p)) \leftrightarrow (p \to \neg q)) \\ &\equiv (p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge q \wedge r) \vee (\neg p \wedge q \wedge \neg r) \\ &\equiv (p \vee q \vee r) \wedge (p \vee q \vee \neg r) \wedge (p \vee \neg q \vee r) \wedge (\neg p \vee \neg q \vee r) \wedge (\neg p \vee \neg q \vee \neg r).\end{aligned}$$

♠

Definition. A set C of propositional connectives is called *complete* if every propositional formula is logically equivalent to one which can be written using only connectives from C.

We know that every formula is logically equivalent to one in DNF, and formulas in DNF use only the connectives $\vee$, $\wedge$, and $\neg$. This gives us the following result.

Corollary 1.2.15. *The set $\{\neg, \wedge, \vee\}$ is a complete set of propositional connectives.*

Suppose that θ is a formula constructed using only the connectives $\neg$, $\wedge$, and $\vee$. Since $(\varphi \vee \psi) \equiv \neg(\neg\varphi \wedge \neg\psi)$ for any formulas φ and ψ, and using Theorem 1.2.8, we may replace in θ every subformula of the form $(\varphi \vee \psi)$ with the formula $\neg(\neg\varphi \wedge \neg\psi)$ and get a new logically equivalent formula that only contains the propositional connectives $\neg$ and $\wedge$. This discussion gives the following result.

Corollary 1.2.16. *The set $\{\neg, \wedge\}$ is a complete set of propositional connectives.*

Furthermore, for any formulas φ and ψ we have $(\varphi \wedge \psi) \equiv \neg(\varphi \to \neg\psi)$. This gives the following result.

Corollary 1.2.17. *The set $\{\neg, \to\}$ is a complete set of propositional connectives.*

Remark. There is a binary connective called *Sheffer stroke* or *NAND* that is denoted by $\uparrow$ whose truth table is given by the condition $\delta[(\varphi \uparrow \psi)] = 0$ exactly when $\delta[\varphi] = 1$ and $\delta[\psi] = 1$ for any truth assignment δ.

Sheffer Stroke Truth Table

φ	ψ	$(\varphi \uparrow \psi)$
1	1	0
1	0	1
0	1	1
0	0	1

One can check the following.

- $\neg\varphi \equiv (\varphi \uparrow \varphi)$.
- $(\varphi \wedge \psi) \equiv ((\varphi \uparrow \psi) \uparrow (\varphi \uparrow \psi))$.
- $(\varphi \vee \psi) \equiv ((\varphi \uparrow \varphi) \uparrow (\psi \uparrow \psi))$.

Thus any propositional formula is logically equivalent to one which only uses the Sheffer stroke.

Corollary 1.2.18. *The set $\{\uparrow\}$ is a complete set of propositional connectives.*

1.3 Logical Consequence

Recall that we say a truth assignment δ satisfies a formula φ when $\delta[\varphi] = 1$, and φ is said to be satisfiable if some truth assignment satisfies φ. We can generalize satisfiability to a set of formulas as follows.

Definition. A set Γ of propositional formulas is *satisfiable* if there exists a truth assignment δ such that $\delta[\gamma] = 1$ for all $\gamma \in \Gamma$.

For example, the set $\{p,\ q,\ (\neg p \vee q)\}$ is satisfiable, however, the set $\{p,\ \neg q,\ (p \to q)\}$ is not satisfiable. The empty set is satisfied by any truth assignment. The next notion describes the phenomenon when the satisfiability of a set of formulas imposes the satisfiability of some formula.

Definition.

- A propositional formula φ *logically implies* a formula ψ if for any truth assignment δ, whenever δ satisfies φ, then δ satisfies ψ. We also say ψ is a *logical consequence* of φ. The notation for logical implication is

$$\varphi \models \psi.$$

- Let Γ be a set of propositional formulas and ψ be a formula. We say that Γ *logically implies* ψ if for any truth assignment δ, whenever $\delta[\gamma] = 1$ for every $\gamma \in \Gamma$, then $\delta[\psi] = 1$. We also say ψ is a *logical consequence* of Γ and write

$$\Gamma \models \psi.$$

- When ψ is not a *logical consequence* of Γ, we write $\Gamma \not\models \psi$.

Remark. If $\Gamma \models \psi$ and Γ is the empty set, we simply write $\models \psi$ and in this case ψ is a tautology since any truth assignments satisfies the empty set.

Example.

- (Modus Ponens) Let φ, ψ be any formulas. Show that

$$\{(\varphi \to \psi),\ \varphi\} \models \psi.$$

Let $\delta : \mathcal{F} \to \{0,1\}$ be any truth assignment where $\delta[(\varphi \to q)] = 1$ and $\delta[\varphi] = 1$. Then

$$1 = \delta[(\varphi \to \psi)] = 1 + \delta[\varphi] + \delta[\varphi] \cdot \delta[\psi] = 1 + 1 + 1 \cdot \delta[\psi] = \delta[\psi].$$

Thus, $\delta[\psi] = 1$ as desired.

- Show that
$$\{(p \wedge r),\, (r \to (p \wedge q))\} \models (q \vee t).$$
Let $\delta : \mathcal{F} \to \{0,1\}$ be any truth assignment where $\delta[(p \wedge r)] = 1$ and $\delta[(r \to (p \wedge q))] = 1$. Consequently, we must have $\delta[r] = 1$. Furthermore,
$$1 = \delta[(r \to (p \wedge q))] = 1 + \delta[r] + \delta[r] \cdot \delta[(p \wedge q)] = 1 + 1 + 1 \cdot \delta[(p \wedge q)] = \delta[(p \wedge q)].$$
So $\delta[(p \wedge q)] = 1$. This forces $\delta[q] = 1$. Finally,
$$\delta[(q \vee t)] = \delta[q] + \delta[t] + \delta[q] \cdot \delta[t] = 1 + \delta[t] + \delta[t] = 1.$$
Thus, δ satisfies $(q \vee t)$ as desired.

- Show that $\{q,\, (r \to \neg p)\} \not\models (q \to r)$.
Consider a truth assignment λ where $\lambda[p] = 1$, $\lambda[q] = 1$, and $\lambda[r] = 0$. Then λ satisfies both q and $(r \to \neg p)$, but does not satisfy $(q \to r)$.

Lemma 1.3.1. *Let φ and ψ be propositional formulas. Then the following hold.*

- $\varphi \models \psi$ *if and only if* $\models (\varphi \to \psi)$.
- $\varphi \equiv \psi$ *if and only if* $\varphi \models \psi$ *and* $\psi \models \varphi$.
- $\Gamma \cup \{\varphi\} \models \psi$ *if and only if* $\Gamma \models (\varphi \to \psi)$.

Proof. We prove the first one. For the forward direction, suppose that $\varphi \models \psi$. We will show that $(\varphi \to \psi)$ is a tautology. Let $\delta : \mathcal{F} \to \{0,1\}$ be an arbitrary truth assignment. If $\delta[\varphi] = 0$, then $\delta[(\varphi \to \psi)] = 1 + \delta[\varphi] + \delta[\varphi] \cdot \delta[\psi] = 1 + 0 + 0 = 1$. Otherwise, $\delta[\varphi] = 1$, and since $\varphi \models \psi$, we must have that $\delta[\psi] = 1$ as well. Then the truth value of the implication is $\delta[(\varphi \to \psi)] = 1 + \delta[\varphi] + \delta[\varphi] \cdot \delta[\psi] = 1 + 1 + 1 = 1$. Thus $(\varphi \to \psi)$ is a tautology as desired.

For the reverse direction, suppose that $(\varphi \to \psi)$ is a tautology. We will show that $\varphi \models \psi$. Take any truth assignment δ where $\delta[\varphi] = 1$. As $\delta[(\varphi \to \psi)] = 1$ we get that
$$1 = \delta[(\varphi \to \psi)] = 1 + \delta[\varphi] + \delta[\varphi] \cdot \delta[\psi] = 1 + 1 + \delta[\psi] = \delta[\psi].$$
Thus, $\delta[\psi] = 1$ as well. This shows that $\varphi \models \psi$. ■

Chapter 2

Soundness and Completeness

The main objective of this chapter is to show that propositional logic admits a proof system which is both sound and complete. These new concepts will gradually become lucid through the chapter.

2.1 Proof Systems

We aim to formalize what constitutes a mathematical proof. In order to formulate mathematical reasoning in a precise way we introduce proof systems which have three constituent parts: (1) a formal language (e.g. propositional logic or first-order logic), (2) a set of axioms, and (3) a set of deduction rules. Using the deduction rules we can manipulate strings of symbols (e.g. formulas) and consequently deduce new formulas from older formulas. There are several types of proof systems in logic, for example, *Hilbert-style* proof systems (many axioms, few deductions rules), proof systems of *Natural Deduction* (few axioms, many deductions rules), and proof systems of *Sequent Calculus* (used in automated reasoning).

Our objective is to capture the idea of a formula ψ being derived from a set Γ of formulas within a particular proof system. Such a derivation is a finite sequence of formulas ending with ψ and where each step of the derivation arises in one of the following ways:

(i) as one of the formulas in Γ;

(ii) as one of the axioms of the proof system (an axiom is a formula previously agreed as allowable in *any* derivation within the system);

(iii) as a consequence of applying a deduction rule of the system to formulas already derived.

We will introduce a Hilbert-style proof system denoted by $\mathcal{S}$ suitable for propositional logic. The proof system will manipulate formulas which contain only the connectives

$\neg$ and $\to$. This can be later justified as the set $\{\neg, \to\}$ is a complete set of propositional connectives, meaning that every propositional formula is logically equivalent to a formula which can be written using only $\neg$ and $\to$ from the connectives.

Let $\mathcal{F}^* \subseteq \mathcal{F}$ be the set of all formulas whose connectives are among $\neg$ and $\to$. The set $\mathcal{F}^*$ is constructed as follows.

Definition.

- We set $\mathcal{F}^*_0 = \mathbf{P}$.
- For each natural number n, we define
$$\mathcal{F}^*_{n+1} = \mathcal{F}^*_n \ \cup\ \{\neg\varphi \mid \varphi \in \mathcal{F}^*_n\} \ \cup\ \{(\varphi \to \psi) \mid \varphi, \psi \in \mathcal{F}^*_n\}.$$
- We define the set $\mathcal{F}^*$ to be
$$\mathcal{F}^* = \bigcup_{n \in \mathbb{N}} \mathcal{F}^*_n.$$

We now introduce the axioms and the deduction rules of our system $\mathcal{S}$. There are three axioms and one deduction rule called *modus ponens*.

Axioms of $\mathcal{S}$

The proof system $\mathcal{S}$ has the following three schemes of axioms.

(Ax 1) $\big(\varphi \to (\psi \to \varphi)\big)$ for any φ, ψ in $\mathcal{F}^*$.

(Ax 2) $\big((\varphi \to (\psi \to \theta)) \to ((\varphi \to \psi) \to (\varphi \to \theta))\big)$ for any φ, ψ, θ in $\mathcal{F}^*$.

(Ax 3) $\big((\neg\varphi \to \neg\psi) \to (\psi \to \varphi)\big)$ for any φ, ψ in $\mathcal{F}^*$.

Deduction rules of $\mathcal{S}$

The proof system $\mathcal{S}$ has one deduction rule.

Modus Ponens. From formulas φ and $(\varphi \to \psi)$, derive the formula ψ.

$$\text{(MP)} \quad \frac{\varphi,\ (\varphi \to \psi)}{\psi}$$

Once we have a proof system we can formalize the notion of a proof in mathematics. We will call this formalization a *formal proof* or *derivation* or *deduction*.

Definition. Let Γ be a set of propositional formulas and let ψ be a propositional formula. A *derivation* (or a *proof*) of ψ from Γ within the system $\mathcal{S}$ is a finite sequence of propositional formulas

$$\varphi_1,\ \varphi_2,\ \ldots,\ \varphi_n,$$

where the last formula φ_n in the sequence is the formula ψ and where each formula φ_k in the sequence satisfies one of the following:

(i) $\varphi_k \in \Gamma$;

(ii) φ_k is one of the axioms of the system $\mathcal{S}$;

(iii) there are formulas φ_i and φ_j in the sequence such that $i < k$ and $j < k$, and $\varphi_j = (\varphi_i \to \varphi_k)$.

Definition. We say that ψ is *derivable* (or *provable*) from Γ if there exists a derivation of ψ from Γ within the system $\mathcal{S}$. We write

$$\Gamma \vdash \psi$$

when ψ is derivable from Γ.

Remark.

- We call the formulas in the set Γ *assumptions*. If $\Gamma \vdash \psi$ and $\Gamma = \emptyset$, then we write $\vdash \psi$ and say ψ is a *theorem of the system $\mathcal{S}$*.
- $\Gamma \nvdash \psi$ means that ψ is not provable from Γ, that is, there exists no derivation of ψ from Γ.

Example. Show that

$$\{p,\, (q \to r)\} \vdash (r \to p)$$

φ_1	p	Assumption
φ_2	$(p \to (r \to p))$	Ax 1
φ_3	$(r \to p)$	MP applied to φ_1, φ_2

♠

Example. Let α be any formula in $\mathcal{F}^*$. Show that $(\alpha \to \alpha)$ is a theorem of $\mathcal{S}$, i.e.,

$$\vdash (\alpha \to \alpha).$$

1.	$\big(\alpha \to (\alpha \to \alpha)\big)$	Ax 1
2.	$\big(\alpha \to ((\alpha \to \alpha) \to \alpha)\big)$	Ax 1
3.	$\big((\alpha \to ((\alpha \to \alpha) \to \alpha)) \to ((\alpha \to (\alpha \to \alpha)) \to (\alpha \to \alpha))\big)$	Ax 2
4.	$((\alpha \to (\alpha \to \alpha)) \to (\alpha \to \alpha))$	MP 2, 3
5.	$(\alpha \to \alpha)$	MP 1, 4

♠

The next example shows that, within the system $\mathcal{S}$, from a formula and its negation one can derive any formula!

Example. Let α, β be any formulas in $\mathcal{F}^*$. Show that

$$\{\alpha,\ \neg\alpha\} \vdash \beta.$$

1.	α	Assumption
2.	$\neg\alpha$	Assumption
3.	$(\neg\alpha \to (\neg\beta \to \neg\alpha))$	Ax 1
4.	$(\neg\beta \to \neg\alpha)$	MP 2, 3
5.	$((\neg\beta \to \neg\alpha) \to (\alpha \to \beta))$	Ax 3
6.	$(\alpha \to \beta)$	MP 4, 5
7.	β	MP 1, 6

♠

Example. Let φ, ψ, θ be any formulas in $\mathcal{F}^*$. Show that

$$\{(\varphi \to (\psi \to \theta)),\ \psi\} \vdash (\varphi \to \theta).$$

1.	$(\varphi \to (\psi \to \theta))$	Assumption
2.	$((\varphi \to (\psi \to \theta)) \to ((\varphi \to \psi) \to (\varphi \to \theta)))$	Ax 2
3.	$((\varphi \to \psi) \to (\varphi \to \theta))$	MP 1, 2
4.	$(\psi \to (\varphi \to \psi))$	Ax 1
5.	ψ	Assumption
6.	$(\varphi \to \psi)$	MP 4, 5
7.	$(\varphi \to \theta)$	MP 3, 6

♠

Example. Let α be any formula. Show that

$$\{\neg\neg\alpha\} \vdash \alpha.$$

1.	$((\neg\neg\neg\neg\alpha \to \neg\neg\alpha) \to (\neg\alpha \to \neg\neg\neg\alpha))$	Ax 3
2.	$((\neg\alpha \to \neg\neg\neg\alpha) \to (\neg\neg\alpha \to \alpha))$	Ax 3
3.	$(\neg\neg\alpha \to (\neg\neg\neg\neg\alpha \to \neg\neg\alpha))$	Ax 1
4.	$\neg\neg\alpha$	Assumption
5.	$(\neg\neg\neg\neg\alpha \to \neg\neg\alpha)$	MP 3, 4
6.	$(\neg\alpha \to \neg\neg\neg\alpha)$	MP 1, 5
7.	$(\neg\neg\alpha \to \alpha)$	MP 2, 6
8.	α	MP 4, 7

♠

Lemma 2.1.1. *Let $\Gamma \subseteq \mathcal{F}^*$ be a set of formulas, and let ψ be a formula in $\mathcal{F}^*$.*

1. *If $\varphi_1, \varphi_2, \ldots, \varphi_n$ is derivation from Γ within the proof system $\mathcal{S}$, then so is the sequence $\varphi_1, \varphi_2, \ldots, \varphi_k$ for every $k \leq n$.*

2. *If $\Gamma \subseteq \Delta \subseteq \mathcal{F}^*$ and $\Gamma \vdash \psi$, then $\Delta \vdash \psi$.*

2.2 The Deduction Theorem

We may call theorems about the proof system $\mathcal{S}$ *metatheorems* to distinguish them from the formal theorems that are derived within the system $\mathcal{S}$. In this section we show an important metatheoreom of $\mathcal{S}$ called the deduction theorem.

Lemma 2.2.1. *Let Γ be a set of formulas from $\mathcal{F}^*$, and φ, ψ be formulas from $\mathcal{F}^*$. If $\Gamma \vdash (\varphi \to \psi)$, then $\Gamma \cup \{\varphi\} \vdash \psi$.*

Proof. Suppose that $\Gamma \vdash (\varphi \to \psi)$. This means that there exists a derivation of $(\varphi \to \psi)$ from Γ, say the derivation is the sequence $\alpha_1, \alpha_2, \ldots, \alpha_n$. We know that $\alpha_n = (\varphi \to \psi)$, and each α_i is either from Γ, an axiom, or deduced by Modus Ponens. We add two more legal steps to this derivation to create a new derivation for ψ from the set $\Gamma \cup \{\varphi\}$ as follows.

1.	α_1	*
2.	α_2	*
$\vdots$	$\vdots$	$\vdots$
$n-1$.	α_{n-1}	*
n.	$(\varphi \to \psi)$	*
$n+1$.	φ	Assumption since $\varphi \in \Gamma \cup \{\varphi\}$
$n+2$.	ψ	MP applied to steps n and $n+1$

Therefore, $\Gamma \cup \{\varphi\} \vdash \psi$ as desired. ■

The converse of the previous lemma is a crucial property of a proof system. It is an extremely useful tool in constructing derivations. This property is well known as the *Deduction Theorem* and it states that if $\Gamma \cup \{\varphi\} \vdash \psi$, then $\Gamma \vdash (\varphi \to \psi)$. Such important property formalizes what we do in everyday mathematics when we aim to prove an implication 'if φ, then ψ'. Towards proving an implication, we assume the hypothesis φ, and then work our way to prove the conclusion ψ. Once this has been accomplished successfully, one concludes that 'if φ, then ψ' holds.

Theorem 2.2.2 (Deduction Theorem for $\mathcal{S}$)**.** *Let Γ be a set of formulas from $\mathcal{F}^*$, and φ, ψ be formulas from $\mathcal{F}^*$.*

$$\textit{If } \Gamma \cup \{\varphi\} \vdash \psi, \textit{ then } \Gamma \vdash (\varphi \to \psi).$$

Proof. Let Γ be a set of formulas from $\mathcal{F}^*$, and φ be a formula from $\mathcal{F}^*$. We will prove the theorem by *mathematical induction on the length of the derivation.* More precisely, we will prove by induction on n the following statement:

$$\text{If } \alpha_1, \alpha_2, \ldots, \alpha_n \text{ is a derivation in } \mathcal{S} \text{ from } \Gamma \cup \{\varphi\}, \text{ then } \Gamma \vdash (\varphi \to \alpha_n).$$

Base case. Suppose $n = 1$. Suppose that α_1 is a derivation from $\Gamma \cup \{\varphi\}$. So the derivation sequence consists only of α_1. It follows that α_1 is either an axiom or an assumption, more precisely, three cases arise where in each case we need to show that $\Gamma \vdash (\varphi \to \alpha_1)$:

(i) α_1 is an axiom.

1.	α_1	Axiom
2.	$(\alpha_1 \to (\varphi \to \alpha_1))$	Ax 1
3.	$(\varphi \to \alpha_1)$	MP 1, 2

(ii) $\alpha_1 \in \Gamma$.

1.	α_1	Assumption from Γ
2.	$(\alpha_1 \to (\varphi \to \alpha_1))$	Ax 1
3.	$(\varphi \to \alpha_1)$	MP 1, 2

(iii) $\alpha_1 = \varphi$.
We proved earlier that $(\varphi \to \varphi)$ is a theorem of the system $\mathcal{S}$, that is, it can be derived using only the axioms and the deduction rule. Thus, $\Gamma \vdash (\varphi \to \varphi)$.

Thus, in all these three cases, we have shown that $\Gamma \vdash (\varphi \to \alpha_1)$ completing the base case.

Induction step. Suppose the result is true for *all* derivations of length n or less. We will show that the result holds for derivations of length $n + 1$, so let

$$\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n,\ \alpha_{n+1}$$

be a derivation from $\Gamma \cup \{\varphi\}$ in the system $\mathcal{S}$. To complete the induction step, we need to show that $\Gamma \vdash (\varphi \to \alpha_{n+1})$. Since the sequence is a derivation we know that α_{n+1} is either an axiom or $\alpha_{n+1} \in \Gamma \cup \{\varphi\}$ or deduced by Modus Ponens. In the first two cases we argue in a similar fashion as in the base case.

We are now left with the possibility that α_{n+1} was deduced by Modus Ponens. Thus, there are formulas α_i and α_j in the sequence such that $i < n + 1$ and $j < n + 1$ and $\alpha_j = (\alpha_i \to \alpha_{n+1})$. Since both $\alpha_1, \alpha_2, \ldots, \alpha_i$ and $\alpha_1, \alpha_2, \ldots, \alpha_j$ are derivations of length at most n from $\Gamma \cup \{\varphi\}$ we know, by induction hypothesis, that $\Gamma \vdash (\varphi \to \alpha_i)$ and $\Gamma \vdash (\varphi \to \alpha_j)$. Let $\beta_1, \beta_2, \ldots, \beta_k$ be a derivation of $(\varphi \to \alpha_i)$ from Γ and let $\theta_1, \theta_2, \ldots, \theta_m$ be a derivation of $(\varphi \to \alpha_j)$ from Γ. We now construct the following derivation:

1.	β_1	$*$
2.	β_2	$*$
$\vdots$	$\vdots$	$\vdots$
k.	$\beta_k = (\varphi \to \alpha_i)$	$*$
$k+1$.	θ_1	$*$
$k+2$.	θ_2	$*$
$\vdots$	$\vdots$	$\vdots$
$k+m$.	$\theta_m = (\varphi \to (\alpha_i \to \alpha_{n+1}))$	$*$
$k+m+1$.	$((\varphi \to (\alpha_i \to \alpha_{n+1})) \to ((\varphi \to \alpha_i) \to (\varphi \to \alpha_{n+1})))$	Ax 2
$k+m+2$.	$((\varphi \to \alpha_i) \to (\varphi \to \alpha_{n+1}))$	MP $k+m,\ k+m+1$
$k+m+3$.	$(\varphi \to \alpha_{n+1})$	MP $k,\ k+m+2$

This is a derivation of $(\varphi \to \alpha_{n+1})$ from Γ, and so $\Gamma \vdash (\varphi \to \alpha_{n+1})$ as desired. This completes the induction step. Therefore, we have shown by induction that if $\alpha_1, \alpha_2, \ldots, \alpha_n$ is a derivation from $\Gamma \cup \{\varphi\}$, then $\Gamma \vdash (\varphi \to \alpha_n)$ for every positive integer n.

Now to prove the theorem, let ψ be any formula and assume that $\Gamma \cup \{\varphi\} \vdash \psi$. It follows that there exists a derivation $\alpha_1, \alpha_2, \ldots, \alpha_n$ of ψ from $\Gamma \cup \{\varphi\}$. By the above we get that $\Gamma \vdash (\varphi \to \alpha_n)$, but $\alpha_n = \psi$, so $\Gamma \vdash (\varphi \to \psi)$ as desired. ■

Example. We proved earlier that $\{\neg\alpha, \alpha\} \vdash \beta$ for any formulas $\alpha, \beta \in \mathcal{F}^*$. By the deduction theorem, we get $\{\neg\alpha\} \vdash (\alpha \to \beta)$. By another application of the deduction theorem we get

$$\vdash (\neg\alpha \to (\alpha \to \beta)).$$

♠

Example (Double-negation Elimination)**.** We proved earlier that $\{\neg\neg\alpha\} \vdash \alpha$ for any formula $\alpha \in \mathcal{F}^*$. By the deduction theorem we get

$$\vdash (\neg\neg\alpha \to \alpha).$$

♠

Example (Double-negation Introduction)**.** Let β be any formula in $\mathcal{F}^*$. Consider the following derivation in the system $\mathcal{S}$.

1.	$(\neg\neg\neg\beta \to \neg\beta)$	Double-negation Elimination
2.	$((\neg\neg\neg\beta \to \neg\beta) \to (\beta \to \neg\neg\beta))$	Ax 3
3.	$(\beta \to \neg\neg\beta)$	MP 1, 2

Therefore,

$$\vdash (\beta \to \neg\neg\beta).$$

♠

Lemma 2.2.3. *Let φ be any formula in $\mathcal{F}^*$. Then within the system $\mathcal{S}$ we have*

$$\vdash ((\neg\varphi \to \varphi) \to \varphi).$$

Proof. Let φ be any formula in $\mathcal{F}^*$. We proved earlier that $\vdash (\neg\alpha \to (\alpha \to \beta))$ for any formulas $\alpha, \beta \in \mathcal{F}^*$. Take α to be the formula φ and β to be the formula $\neg(\neg\varphi \to \varphi)$ to get

$$\vdash (\neg\varphi \to (\varphi \to \neg(\neg\varphi \to \varphi))).$$

We now show that $\{(\neg\varphi \to \varphi)\} \vdash \varphi$ by construction the following derivation.

1.	$(\neg\varphi \to (\varphi \to \neg(\neg\varphi \to \varphi)))$	Theorem of $\mathcal{S}$
2.	$((\neg\varphi \to (\varphi \to \neg(\neg\varphi \to \varphi))) \to ((\neg\varphi \to \varphi) \to (\neg\varphi \to \neg(\neg\varphi \to \varphi))))$	Ax 2
3.	$((\neg\varphi \to \varphi) \to (\neg\varphi \to \neg(\neg\varphi \to \varphi)))$	MP 1, 2
4.	$(\neg\varphi \to \varphi)$	Assumption
5.	$(\neg\varphi \to \neg(\neg\varphi \to \varphi))$	MP 3, 4
6.	$((\neg\varphi \to \neg(\neg\varphi \to \varphi)) \to ((\neg\varphi \to \varphi) \to \varphi))$	Ax 3
7.	$((\neg\varphi \to \varphi) \to \varphi)$	MP 5, 6
8.	φ	MP 4, 7

The derivation above shows that $\{(\neg\varphi \to \varphi)\} \vdash \varphi$, and by the deduction theorem it follows that $\vdash ((\neg\varphi \to \varphi) \to \varphi)$, that is, the formula $((\neg\varphi \to \varphi) \to \varphi)$ is a theorem of the system $\mathcal{S}$. ■

Lemma 2.2.4. *If $\Gamma \vdash \varphi$ and $\Delta \cup \{\varphi\} \vdash \psi$, then $\Gamma \cup \Delta \vdash \psi$.*

Here is another important metatheorem of the system $\mathcal{S}$.

Theorem 2.2.5 (Proof by Contradiction)**.** *Let Γ be a set of formulas from $\mathcal{F}^*$, and φ, ψ be formulas from $\mathcal{F}^*$.*

If $\Gamma \cup \{\neg\varphi\} \vdash \psi$ and $\Gamma \cup \{\neg\varphi\} \vdash \neg\psi$, then $\Gamma \vdash \varphi$.

Proof. Assume that $\Gamma \cup \{\neg\varphi\} \vdash \psi$ and $\Gamma \cup \{\neg\varphi\} \vdash \neg\psi$. Applying the deduction theorem to the second one we obtain $\Gamma \vdash (\neg\varphi \to \neg\psi)$. Now we add additional steps to this derivation to obtain the following derivation from Γ.

1.	$(\neg\varphi \to \neg\psi)$	*
2.	$((\neg\varphi \to \neg\psi) \to (\psi \to \varphi))$	Ax 3
3.	$(\psi \to \varphi)$	MP 1, 2

Therefore, $\Gamma \vdash (\psi \to \varphi)$. By Lemma 2.2.1, we obtain that $\Gamma \cup \{\psi\} \vdash \varphi$. Since $\Gamma \cup \{\neg\varphi\} \vdash \psi$ and $\Gamma \cup \{\psi\} \vdash \varphi$, by Lemma 2.2.4, it follows that $\Gamma \cup \{\neg\varphi\} \vdash \varphi$. By the deduction theorem we obtain that $\Gamma \vdash (\neg\varphi \to \varphi)$. Adding the following steps to this derivation we get the following derivation from Γ.

1.	$(\neg\varphi \to \varphi)$	*
2.	$((\neg\varphi \to \varphi) \to \varphi)$	Lemma 2.2.3
3.	φ	MP 1, 2

Therefore, $\Gamma \vdash \varphi$ as desired. ■

Theorem 2.2.6 (Proof by Contradiction II)**.** *Let Γ be a set of formulas from $\mathcal{F}^*$, and φ, ψ be formulas from $\mathcal{F}^*$.*

$$\textit{If } \Gamma \cup \{\varphi\} \vdash \psi \textit{ and } \Gamma \cup \{\varphi\} \vdash \neg\psi, \textit{ then } \Gamma \vdash \neg\varphi.$$

Proof. Assume that $\Gamma \cup \{\varphi\} \vdash \psi$ and $\Gamma \cup \{\varphi\} \vdash \neg\psi$. We proved earlier that $\{\neg\neg\varphi\} \vdash \varphi$. Using Lemma 2.2.4, from $\{\neg\neg\varphi\} \vdash \varphi$ and $\Gamma \cup \{\varphi\} \vdash \psi$ we obtain $\Gamma \cup \{\neg\neg\varphi\} \vdash \psi$. Similarly, from $\{\neg\neg\varphi\} \vdash \varphi$ and $\Gamma \cup \{\varphi\} \vdash \neg\psi$ we obtain $\Gamma \cup \{\neg\neg\varphi\} \vdash \neg\psi$. By the first version of Proof by Contradiction, we obtain that $\Gamma \vdash \neg\varphi$. ■

Example. Show that $\{\neg(\theta \to \neg\psi)\} \vdash \psi$.

We will use proof by contradiction. Towards this we will show that $\{\neg(\theta \to \neg\psi), \neg\psi\} \vdash \chi$ and $\{\neg(\theta \to \neg\psi), \neg\psi\} \vdash \neg\chi$ for some formula χ. The following derivation shows that $\{\neg(\theta \to \neg\psi), \neg\psi\} \vdash (\theta \to \neg\psi)$.

1.	$\neg(\theta \to \neg\psi)$	Assumption
2.	$\neg\psi$	Assumption
3.	$(\neg\psi \to (\theta \to \neg\psi))$	Ax 1
4.	$(\theta \to \neg\psi)$	MP 2, 3

Moreover, it is obvious that $\{\neg(\theta \to \neg\psi), \neg\psi\} \vdash \neg(\theta \to \neg\psi)$. Therefore, by proof by contradiction, we conclude that $\{\neg(\theta \to \neg\psi)\} \vdash \psi$. ♠

2.3 The Soundness Theorem

Our goal is to show that our proof system $\mathcal{S}$ enjoys two important metatheorems: soundness and completeness. Together they show that the notions of provability (syntax) and logical consequence (semantics) match together. Notice that these two notions describe the situation when one statement follows from a set of statements. We first show that the system $\mathcal{S}$ is sound, meaning that a derivation in $\mathcal{S}$ corresponds to a logical consequence.

Lemma 2.3.1. *All instances of the three axioms of the proof system $\mathcal{S}$ are tautologies.*

- $(\varphi \to (\psi \to \varphi))$
- $((\varphi \to (\psi \to \theta)) \to ((\varphi \to \psi) \to (\varphi \to \theta)))$
- $((\neg\varphi \to \neg\psi) \to (\psi \to \varphi))$

Proof. We will show that Ax 1 is a tautology. Choose any formulas φ and ψ and let $\delta : \mathcal{F} \to \{0,1\}$ be any truth assignment. Then

$$\begin{aligned}
\delta[(\varphi \to (\psi \to \varphi))] &= 1 + \delta[\varphi] + \delta[\varphi] \cdot \delta[(\psi \to \varphi)] \\
&= 1 + \delta[\varphi] + \delta[\varphi] \cdot (1 + \delta[\psi] + \delta[\psi] \cdot \delta[\varphi]) \\
&= 1 + \delta[\varphi] + \delta[\varphi] + \delta[\varphi] \cdot \delta[\psi] + \delta[\psi] \cdot \delta[\varphi] \\
&= 1 + \delta[\varphi] \cdot (1+1) + (\delta[\varphi] \cdot \delta[\psi]) \cdot (1+1) \\
&= 1 + 0 + 0 = 1.
\end{aligned}$$

■

Theorem 2.3.2 (Soundess theorem for $\mathcal{S}$). *Let Γ be a set of propositional formulas and let ψ be any propositional formula. If Γ proves ψ, then Γ logically implies ψ. In symbols,*

$$\text{If } \Gamma \vdash \psi, \text{ then } \Gamma \models \psi.$$

Proof. We will show by mathematical induction on the length n of the derivation the following:

If $\alpha_1, \alpha_2, \ldots, \alpha_n$ is a derivation from Γ in the system $\mathcal{S}$ and δ is a truth assignment that satisfies Γ, then δ satisfies α_n.

Base case. Suppose that $n = 1$ and let δ be a truth assignment which satisfies Γ, that is, $\delta[\gamma] = 1$ for every $\gamma \in \Gamma$. We need to show that $\delta[\alpha_1] = 1$. Observe that α_1 being the first formula of a derivation either belongs to Γ or is an axiom. For the former case, as δ satisfies every formula in Γ we get $\delta[\alpha_1] = 1$. For the latter case, by Lemma 2.3.1, we know that axioms of $\mathcal{S}$ are tautologies, and so $\delta[\alpha_1] = 1$ as well.

Induction step. Suppose that result holds for all derivations of length at most n. That is, if $\beta_1, \beta_2, \ldots, \beta_k$ is a derivation from Γ and $k \leq n$ and δ satisfies Γ, then $\delta[\beta_k] = 1$. Consider a derivation $\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_{n+1}$ from Γ and let δ be a truth assignment which satisfies Γ. We need to show that δ satisfies α_{n+1}. The formula α_{n+1} being in a derivation either belongs to Γ or an axiom or derived by Modus Ponens. If α_{n+1} is in Γ or an axiom, then δ satisfies α_{n+1} as in the base case.

Otherwise, α_{n+1} is deduced in the derivation by an application of Modus Ponens. So there are formulas α_i and α_j in the derivation such that $i < n+1$ and $j < n+1$ and $\alpha_j = (\alpha_i \rightarrow \alpha_{n+1})$. Since both $\alpha_1, \alpha_2, \ldots, \alpha_i$ and $\alpha_1, \alpha_2, \ldots, \alpha_j$ are derivations of length at most n from Γ we know, by induction hypothesis, that $\delta[\alpha_i] = 1$ and $\delta[\alpha_j] = 1$. Therefore,

$$1 = \delta[\alpha_j] = \delta[(\alpha_i \rightarrow \alpha_{n+1})] = 1 + \delta[\alpha_i] + \delta[\alpha_i] \cdot \delta[\alpha_{n+1}] = 1 + 1 + \delta[\alpha_{n+1}] = \delta[\alpha_{n+1}].$$

Thus, δ satisfies α_{n+1}. This completes the induction step.

Finally to show the soundness of the system $\mathcal{S}$, assume that $\Gamma \vdash \psi$. We need to show that $\Gamma \models \psi$. Towards this end, let δ be any derivation which satisfies Γ. Since $\Gamma \vdash \psi$, there exists a derivation $\alpha_1, \alpha_2, \ldots, \alpha_n$ of ψ from Γ in the system $\mathcal{S}$. It follows that $\alpha_n = \psi$, and by the above we obtain that δ satisfies ψ. So every truth assignment which satisfies Γ satisfies ψ, that is, $\Gamma \models \psi$. ■

If we take Γ in the soundness theorem to be the empty set we obtain the following.

Corollary 2.3.3. *Any theorem of the system $\mathcal{S}$ is a tautology. In symbols,*

$$\textit{if } \vdash \psi, \textit{ then } \models \psi.$$

The contrapositive of the soundness theorem is useful in showing that some derivations do not exist. We state it here.

Corollary 2.3.4. *If $\Gamma \not\models \psi$, then $\Gamma \nvdash \psi$.*

Example. The formula $((q \rightarrow p) \rightarrow q)$ is not derivable in $\mathcal{S}$ from the set $\{(p \rightarrow q)\}$.

Consider a truth assignment δ where $\delta[p] = 0$ and $\delta[q] = 0$. Then δ satisfies $(p \rightarrow q)$ but does not satisfy $((q \rightarrow p) \rightarrow q)$. Thus, $\{(p \rightarrow q)\}$ does not logically imply $((q \rightarrow p) \rightarrow q)$, that is, $\{(p \rightarrow q)\} \not\models ((q \rightarrow p) \rightarrow q)$. By the soundness theorem we get that $\{(p \rightarrow q)\} \nvdash ((q \rightarrow p) \rightarrow q)$. ♠

Definition.

- A proof system is *inconsistent* if there exists a formula θ such that within the system we have $\vdash \theta$ and $\vdash \neg\theta$.
- A proof system is *consistent* if it is not inconsistent.

In a similar fashion, we may talk about a set of formulas being inconsistent.

Definition.

- A set Γ of formulas is *inconsistent* if there exists a formula θ for which both $\Gamma \vdash \theta$ and $\Gamma \vdash \neg\theta$.
- A set Γ of formulas is *consistent* if it is not inconsistent.

Remark. A proof system is inconsistent if and only if the empty set is inconsistent.

Example.

- The set $\Gamma = \{p, \neg p\}$ is inconsistent because $\Gamma \vdash p$ and $\Gamma \vdash \neg p$.
- The set $\Delta = \{\neg(p \to q), \neg(q \to r)\}$ is inconsistent because $\Delta \vdash q$ and $\Delta \vdash \neg q$.

♠

Showing that a set of axioms is consistent is of major importance in mathematics, as otherwise, one can prove all formulas in an inconsistent system, meaning that no valuable information is gained by such formal frameworks. The soundness theorem assures us that our proof system $\mathcal{S}$ is consistent, a status that is much desirable.

Theorem 2.3.5. *The proof system $\mathcal{S}$ is consistent.*

Proof. Suppose for the sake of contradiction that $\mathcal{S}$ was inconsistent. Then there exists a formula θ such that $\vdash \theta$ and $\vdash \neg\theta$. By the soundness theorem, we get that $\models \theta$ and $\models \neg\theta$, meaning that both θ and $\neg\theta$ are tautologies, this cannot happen. So the system $\mathcal{S}$ is indeed consistent. ■

Lemma 2.3.6. *A set Γ of formulas is inconsistent if and only if $\Gamma \vdash \psi$ for every formula ψ.*

Proof. Suppose the Γ is inconsistent and let ψ be any arbitrary formula. So there exists a formula θ for which both $\Gamma \vdash \theta$ and $\Gamma \vdash \neg\theta$. We previously proved that $\{\theta, \neg\theta\} \vdash \psi$. Using these three derivations one can construct a derivation of the formula ψ from Γ. Therefore, $\Gamma \vdash \psi$.

For the reverse direction, suppose that $\Gamma \vdash \psi$ for every formula ψ. Then $\Gamma \vdash p$ and $\Gamma \vdash \neg p$. Therefore, Γ is inconsistent. ■

Corollary 2.3.7. *A set Γ of formulas is consistent if and only if $\Gamma \nvdash \varphi$ for some formula φ.*

Lemma 2.3.8. *Let Γ be a set of formulas and φ be a formula. Then*

$$\Gamma \cup \{\neg\varphi\} \textit{ is inconsistent if and only if } \Gamma \vdash \varphi.$$

Proof. Suppose that $\Gamma \cup \{\neg\varphi\}$ is inconsistent. Then there is a formula θ for which both $\Gamma \cup \{\neg\varphi\} \vdash \theta$ and $\Gamma \cup \{\neg\varphi\} \vdash \neg\theta$. By proof by contradiction, we obtain that $\Gamma \vdash \varphi$.

For the other direction, suppose that $\Gamma \vdash \varphi$. Then obviously, $\Gamma \cup \{\neg\varphi\} \vdash \varphi$ and $\Gamma \cup \{\neg\varphi\} \vdash \neg\varphi$. Therefore $\Gamma \cup \{\neg\varphi\}$ is inconsistent. ■

Corollary 2.3.9. *Let Γ be a set of formulas and φ be a formula. Then*

$$\Gamma \cup \{\neg\varphi\} \text{ is consistent if and only if } \Gamma \nvdash \varphi.$$

Lemma 2.3.10. *Let Γ be set of formulas. If Γ is consistent and $\Gamma \vdash \varphi$, then $\Gamma \cup \{\varphi\}$ is consistent as well.*

Proof. Suppose that Γ is consistent and $\Gamma \vdash \varphi$. For the sake of contradiction, suppose that $\Gamma \cup \{\varphi\}$ is inconsistent. Thus, $\Gamma \cup \{\varphi\} \vdash \theta$ and $\Gamma \cup \{\varphi\} \vdash \neg\theta$ for some formula θ. By the deduction theorem, we get that $\Gamma \vdash (\varphi \to \theta)$ and $\Gamma \vdash (\varphi \to \neg\theta)$. Since $\Gamma \vdash \varphi$ we obtain that $\Gamma \vdash \theta$ and $\Gamma \vdash \neg\theta$ meaning that Γ is inconsistent, a contradiction. Therefore, $\Gamma \cup \{\varphi\}$ must be consistent. ■

The following theorem shows that soundness of a formal system can be expressed in terms of consistency and satisfiability of sets of formulas.

Theorem 2.3.11. *The following statements are equivalent in a proof system.*

(I) For all sets of formulas Γ and all formulas ψ, if $\Gamma \vdash \psi$, then $\Gamma \models \psi$.

(II) For all sets of formulas Γ, if Γ is inconsistent, then Γ is not satisfiable.

Proof. (I) $\Rightarrow$ (II). Assume statement (I), that is, suppose that the system is sound. Assume that Γ is an inconsistent set of formulas. So there exists a formula θ for which both $\Gamma \vdash \theta$ and $\Gamma \vdash \neg\theta$. By soundness, we obtain that $\Gamma \models \theta$ and $\Gamma \models \neg\theta$. Now let δ be any truth assignment and, for the sake of contradiction, assume that δ satisfies Γ. By definition of logical consequence, we get that δ satisfies both θ and $\neg\theta$. But also as δ respects negation, we must have $\delta[\neg\theta] = 1 + \delta[\theta] = 1 + 1 = 0$, and so δ does not satisfy $\neg\theta$, a contradiction. Thus, no truth assignment satisfies Γ, and so Γ is not satisfiable.

(II) $\Rightarrow$ (I). Suppose statement (II) holds. We need to show that the proof system is sound. Let Γ be a set of formulas and ψ be a formula, and assume that $\Gamma \vdash \psi$. We need to show that $\Gamma \models \psi$. Let δ be any truth assignment and suppose that δ satisfies Γ. Since $\Gamma \vdash \psi$, we infer by Lemma 2.3.8 that $\Gamma \cup \{\neg\psi\}$ is inconsistent. It follows from (II) that $\Gamma \cup \{\neg\psi\}$ is not satisfiable, and so δ cannot satisfy $\neg\psi$ because if it does then it will satisfy all formulas in $\Gamma \cup \{\neg\psi\}$. So $\delta[\neg\psi] = 0$ implying that $\delta[\psi] = 1$, and thus δ satisfies ψ. So every truth assignment which satisfies Γ satisfies ψ, that is, $\Gamma \models \psi$. ■

As we have proved the soundness theorem for our proof system $\mathcal{S}$ we obtain the following result describing another connection between syntax and semantics of propositional logic.

Corollary 2.3.12. *For all sets of formulas* Γ*, if* Γ *is inconsistent in* $\mathcal{S}$*, then* Γ *is not satisfiable.*

We may also express the previous result via its contrapositive.

Corollary 2.3.13. *For all sets of formulas* Γ*, if* Γ *is satisfiable, then* Γ *is consistent.*

Example. The set $\{\neg(p \to q), \neg(r \to q)\}$ is consistent because it is satisfiable by any truth assignment δ for which $\delta[p] = 1$, $\delta[q] = 0$, and $\delta[r] = 1$. ♠

2.4 The Completeness Theorem

We have come to a point to establish the most important property of our formal system: the *completeness theorem.* It states that the proof system is powerful enough to derive all logical consequences. It is in no way obvious that the axioms and the deduction rule of the proof system $\mathcal{S}$ are sufficient to deal with all logical consequences.

Definition. A set Γ of propositional formulas is called *complete* if

(i) Γ is consistent and

(ii) for each formula φ, either $\Gamma \vdash \varphi$ or $\Gamma \vdash \neg\varphi$.

Lemma 2.4.1. *Suppose that Γ is a complete set of propositional formulas. Then for any formulas φ and ψ in $\mathcal{F}^*$ we have:*

(i) $\Gamma \vdash \neg\varphi$ *if and only if* $\Gamma \nvdash \varphi$.

(ii) $\Gamma \vdash (\varphi \to \psi)$ *if and only if* $\Gamma \vdash \neg\varphi$ *or* $\Gamma \vdash \psi$.

Proof. Let Γ be a complete set of formulas and let φ and ψ be formulas in $\mathcal{F}^*$.

(i) ($\Rightarrow$) Suppose that $\Gamma \vdash \neg\varphi$. Since Γ is consistent, $\Gamma \nvdash \varphi$.

($\Leftarrow$) Suppose $\Gamma \nvdash \varphi$. Since Γ is complete, $\Gamma \vdash \neg\varphi$.

(ii) ($\Rightarrow$) Suppose $\Gamma \vdash (\varphi \to \psi)$. If $\Gamma \vdash \neg\varphi$, we are done. Otherwise, suppose that $\Gamma \nvdash \neg\varphi$. As Γ is complete we must have $\Gamma \vdash \varphi$. Using the derivations of $\Gamma \vdash (\varphi \to \psi)$ and $\Gamma \vdash \varphi$, and an application of Modus Ponens we may construct a derivation for $\Gamma \vdash \psi$.

($\Leftarrow$) Case (i) Suppose $\Gamma \vdash \neg\varphi$. We proved earlier that $\vdash (\neg\varphi \to (\varphi \to \psi))$. Thus, using these two derivations together with an application of Modus Ponens we construct a derivation of $\Gamma \vdash (\varphi \to \psi)$.
Case (ii) Suppose that $\Gamma \vdash \psi$. Using this derivation together with an instance $(\psi \to (\varphi \to \psi))$ of Axiom 1, and an application of Modus Ponens we get that $\Gamma \vdash (\varphi \to \psi)$.

■

Theorem 2.4.2. *If Γ is a complete set of propositional formulas, then Γ is satisfiable.*

Proof. Let Γ be a complete set of formulas. Let the set of propositional variables be $\mathbf{P} = \{p_0, p_1, p_2, \ldots\}$. As each propositional variable is a formula and as Γ is complete we get that either $\Gamma \vdash p_i$ or $\Gamma \vdash \neg p_i$, but not both, for every $i \in \mathbb{N}$. Let $\delta : \mathcal{F} \to \{0, 1\}$ be the truth assignment determined by setting:

$$\delta[p_i] = \begin{cases} 1 & \text{if } \Gamma \vdash p_i; \\ 0 & \text{if } \Gamma \vdash \neg p_i. \end{cases}$$

Claim. We claim that for any formula $\varphi \in \mathcal{F}^*$ we have:

$$\delta[\varphi] = 1 \text{ if and only if } \Gamma \vdash \varphi.$$

We prove the claim by induction on formulas. By definition of the truth assignment δ we have that $\delta[p] = 1$ if and only if $\Gamma \vdash p$ for every propositional variable $p \in \mathbf{P}$.

Now let φ, ψ be formulas in $\mathcal{F}^*$ for which the claim is satisfied. We have to show that the claim holds for $\neg\varphi$ and $(\varphi \to \psi)$. For the first,

$$\begin{aligned} \delta[\neg\varphi] = 1 &\Leftrightarrow \delta[\varphi] = 0 \\ &\Leftrightarrow \Gamma \nvdash \varphi && \text{(Induction hypothesis)} \\ &\Leftrightarrow \Gamma \vdash \neg\varphi. && \text{(Since } \Gamma \text{ is complete)} \end{aligned}$$

It remains to show that the claim holds for the formula $(\varphi \to \psi)$.

$$\begin{aligned} \delta[(\varphi \to \psi)] = 1 &\Leftrightarrow \delta[\varphi] = 0 \text{ or } \delta[\psi] = 1 \\ &\Leftrightarrow \Gamma \nvdash \varphi \text{ or } \Gamma \vdash \psi && \text{(Induction hypothesis)} \\ &\Leftrightarrow \Gamma \vdash \neg\varphi \text{ or } \Gamma \vdash \psi && \text{(Since } \Gamma \text{ is complete)} \\ &\Leftrightarrow \Gamma \vdash (\varphi \to \psi). && \text{(By Lemma 2.4.1(ii)).} \end{aligned}$$

This establishes the claim. Now, choose any formula $\gamma \in \Gamma$. Obviously, $\Gamma \vdash \gamma$. By the claim $\delta[\gamma] = 1$. So δ satisfies every formula in Γ and so Γ is satisfiable. ■

Theorem 2.4.3. *Suppose that Γ is a consistent set of formulas. Then there exists a complete set of formulas $\tilde{\Gamma}$ such that $\Gamma \subseteq \tilde{\Gamma}$.*

Proof. Start with any consistent set of formulas Γ. Since $\mathcal{F}^*$ is a countably infinite set, we can enumerate all its formulas as

$$\mathcal{F}^* = \{\varphi_1, \varphi_2, \varphi_3, \ldots\}.$$

We proceed by constructing a chain of consistent sets of formulas

$$\Gamma_0 \subseteq \Gamma_1 \subseteq \Gamma_2 \subseteq \cdots \subseteq \Gamma_n \subseteq \Gamma_{n+1} \subseteq \cdots$$

as follows. We start by taking $\Gamma_0 = \Gamma$, and so Γ_0 is consistent. Then suppose the consistent set Γ_n has been constructed, we construct a new consistent set Γ_{n+1} as follows.

$$\Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{\varphi_{n+1}\} & \text{if } \Gamma_n \vdash \varphi_{n+1}\,; \\ \Gamma_n \cup \{\neg\varphi_{n+1}\} & \text{if } \Gamma_n \nvdash \varphi_{n+1}\,. \end{cases}$$

Observe that if $\Gamma_n \vdash \varphi_{n+1}$, then $\Gamma_n \cup \{\varphi_{n+1}\}$ is consistent by Lemma 2.3.10, and if $\Gamma_n \nvdash \varphi_{n+1}$, then $\Gamma_n \cup \{\neg\varphi_{n+1}\}$ is consistent by Corollary 2.3.9. Thus, Γ_{n+1} is consistent. Furthermore, by construction, we have that either $\varphi_n \in \Gamma_n$ or $\neg\varphi_n \in \Gamma_n$ for every positive integer n. Finally, let

$$\tilde{\Gamma} = \bigcup_{n \in \mathbb{N}} \Gamma_n.$$

Clearly, $\Gamma = \Gamma_0 \subseteq \tilde{\Gamma}$. We claim that $\tilde{\Gamma}$ is a complete set of formulas.

Suppose for the sake of contradiction that $\tilde{\Gamma}$ is inconsistent. Then there exists a formulas θ such that $\tilde{\Gamma} \vdash \theta$ and $\tilde{\Gamma} \vdash \neg\theta$. Since derivations are finite, these two derivations used finitely many formulas (assumptions) from $\tilde{\Gamma}$. So there is some m large enough such that all these assumptions are members of Γ_m, and consequently, we obtain that $\Gamma_m \vdash \theta$ and $\Gamma_m \vdash \neg\theta$. This shows that Γ_m is inconsistent, contradicting that every Γ_n is consistent by construction. Therefore, $\tilde{\Gamma}$ must be consistent.

It remains to show that $\tilde{\Gamma}$ proves any formula or its negation. Let ψ be any formula in $\mathcal{F}^*$. Then ψ appears in the enumeration above, say $\psi = \varphi_n$ for some n. At stage n of construction, we ensured that either $\varphi_n \in \Gamma_n$ or $\neg\varphi_n \in \Gamma_n$. It follows that $\varphi_n \in \tilde{\Gamma}$ or $\neg\varphi_n \in \tilde{\Gamma}$. Thus obviously we obtain that $\tilde{\Gamma} \vdash \varphi_n$ or $\tilde{\Gamma} \vdash \neg\varphi_n$. Therefore, $\tilde{\Gamma} \vdash \psi$ or $\tilde{\Gamma} \vdash \neg\psi$ as desired.

We have shown that $\tilde{\Gamma}$ is complete and contains Γ as a subset. ■

Corollary 2.4.4. *If Γ is a consistent set of propositional formulas, then Γ is satisfiable.*

Proof. Let Γ be a consistent set of formulas. By Theorem 2.4.3, there exists a complete set of formulas $\tilde{\Gamma}$ such that $\Gamma \subseteq \tilde{\Gamma}$. By Theorem 2.4.2, we get that $\tilde{\Gamma}$ is satisfiable. Thus, there exists a truth assignment δ which satisfies every formula $\tilde{\Gamma}$. Since $\Gamma \subseteq \tilde{\Gamma}$ the same truth assignment δ satisfies every formula in Γ. So Γ is satisfiable. ■

We now have all the tools to prove the completeness theorem for the proof system $\mathcal{S}$. The completeness theorem says that the proof system $\mathcal{S}$ with its three schemes of axioms and its Modus Ponens deduction rule is complete, that is, it is rich enough to produce a formal derivation for any logical consequence in propositional logic.

Theorem 2.4.5 (Completeness Theorem for $\mathcal{S}$). *Let Γ be any set of propositional formulas and let ψ be any propositional formula. Then if Γ logically implies ψ, then Γ proves ψ. In symbols,*

$$\textit{if } \Gamma \models \psi, \textit{ then } \Gamma \vdash \psi.$$

Proof. We will show the contrapositive of the theorem. Suppose that $\Gamma \nvdash \psi$. It follows by Corollary 2.3.9 that $\Gamma \cup \{\neg\psi\}$ is consistent. By Corollary 2.4.4, we get that $\Gamma \cup \{\neg\psi\}$ is satisfiable. Thus, there exists a truth assignment δ that satisfies every formula in $\Gamma \cup \{\neg\psi\}$. This δ satisfies the set Γ and satisfies the formula $\neg\psi$, and so δ satisfies Γ and does not satisfy ψ. Therefore, $\Gamma \not\models \psi$ as desired. ■

Corollary 2.4.6 (Soundness and Completeness of $\mathcal{S}$)**.** *Let Γ be any set of propositional formulas and ψ be a propositional formula.*

$$\Gamma \vdash \psi \textit{ if and only if } \Gamma \models \psi.$$

Corollary 2.4.7. *Let Γ be any set of propositional formulas.*

$$\Gamma \textit{ is satisfiable if and only if } \Gamma \textit{ is consistent.}$$

Next we investigate an interesting characteristic of sets of formulas and its relation with complete sets.

Definition. A set Σ of formulas is called *maximal consistent* if

(i) Σ is consistent and

(ii) for any consistent set Δ of formulas, if $\Sigma \subseteq \Delta$, then $\Sigma = \Delta$.

Observe that Σ being maximal consistent means that there is no consistent set of formulas which properly contains Σ. That is, every proper superset of Σ is inconsistent. We now show that maximal consistent sets are closed under derivability.

Lemma 2.4.8. *If Σ is maximal consistent, then Σ contains all the formulas derivable from Σ, that is, if $\Sigma \vdash \varphi$, then $\varphi \in \Sigma$ for any formula φ.*

Proof. Let Σ be a maximal consistent set of formulas and let φ be a formula. Suppose that $\Sigma \vdash \varphi$. Since Σ is consistent, by Lemma 2.3.10, we get that $\Sigma \cup \{\varphi\}$ is consistent as well. Clearly, $\Sigma \subseteq \Sigma \cup \{\varphi\}$, and so by maximality of Σ, we get that $\Sigma = \Sigma \cup \{\varphi\}$. Thus, $\varphi \in \Sigma$. ■

The next two theorems show that complete sets and maximal consistent sets are essentially the same notions.

Theorem 2.4.9. *Let Γ be a complete set of formulas. Then the set*

$$\Sigma = \{\varphi \in \mathcal{F}^* \mid \Gamma \vdash \varphi\}$$

is maximal consistent.

Proof. Suppose that Γ is a complete set of formulas. We have to show that the set $\Sigma = \{\varphi \in \mathcal{F}^* \mid \Gamma \vdash \varphi\}$ is maximal consistent. By definition of Σ one can show that $\Gamma \vdash \varphi$ if and only if $\Sigma \vdash \varphi$ for any formula φ. In other words, Γ and Σ prove exactly the same set of formulas. Consequently, as Γ is consistent it follows that Σ

is consistent as well. It remains to show that Σ is maximal. Let $\Sigma \subseteq \Delta$ where Δ is a consistent set of formulas. Choose any formula $\varphi \in \Delta$. Since Γ is complete, $\Gamma \vdash \varphi$ or $\Gamma \vdash \neg\varphi$. If $\Gamma \vdash \neg\varphi$, it follows that $\neg\varphi \in \Sigma$, and thus both $\varphi, \neg\varphi \in \Delta$, contradicting that Δ is consistent. Therefore, it must be the case that $\Gamma \vdash \varphi$, and in this case we get that $\varphi \in \Sigma$. Thus, $\Delta \subseteq \Sigma$. So $\Sigma = \Delta$ establishing that Σ is maximal consistent. ■

Theorem 2.4.10. *If Σ is a maximal consistent set of formulas, then Σ is complete.*

Proof. Let Σ be a maximal consistent set of formulas. By definition, Σ is consistent. It remains to show that Σ proves any formula or its negation. Let φ be any formula. If $\Sigma \vdash \varphi$, we are done. Otherwise, $\Sigma \nvdash \varphi$ and so, by Corollary 2.3.9, it follows that $\Sigma \cup \{\neg\varphi\}$ is consistent. Since $\Sigma \subseteq \Sigma \cup \{\neg\varphi\}$ and Σ is maximal consistent, we get that $\Sigma = \Sigma \cup \{\neg\varphi\}$, and so $\neg\varphi \in \Sigma$. Thus, $\Sigma \vdash \neg\varphi$. ■

By Theorem 2.4.10 and Lemma 2.4.8 we obtain the following characterization of maximal consistent sets.

Corollary 2.4.11. *A set Σ of formulas is maximal consistent if and only if*

(i) Σ is consistent and

(ii) for each formula φ, either $\varphi \in \Sigma$ or $\neg\varphi \in \Sigma$.

We can now give plenty of examples of complete sets.

Example. Choose any truth assignment $\delta : \mathbf{P} \to \{0, 1\}$. Consider the set

$$\Sigma_\delta = \{\varphi \in \mathcal{F}^* \mid \delta[\varphi] = 1\}.$$

Then Σ_δ is satisfiable by δ, and so consistent by Corollary 2.3.13. Moreover, by definition of Σ_δ, exactly one of $\varphi \in \Sigma_\delta$ or $\neg\varphi \in \Sigma_\delta$ holds for every formula φ in $\mathcal{F}^*$. Therefore, Σ_δ is maximal consistent and is complete as well. ♠

Definition. We say that two proof systems $\mathcal{S}_1$ and $\mathcal{S}_2$ are *equivalent* if for any set of formulas Γ and any formula ψ we have that

$$\Gamma \vdash_{\mathcal{S}_1} \psi \text{ if and only if } \Gamma \vdash_{\mathcal{S}_2} \psi.$$

Theorem 2.4.12. *Suppose that $\mathcal{S}_1$ and $\mathcal{S}_2$ are proof systems for propositional logic which are both sound and complete. Then $\mathcal{S}_1$ and $\mathcal{S}_2$ are equivalent.*

Proof. Let Γ be a set of formulas and let ψ be a formula. Then

$$\begin{aligned} \Gamma \vdash_{\mathcal{S}_1} \psi &\xRightarrow[\mathcal{S}_1]{\text{Soundness}} \Gamma \models \psi \\ &\xRightarrow[\mathcal{S}_2]{\text{Completeness}} \Gamma \vdash_{\mathcal{S}_2} \psi. \end{aligned}$$

Thus, if $\Gamma \vdash_{\mathcal{S}_1} \psi$, then $\Gamma \vdash_{\mathcal{S}_2} \psi$. Similarly,

$$\begin{aligned} \Gamma \vdash_{\mathcal{S}_2} \psi \ & \xRightarrow[\mathcal{S}_2]{\text{Soundness}} \Gamma \models \psi \\ & \xRightarrow[\mathcal{S}_1]{\text{Completeness}} \Gamma \vdash_{\mathcal{S}_1} \psi. \end{aligned}$$

Therefore, $\mathcal{S}_1$ and $\mathcal{S}_2$ are equivalent. ■

Corollary 2.4.13. *Any sound and complete proof system of propositional logic is equivalent to our proof system $\mathcal{S}$ introduced in Section 2.1.*

2.5 The Compactness Theorem

Let Γ be a (possibly infinite) set of formulas and let ψ be a formula. Furthermore, suppose that $\Gamma \models \psi$. By the completeness theorem, we obtain $\Gamma \vdash \psi$. By definition of a derivation being a finite sequence of formulas, there is a finite subset $\Delta \subseteq \Gamma$ such that $\Delta \vdash \psi$. By the soundness theorem, we get that $\Delta \models \psi$. In summary, if $\Gamma \models \psi$, then $\Delta \models \psi$ for some finite subset $\Delta \subseteq \Gamma$.

Definition. A set Γ of formulas is called *finitely satisfiable* if every finite subset of Γ is satisfiable.

Theorem 2.5.1 (Compactness Theorem)**.** *A set of propositional formulas Γ is satisfiable if and only if Γ is finitely satisfiable.*

Proof. The forward direction is obvious.

For the converse, suppose that a set of formulas Γ is finitely satisfiable. For the sake of contradiction, assume that Γ is not satisfiable. By Corollary 2.4.4 we know that Γ is inconsistent. Thus there exists a formula θ such that $\Gamma \vdash \theta$ and $\Gamma \vdash \neg\theta$. As derivations use finitely many formulas, these two derivations have used finitely many assumptions from Γ. Let $\Delta \subseteq \Gamma$ be the set of the finitely many assumptions from Γ used in both of these derivations. Thus, $\Delta \vdash \theta$ and $\Delta \vdash \neg\theta$. It follows that Δ is a finite inconsistent set of Γ. By Corollary 2.3.12, we know that Δ is not satisfiable. So we get that Δ is a finite set of Γ which is not satisfiable, contradicting that Γ is finitely satisfiable. Therefore, Γ must be satisfiable. ■

Corollary 2.5.2. *If Γ is not satisfiable, then there exists a finite subset $\Delta \subseteq \Gamma$ which is not satisfiable.*

Example. Suppose that Γ is an infinite set of formulas such that every truth assignment satisfies at least one formula in Γ. That is, for any truth assignment δ, there is $\varphi \in \Gamma$ such that $\delta[\varphi] = 1$. Show that Γ has a finite subset with the same property.

Consider the set $\Sigma = \{\neg\varphi \mid \varphi \in \Gamma\}$. By definition of Σ and the property of Γ, any truth assignment does not satisfy at least one formula in Σ. So Σ is not satisfiable. By Compactness Theorem, there is a finite subset $\Delta \subseteq \Sigma$ which is not satisfiable. Then the set $\Lambda = \{\varphi \in \Gamma \mid \neg\varphi \in \Delta\}$ is a finite subset of Γ where every truth assignment satisfies at least one formula in Λ. ♠

Chapter 3

First-Order Logic

In everyday mathematics we study mathematical objects together with the properties they satisfy or they do not satisfy. *First-order logic* (FOL), also called *predicate logic*, is one attempt towards the formalization of this mathematical activity. The syntax of first-order logic consists of certain strings of symbols called *terms* and *formulas*. Terms name elements of the mathematical object of interest, and formulas describe properties of these elements. Terms and formulas are strings built from an alphabet of symbols that includes propositional connectives and parentheses, as in propositional logic, together with the universal quantifier $\forall$, the existential quantifier $\exists$, and an infinite stock of variables $x_0, x_1, x_2, \ldots$; these make the innovation of first-order logic. The alphabet has also more symbols from the *first-order language* that is appropriate to describe the mathematical structure under consideration. The semantics of first-order logic defines the satisfaction of a formula in a structure.

The name "first-order" refers to the fact that quantifiers range over elements of the structure under study. For instance when we say that every real number has an additive inverse, we quantify over the members of the set $\mathbb{R}$. On the other hand, there are mathematical properties that cannot be expressed in first-order logic. For example, to say that the field of real numbers is complete we say that "every nonempty subset of $\mathbb{R}$ which is bounded above has a least upper bound (supremum)". Here the universal quantifier ranges over subsets of $\mathbb{R}$ and not over elements of $\mathbb{R}$. This is allowed in *second-order logic*.

3.1 First-Order Structures

3.1.1 Languages and Structures

We aim to develop a logic capable of describing mathematical objects such as the following.

1. The ordered field of real numbers $(\mathbb{R},\, 0, 1, +, \cdot, <)$.
2. The field of complex numbers $(\mathbb{C},\, 0, 1, +, \cdot)$.
3. The field $(\mathbb{Z}_7,\, 0, 1, +, \cdot)$ of integers modulo 7.
4. The exponential field of real numbers $(\mathbb{R},\, 0, 1, +, \cdot, \exp)$.
5. The structure $(A,\, \sim)$ where A is a set and $\sim$ is an equivalence relation on A.
6. The complete graph $(K_n,\, E)$ where E is the edge relation.
7. The discrete linear order of the integers $(\mathbb{Z},\, <)$.
8. The dense linear order of the rationals $(\mathbb{Q},\, <)$.
9. The additive group of integers $(\mathbb{Z},\, 0, +)$.
10. The arithmetic of natural numbers $(\mathbb{N},\, 0, 1, +, \cdot)$.
11. The boolean algebra $(\mathcal{P}(A),\, \emptyset, A, \cap, \cup, {}^c)$ for any set A.
12. The Lindenbaum-Tarski algebra $(\mathcal{F}/\equiv,\, \bot, \top, \wedge, \vee, \neg)$ where $\mathcal{F}/\equiv$ is the set of equivalence classes given by the logical equivalence relation on the set $\mathcal{F}$ of propositional formulas.

All of the mathematical structures above consist of a set on which some functions and relations are defined along with some distinguished elements. For example, the field of real numbers consists of the set $\mathbb{R}$ with 0 and 1 as distinguished elements and it has addition $+$ and multiplication $\cdot$ as functions from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$, and has the usual order $\leq$ of real numbers as a binary relation on $\mathbb{R}$, that is, a subset of $\mathbb{R} \times \mathbb{R}$.

Definition. A *first-order language* $\mathcal{L}$ is a set of symbols of three kinds: constant symbols c_i, function symbols f_j, and relation symbols R_k, more precisely,

$$\mathcal{L} = \{c_i \mid i \in I\} \cup \{f_j \mid j \in J\} \cup \{R_k \mid k \in K\}$$

for some indexing sets I, J, K. Moreover, every function symbol and every relation symbol is associated with a positive integer called its *arity.*

A first-order language is also called a *signature* or a *vocabulary.* When the arity is 1 we say *unary*, when the arity is 2 we say *binary*, and when the arity is 3 we say *ternary.* We follow the convention that first-order languages contain the equality symbol "=" which is a binary relation symbol. A first-order language which contains no constant symbols and no function symbols is called a *relational language.*

When the symbols of a first-order language are interpreted as their intended meaning the result will be a *first-order structure.* All of the mathematical objects given at the beginning of this section are first-order structures. In general, we choose a first-order language appropriate for the type of structure under consideration. We adapt our choice of constant symbols, function symbols, and relation symbols in $\mathcal{L}$ to suit the structure being dealt with.

Definition. Let $\mathcal{L}$ be a first-order language. An $\mathcal{L}$-structure $\mathcal{M}$ consists of

- a nonempty set M, called the *domain* or the *underlying set* of the structure $\mathcal{M}$;
- for each constant symbol $c \in \mathcal{L}$, a distinguished element $c^{\mathcal{M}}$ in M called the *interpretation* of the symbol c in the structure $\mathcal{M}$;
- for each function symbol $f \in \mathcal{L}$ of arity n, a function $f^{\mathcal{M}} : M^n \to M$ called the interpretation of the symbol f in the structure $\mathcal{M}$;
- for each relation symbol $R \in \mathcal{L}$ of arity n, a subset $R^{\mathcal{M}} \subseteq M^n$ called the interpretation of the symbol R in the structure $\mathcal{M}$.

Recall that a *relation* of arity n on a set M is a subset of the cartesian product M^n and a *predicate* is a symbol representing a relation.

Remark. The interpretation of the equality symbol in a structure $\mathcal{M}$ is the equality relation on its underlying set M. That is, $=^{\mathcal{M}}$ is a subset of M^2 and it is the set

$$\{(a, a) \mid a \in M\}.$$

Usually, we first choose a first order language $\mathcal{L}$. Then to describe an $\mathcal{L}$-structure we give its underlying set, and then provide the interpretations of the various symbols of the language preferably in the same order these symbols were presented. It is common to use calligraphic letters (usually $\mathcal{M}$ or $\mathcal{N}$) to denote a $\mathcal{L}$-structures and use their corresponding Latin letters (M or N) to denote the underlying set of the structure. More generally, for a first-order language

$$\mathcal{L} = \{c_i \mid i \in I\} \cup \{f_j \mid j \in J\} \cup \{R_k \mid k \in K\}$$

we present an $\mathcal{L}$-structure by writing

$$\mathcal{M} = (M, c_i^{\mathcal{M}}, f_j^{\mathcal{M}}, R_k^{\mathcal{M}})_{i\in I,\, j\in J,\, k\in K}.$$

Example. Consider the first-order language $\mathcal{L} = \{a, c, f, g, h, R\}$ where a, c are constant symbols, f, g are binary function symbols, h is a unary function symbol, and R is a binary relation symbol. An example of an $\mathcal{L}$-structure is the exponential field of real numbers

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

From this we understand that the underlying set of the $\mathcal{L}$-structure $\mathcal{M}$ is the set $\mathbb{R}$ of real numbers, and in which the interpretations of the symbols in $\mathcal{L}$ are as follows:

- Interpret a as the real number 0, that is, $a^{\mathcal{M}}$ is 0.
- Interpret c as the real number 1, that is, $c^{\mathcal{M}}$ is 1.
- Interpret f as the addition function $f^{\mathcal{M}} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ where $f^{\mathcal{M}}(x, y) = x + y$.
- Interpret g as the multiplication function $g^{\mathcal{M}} : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ where $g^{\mathcal{M}}(x, y) = x{\cdot}y$.
- Interpret h as the exponential function $h^{\mathcal{M}} : \mathbb{R} \to \mathbb{R}$ where $h^{\mathcal{M}}(x) = e^x$.
- Interpret R as the usual order relation on real numbers, so $R^{\mathcal{M}} = \{(x, y) \mid x \leq y\}$.

♠

Example. Consider the first-order language $\mathcal{L} = \{a, c, f, g, h, L\}$ as in the previous example and let us introduce another $\mathcal{L}$-structure. Let $A = \{1, 2, 3\}$ and consider the $\mathcal{L}$-structure given by

$$\mathcal{N} = (\mathcal{P}(A),\, \emptyset, A, \cap, \cup,\, ^c, \subseteq).$$

The underlying set of the structure $\mathcal{N}$ is the set

$$N = \mathcal{P}(A) = \big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\big\},$$

and the interpretations in the structure $\mathcal{N}$ of the symbols of $\mathcal{L}$ are as follows.

- $a^{\mathcal{N}}$ is $\emptyset$.
- $c^{\mathcal{N}}$ is $\{1, 2, 3\}$.
- $f^{\mathcal{N}} : N \times N \to N$ where $f^{\mathcal{N}}(x, y) = x \cap y$.
- $g^{\mathcal{N}} : N \times N \to N$ where $g^{\mathcal{N}}(x, y) = x \cup y$.
- $h^{\mathcal{N}} : N \to N$ where $h^{\mathcal{N}}(x) = x^c$ where x^c is the complement of the subset x.
- Interpret R as the subset relation, so $R^{\mathcal{N}} = \{(x, y) \mid x \subseteq y\}$.

♠

Remark. It is important to make a clear distinction between a symbol s of a first-order language $\mathcal{L}$ and its interpretation $s^{\mathcal{M}}$ in an $\mathcal{L}$-structure $\mathcal{M}$. As we have seen above, the same symbol has different interpretations in various $\mathcal{L}$-structures.

Exercise. Build a first-order language $\mathcal{L}$ of your choice and present various $\mathcal{L}$-structures which need not to be known structures!

3.1.2 Substructures

Fix a first-order language $\mathcal{L}$. Let $\mathcal{M}$ be an $\mathcal{L}$-structure with domain M and let $\mathcal{N}$ be an $\mathcal{L}$-structure with domain N. We say that $\mathcal{M}$ is a substructure $\mathcal{N}$ when $M \subseteq N$ and the interpretations of the symbols of $\mathcal{L}$ in the structure $\mathcal{M}$ are the restrictions of their interpretations in $\mathcal{N}$ to the subset M. More precisely, we present the following definition.

Definition. Let $\mathcal{L}$ be a first-order language, and let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. We say that $\mathcal{M}$ is a *substructure* of $\mathcal{N}$, if and only if the following conditions are satisfied:

- $M \subseteq N$ where M is the domain of $\mathcal{M}$ and N is the domain $\mathcal{N}$;
- for each constant symbol $c \in \mathcal{L}$, we have that $c^{\mathcal{M}} = c^{\mathcal{N}}$;
- for each function symbol $f \in \mathcal{L}$ of arity n, we have that $f^{\mathcal{M}}(\bar{a}) = f^{\mathcal{N}}(\bar{a})$ for every n-tuple $\bar{a} \in M^n$, that is, $f^{\mathcal{M}} = f^{\mathcal{N}}\restriction_{M^n}$;
- for each relation symbol $R \in \mathcal{L}$ of arity n, we have that

  $$R^{\mathcal{M}} = R^{\mathcal{N}} \cap M^n,$$

  that is, for every n-tuple $\bar{a} \in M^n$ we have $\bar{a} \in R^{\mathcal{M}}$ if and only if $\bar{a} \in R^{\mathcal{N}}$.

Remark. When $\mathcal{M}$ is a substructure of $\mathcal{N}$ we write $\mathcal{M} \subseteq \mathcal{N}$, and we also say that $\mathcal{N}$ is an *extension* of $\mathcal{M}$.

Let us state some consequences of the definition above. Suppose that $\mathcal{M}$ is a substructure of $\mathcal{N}$. If c is a constant symbol, then $c^{\mathcal{N}} \in M$ since $c^{\mathcal{M}} = c^{\mathcal{N}}$. If f is a function symbol, then the restriction of the function $f^{\mathcal{N}}$ to the subset M^n must be the function $f^{\mathcal{M}} : M^n \to M$. This means that the subset M of the set N is closed under the function $f^{\mathcal{N}}$. From this we conclude that for a subset $A \subseteq N$ to be the underlying set of a substructure of $\mathcal{N}$ it must include all the interpretations in $\mathcal{N}$ of constant symbols of $\mathcal{L}$ and is closed under all functions of the structure $\mathcal{N}$.

Example. Consider the language $\mathcal{L} = \{c, f, R\}$ where c is a constant symbol, f is a unary function symbol, and R is a binary relation symbol. Now consider the following $\mathcal{L}$-structures.

- $\mathcal{M} = (M, c^{\mathcal{M}}, f^{\mathcal{M}}, R^{\mathcal{M}})$ where $M = \{1, 2, 3\}$ and $c^{\mathcal{M}} = 2$ and the function $f^{\mathcal{M}}$ is given by $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 1$, and the relation is

  $$R^{\mathcal{M}} = \{(1, 1), (1, 2), (2, 1), (3, 2)\}.$$

- $\mathcal{N} = (N, c^{\mathcal{N}}, f^{\mathcal{N}}, R^{\mathcal{N}})$ where $N = \{0,1,2,3,4,5\}$ and $c^{\mathcal{N}} = 2$ and the function $f^{\mathcal{N}}$ is given by $0 \mapsto 1$, $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 1$, $4 \mapsto 2$, $5 \mapsto 3$, and the relation is
$$R^{\mathcal{N}} = \{(0,1),(1,1),(1,4),(1,2),(2,1),(3,2),(4,5),(5,3)\}.$$
- $\mathcal{A} = (A, c^{\mathcal{A}}, f^{\mathcal{A}}, R^{\mathcal{A}})$ where $A = \{0,1,2,3,4,5\}$ and $c^{\mathcal{A}} = 2$ and the function $f^{\mathcal{A}}$ is given by $0 \mapsto 1$, $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 1$, $4 \mapsto 5$, $5 \mapsto 4$, and the relation is
$$R^{\mathcal{A}} = \{(0,2),(1,1),(1,2),(2,1),(1,3),(3,2),(4,5),(5,3)\}.$$

Then $\mathcal{M}$ is a substructure of $\mathcal{N}$, and $\mathcal{M}$ is not a substructure of $\mathcal{A}$. Moreover, neither of $\mathcal{N}$ and $\mathcal{A}$ is a substructure of the other. ♠

Example. Consider the language $\mathcal{L} = \{a, c, f, g, R\}$ where a, c are constant symbols, f, g are binary function symbols, and R is a binary relation symbol. All of the structures below are $\mathcal{L}$-structures.

- $(\mathbb{Q}, 0, 1, +, \cdot, <)$ is a substructure of $(\mathbb{R}, 0, 1, +, \cdot, <)$.
- $(\mathbb{Q}, 0, 1, +, \cdot, <)$ is not a substructure of $(\mathbb{R}, 0, 3, +, \cdot, <)$.
- $(\mathbb{Z}, 0, 1, +, \cdot, <)$ is a substructure of $(\mathbb{Q}, 0, 1, +, \cdot, <)$.
- $(\mathbb{Z}, 0, 1, +, \cdot, \mid)$ where $\mid$ is the divisibility relation is not a substructure of $(\mathbb{Q}, 0, 1, +, \cdot, <)$.
- $(\mathbb{Z}_7, 0, 1, +, \cdot, <)$ where $+, \cdot$ are addition and multiplication modulo 7 respectively is not a substructure of $(\mathbb{Z}_9, 0, 1, \oplus, \star, <)$ where $\oplus, \star$ are addition and multiplication modulo 9 respectively.

♠

Example (Graph theory)**.** Recall that a *graph* is a set equipped with an irreflexive symmetric binary relation. Consider a first-order language $\mathcal{L} = \{E\}$ where E is a binary relation symbol. Suppose that G and H are graphs, then we can view them as $\mathcal{L}$-structures by interpreting E as the edge relation in each one of them. Then G is a substructure of H if and only if G is an *induced subgraph* of H.

3.1.3 Isomorphisms

An isomorphism between first-order structures over the same language is a bijective map which respects the constants, functions, and relations of these these structures.

Definition. Let $\mathcal{L}$ be a first-order language, and let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures with underlying sets M and N, respectively. An *isomorphism* from $\mathcal{M}$ to $\mathcal{N}$ is a bijective map $h : M \to N$ such that the following conditions are satisfied:

- for every constant symbol $c \in \mathcal{L}$, we have that
$$h(c^{\mathcal{M}}) = c^{\mathcal{N}};$$
- for every function symbol $f \in \mathcal{L}$ of arity n, and for every n-tuple $(a_1, a_2, \ldots, a_n)$ in M^n we have that
$$h(f^{\mathcal{M}}(a_1, a_2, \ldots, a_n)) = f^{\mathcal{N}}(h(a_1), h(a_2), \ldots, h(a_n));$$
- for every relation symbol $R \in \mathcal{L}$ of arity n, and for every n-tuple $(a_1, a_2, \ldots, a_n)$ in M^n we have that
$$(a_1, a_2, \ldots, a_n) \in R^{\mathcal{M}} \quad \text{if and only if} \quad (h(a_1), h(a_2), \ldots, h(a_n)) \in R^{\mathcal{N}}.$$

When there is an isomorphism from an $\mathcal{L}$-structure $\mathcal{M}$ to an $\mathcal{L}$-structure $\mathcal{N}$ we say that the structures are *isomorphic* and we write $\mathcal{M} \cong \mathcal{N}$.

Definition. An *automorphism* of a structure $\mathcal{M}$ is an isomorphism from $\mathcal{M}$ to $\mathcal{M}$.

Remark. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Then the set $\mathrm{Aut}(\mathcal{M})$ of all automorphisms of $\mathcal{M}$ is a group under function composition.

Definition. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. An *embedding* of $\mathcal{M}$ in $\mathcal{N}$ is an isomorphism from $\mathcal{M}$ to a substructure of $\mathcal{N}$.

Example. Consider the first-order language $\mathcal{L} = \{c, f\}$ where c is a constant symbol and f is a binary function symbol. Then
$$(\mathbb{R}, 0, +) \cong (\mathbb{R}^+, 1, \cdot)$$
where $\mathbb{R}^+$ is the set of positive reals. The map $x \mapsto e^x$ is one isomorphism.

Example. Consider the first-order language $\mathcal{L} = \{E\}$ where E is a binary relation symbol. Consider the structure $\mathcal{A} = (A, E^{\mathcal{A}})$ where $A = \{0, 1, 2, 3, 4, 6\}$ and $E^{\mathcal{A}} = \{(x, y) \in A^2 \mid x \equiv y \pmod 3\}$ and the structure $\mathcal{B} = (B, E^{\mathcal{B}})$ where $B = \{3, 4, 6, 8, 11, 12\}$ and $E^{\mathcal{B}} = \{(x, y) \in B^2 \mid x \equiv y \pmod 4\}$. Then $\mathcal{A} \cong \mathcal{B}$. A map witnessing this isomorphism is the map $h : A \to B$ given by $0 \mapsto 4$, $1 \mapsto 3$, $2 \mapsto 6$, $3 \mapsto 8$, $4 \mapsto 11$, and $6 \mapsto 12$. Check that for any $x, y \in A$ we have that
$$(x, y) \in E^{\mathcal{A}} \ \text{ if and only if } \ (h(x), h(y)) \in E^{\mathcal{B}}.$$

3.2 Syntax of First-Order Logic

We first describe the alphabet from which we will build the terms and formulas in first-order logic. The alphabet $\mathcal{A}_{\mathcal{L}}$ of a first-order language $\mathcal{L}$ is the set of the following symbols:

- The propositional connectives: negation $\neg$, conjunction $\wedge$, disjunction $\vee$, implication $\rightarrow$, and equivalence $\leftrightarrow$.
- Opening parenthesis '(', closing parenthesis ')', and comma ', '.
- Countably infinite list of variables $x_0, x_1, x_2, x_3, \ldots$.
- The universal quantifier $\forall$ that is read as '*for all*'.
- The existential quantifier $\exists$ that is read as '*there exists at least one*'.
- The constant symbols, function symbols, and relation symbols of $\mathcal{L}$.

Thus,

$$\mathcal{A}_{\mathcal{L}} = \{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\} \cup \{\,),(,\,,\} \cup \{\,\forall, \exists\,\} \cup \{x_0, x_1, x_2, \ldots\} \cup \mathcal{L}.$$

The symbols in $\mathcal{L}$ are sometimes called the *non-logical* symbols of the alphabet. We also use the symbols x, y, z for variables as well. Next, we aim to define special words over this alphabet which will be called terms and first-order formulas. Terms are constructed with the intention that when they are interpreted will denote elements of an $\mathcal{L}$-structure. Formulas are built with the intention of describing properties of an $\mathcal{L}$-structure.

3.2.1 Terms

The ingredients to build terms are variables, constant symbols, and function symbols. We define terms inductively as follows.

Let $\mathcal{L}$ be a first-order language. A *term* of $\mathcal{L}$ is constructed as follows.

- Any constant symbol of $\mathcal{L}$ is a term.
- Any variable from $\{x_0, x_1, x_2, \ldots\}$ is a term.
- If $f \in \mathcal{L}$ is an n-ary function symbol and $t_1, t_2, \ldots, t_n$ are terms, then the word $f(t_1, t_2, \ldots, t_n)$ is a term as well.
- Nothing else is a term.

More precisely, the set of all terms is the union of a chain of subsets defined inductively as follows.

Definition.

- Put $\mathcal{T}_0 = \{c \in \mathcal{L} \mid c \text{ is a constant symbol}\} \cup \{x_0, x_1, x_2, x_3, \ldots\}$.
- For any $k \in \mathbb{N}$, we define $\mathcal{T}_{k+1}$ to be the set
$$\mathcal{T}_k \cup \{f(t_1, \ldots, t_n) \mid f \in \mathcal{L} \text{ is an } n\text{-ary function symbol and } t_1, \ldots, t_n \in \mathcal{T}_k\}.$$
- The set $\mathcal{T}$ of all terms of $\mathcal{L}$ is defined as
$$\mathcal{T} = \bigcup_{k \in \mathbb{N}} \mathcal{T}_k.$$

Observe that if the language has no function symbols at all, then $\mathcal{T} = \mathcal{T}_0$, meaning that the only terms of the language are the variables and the constant symbols. A term which has no occurrences of variables is called a *closed term.* As the reader would expect, the *height* of a term t is the least natural number k such that $t \in \mathcal{T}_k$.

Lemma 3.2.1. *The set $\mathcal{T}$ of all terms of $\mathcal{L}$ is the smallest subset of the set $\mathcal{W}(\mathcal{A}_{\mathcal{L}})$ of all words over the alphabet $\mathcal{A}_{\mathcal{L}}$ which contains all the constant symbols and variables, and is closed under all operations of the form*

$$[w_1, w_2, \ldots, w_n] \mapsto f(w_1, w_2, \ldots, w_n)$$

where f is an n-ary function symbol of $\mathcal{L}$.

Example. Suppose that $\mathcal{L}$ has two constant symbols a, c, one unary function symbol f, one binary function symbol g, and one ternary function symbol h. Then the following are terms of $\mathcal{L}$.

- $a,\ c,\ x,\ y,\ z.$
- $f(a),\ f(c),\ f(x),\ f(y).$
- $g(a,a),\ g(a,c),\ g(c,x),\ g(y,a),\ g(x,y).$
- $h(a,c,c),\ h(x,y,z),\ h(y,z,c),\ h(z,z,z),\ h(a,a,a),\ h(x,x,y),\ h(a,x,c).$
- $f(f(a)),\ f(f(x)),\ f(g(a,c)),\ f(g(x,y)),\ f(h(a,a,a)),\ f(h(y,z,c)).$
- $g(c,f(x)),\ g(f(x),f(y)),\ g(f(c),g(a,a)),\ g(g(c,x),g(x,y)),\ g(f(c),h(a,c,c)),$ $g(h(a,x,c),g(a,c)),\ g(h(y,z,c),h(a,a,a)).$
- $h(a,x,f(c)),\ h(a,f(y),g(x,y)),\ h(f(x),g(x,y),h(x,y,z)),$ $h(h(a,c,z),h(x,y,z),h(c,c,c)).$
- $f(f(f(x))),\ f(f(g(a,c))),\ f(f(h(y,z,c))).$

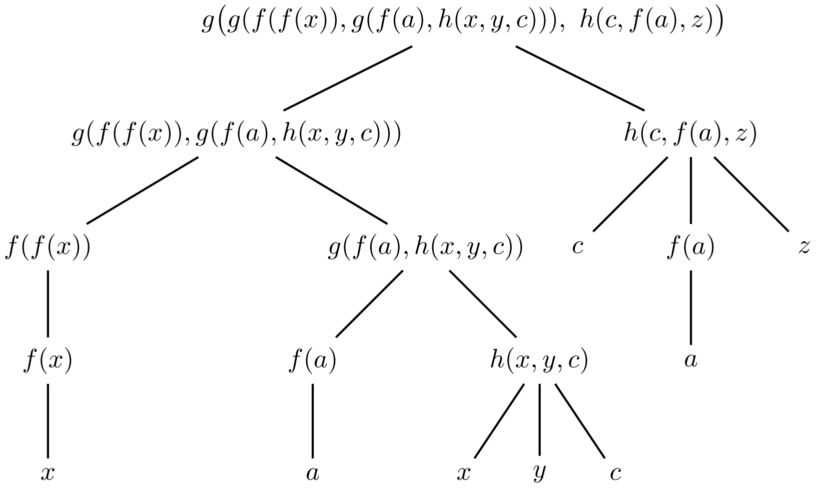
- $g(g(c,f(x)),g(f(c),g(a,a)))$, $g(g(g(c,x),g(x,y)),g(f(c),h(a,c,c)))$, $g(h(f(x),g(x,y),h(x,y,z)),h(h(a,c,z),h(x,y,z),h(c,c,c)))$.
- $h(f(f(c)),g(f(c),g(a,a)),h(a,f(y),g(x,y)))$, $h(h(a,x,f(c)),h(h(a,a,a),h(x,x,x),h(c,c,c)),h(g(a,c),g(x,y),h(c,c,c)))$.

♠

Example. Below is the decomposition tree of the term

$$g\big(g(f(f(x)),g(f(a),h(x,y,c))),\ h(c,f(a),z)\big).$$

Observe that the leaves of the tree are either constant symbols or variables. Moreover, the number of branches going down a non-leaf node is exactly the arity of the function symbol used at that stage of construction.



♠

Notation. Suppose that $t \in \mathcal{T}$ is a term of some language and $i_1, i_2, \ldots, i_k$ are distinct natural numbers. We will use the notation $t = t(x_{i_1}, x_{i_2}, \ldots, x_{i_k})$ to indicate that the variables that occur in the term t are among the variables $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$. For example, if the term $t = g(f(x_2), h(x_4, c, x_2))$, then we write $t(x_2, x_4)$ or $t(x_0, x_1, x_2, x_3, x_4)$.

3.2.2 First-Order Formulas

Definition. Suppose that R is an n-ary relation symbol and $t_1, t_2, \ldots, t_n$ are terms of a first-order language $\mathcal{L}$. Then the word

$$R(t_1, t_2, \ldots, t_n)$$

is called an *atomic formula.*

Notation. When dealing with the equality symbol the atomic formula $= (t_1, t_2)$ may be written as $t_1 = t_2$ or $(t_1 = t_2)$.

Example. Let $\mathcal{L} = \{c, f, g, P, R\}$ be a first-order language consisting of a constant symbol c, a unary function symbol f, a binary function symbol g, a unary relation symbol P and a binary relation symbol R. The following are atomic formulas.

- $P(c)$, $P(x)$, $R(c, c)$, $R(c, x)$, $R(x, y)$, $x = y$, $c = c$, $x = c$.
- $P(f(c))$, $P(f(x))$, $P(g(c, x))$, $R(f(x), f(y))$, $R(f(c), g(x, y))$, $f(x) = c$, $g(x, y) = z$, $g(x, y) = c$, $f(y) = g(x, y)$.
- $P(g(f(x), g(c, c)))$, $R(g(c, c), g(f(x), f(c)))$, $g(f(x), g(c, c)) = f(y)$.

We have arrived to the point of describing the process of constructing a *first-order formula*, which is also called an *$\mathcal{L}$-formula* or a *well-formed formula* (WFF or wff).

- An atomic formula is a first-order formula.
- Suppose that φ and ψ are first-order formulas, then

  $$\neg\varphi,\ (\varphi \wedge \psi),\ (\varphi \vee \psi),\ (\varphi \rightarrow \psi),\ (\varphi \leftrightarrow \psi)$$

  are first-order formulas as well.
- Suppose that φ is first-order formula and x is a variable, then

  $$\forall x\, \varphi \quad \text{and} \quad \exists x\, \varphi$$

  are first-order formulas as well.
- Nothing else is a first-order formula.

We say that the formula $\forall x\, \varphi$ is the *universal quantification* of the formula φ with respect to the variable x. Similarly, the formula $\exists x\, \varphi$ is said to be the *existential quantification* of the formula φ with respect to the variable x. Alternatively, to describe the construction of $\forall x\, \varphi$ from φ we say that the variable x has been *universally quantified.* Similarly, in the construction of $\exists x\, \varphi$ from φ we say that the variable x has been *existentially quantified.* Let us define the set of all first-order formulas $\mathcal{F}$ in a precise way.

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

- We set $\mathcal{F}_0$ to be the set of all atomic formulas of $\mathcal{L}$.
- For each natural number n, we define

  $$\begin{aligned}\mathcal{F}_{n+1} = \mathcal{F}_n\ &\cup\ \{\neg\varphi \mid \varphi \in \mathcal{F}_n\}\ \cup\ \{(\varphi \wedge \psi) \mid \varphi, \psi \in \mathcal{F}_n\} \cup \{(\varphi \vee \psi) \mid \varphi, \psi \in \mathcal{F}_n\}\\ &\cup\ \{(\varphi \rightarrow \psi) \mid \varphi, \psi \in \mathcal{F}_n\}\ \cup\ \{(\varphi \leftrightarrow \psi) \mid \varphi, \psi \in \mathcal{F}_n\}\\ &\cup\ \{\forall x_i\, \varphi \mid \varphi \in \mathcal{F}_n, i \in \mathbb{N}\}\ \cup\ \{\exists x_i\, \varphi \mid \varphi \in \mathcal{F}_n, i \in \mathbb{N}\}.\end{aligned}$$

- We define the set $\mathcal{F}$ of all first-order formulas of $\mathcal{L}$ to be

$$\mathcal{F} = \bigcup_{n\in\mathbb{N}} \mathcal{F}_n.$$

We also say an $\mathcal{L}$-formula for a first-order formula of a language $\mathcal{L}$. The formula $\neg(t_1 = t_2)$ may be written as $t_1 \neq t_2$ or $(t_1 \neq t_2)$. As it was done previously, the *height* of a first-order formula φ, denoted by $h[\varphi]$, is the least natural number n such that $\varphi \in \mathcal{F}_n$.

Example. Let $\mathcal{L} = \{c, f, g, P, R\}$ be a first-order language consisting of a constant symbol c, a unary function symbol f, a binary function symbol g, a unary relation symbol P and a binary relation symbol R. The following are first-order formulas.

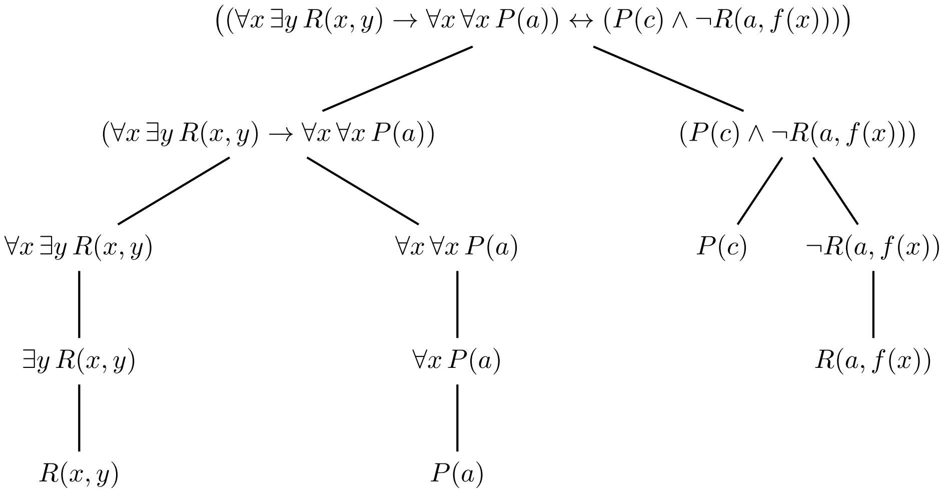
- $\neg P(c)$, $\neg R(c, y)$, $(P(c) \wedge P(x))$, $(P(x) \leftrightarrow R(c, x))$, $(R(x, y) \to x = y)$.
- $(P(f(c)) \wedge P(f(x)))$, $(P(g(c, x)) \vee R(f(x), f(y)))$
- $\forall x\, P(c)$, $\forall x\, P(x)$, $\forall x\, P(y)$, $\forall x\, R(c, y)$, $\forall y\, (P(x) \leftrightarrow R(c, x))$.
- $\exists x\, P(c)$, $\exists x\, P(x)$, $\exists x\, P(z)$, $\exists x\, (P(g(c, y)) \vee R(f(x), f(y)))$.
- $\forall x\, \forall x\, \forall x\, \forall x\, P(x)$, $\forall x_0\, \forall x_1\, \forall x_2\, \forall x_3\, R(x_2, x_3)$, $\exists x\, \exists y\, R(x, c)$, $\exists y\, \exists z\, R(x, x)$.
- $\forall x\, \exists y\, \exists x\, \forall y\, (P(x) \vee R(c, z))$, $\forall x_3\, \exists x_1\, \forall x_0\, \exists x_0\, (P(x_2) \vee R(c, x_0))$.

♠

Example. Below is the decomposition tree of the first-order formula

$$\big((\forall x\, \exists y\, R(x, y) \to \forall x\, \forall x\, P(a)) \leftrightarrow (P(c) \wedge \neg R(a, f(x)))\big).$$

Observe that the leaves of the tree are atomic formulas.


$\big((\forall x\, \exists y\, R(x, y) \to \forall x\, \forall x\, P(a)) \leftrightarrow (P(c) \wedge \neg R(a, f(x)))\big)$
$(\forall x\, \exists y\, R(x, y) \to \forall x\, \forall x\, P(a))$
$(P(c) \wedge \neg R(a, f(x)))$
$\forall x\, \exists y\, R(x, y)$
$\forall x\, \forall x\, P(a)$
$P(c)$
$\neg R(a, f(x))$
$\exists y\, R(x, y)$
$\forall x\, P(a)$
$R(a, f(x))$
$R(x, y)$
$P(a)$

♠

3.2.3 Free Variables and Bound Variables

An occurrence of a variable x in a first-order formula is either free or bound. An occurrence of a variable x is bound if it falls under the action of a quantifier $\forall x$ or $\exists x$. In other words, for any formula φ, all the occurrences of a variable x in the formulas $\forall x\, \varphi$ and $\exists x\, \varphi$ are said to be bound. An occurrence of a variable x which is not controlled by these quantifiers is said to be free. Generally some occurrences of a given variable in a formula could be free and other occurrences of the same variable may be bound. Let us define these concepts precisely.

Definition. Suppose that x and y are distinct variables from the set $\{x_0, x_1, x_2, \ldots\}$. We define whether an occurrence of the variable x in a formula φ is *free* or *bound* inductively as follows.

- If φ is an atomic formula, then every occurrence of x in φ is free.
- If $\varphi = \neg\psi$, then the free occurrences of x in φ are exactly those free occurrences of x in ψ.
- If $\varphi = (\psi \diamond \theta)$ where $\diamond$ is a binary propositional connective, then the free occurrences of x in φ are exactly those free occurrences of x in ψ and the free occurrences of x in θ.
- If $\varphi = \forall y\, \psi$ or $\varphi = \exists y\, \psi$, then the free occurrences of x in φ are exactly those free occurrences of x in ψ.
- If $\varphi = \forall x\, \psi$ or $\varphi = \exists x\, \psi$, then none of the occurrences of x in φ is free. An occurrence of a variable which is not free is called bound.

Definition. A variable in a first-order formula is called a *free variable* if it has at least one free occurrence in the formula.

Definition. A first-order formula which has no free variables is called a *sentence* or a *closed formula.*

Example. Let x, y, z, w be distinct variables, and let $\mathcal{L} = \{c, f, R\}$ be a language where c is a constant symbol, f is a unary function symbol, and R is a binary relation symbol. In the first-order formulas of $\mathcal{L}$ below, the **dotted** occurrences of variables are **free**, and the un-dotted occurrences of variables are bound.

- $R(\dot{x}, \dot{y})$ (Both x and y are free variables.)
- $\exists z\, R(\dot{x}, \dot{y})$ (Both x and y are free variables.)
- $\exists x\, R(x, \dot{y})$ (Only y is free variables.)
- $\exists y\, \exists x\, R(x, y)$ (This is an $\mathcal{L}$-sentence, i.e., no free variables.)

- $R(c, f(c))$ (This is an $\mathcal{L}$-sentence.)
- $(\exists x\, R(x, \dot{y}) \wedge \neg R(\dot{x}, \dot{y}))$ (Both x and y are free variables.)
- $(\exists x\, R(x, \dot{y}) \wedge \forall y\, R(\dot{x}, y))$ (Both x and y are free variables.)
- $\forall y\, (\exists x\, R(x, y) \wedge \forall y\, R(\dot{x}, y))$ (Only x is a free variables.)
- $(\exists x\, (R(x, \dot{y}) \to R(c, \dot{z})) \vee \forall z\, \exists y\, (R(\dot{x}, y) \leftrightarrow R(\dot{x}, z)))$ (x, y, z are free variables.)
- $\forall x\, (\exists y\, \forall x\, (R(y, x) \to x = \dot{w}) \wedge \forall w (\exists z\, (R(\dot{y}, z) \vee R(f(x), c)) \wedge \dot{z} = \dot{z}))$

♠

Notation. Suppose that φ is a first-order formula and $i_1, i_2, \ldots, i_n$ are distinct natural numbers. Then we write

$$\varphi = \varphi(x_{i_1}, x_{i_2}, \ldots, x_{i_n})$$

to indicate that the free variables of φ are among the variables $x_{i_1}, x_{i_2}, \ldots, x_{i_n}$.

Definition. In the formula $\forall x\, \varphi$ the occurrence of x occurring immediately after the quantifier $\forall$ together with all free occurrences of x in the formula φ are said to be *within the scope* of the quantifier $\forall x$. Similarly, we treat the formula $\exists x\, \varphi$.

Observe that an occurrence of a variable can be within the scope of at most one quantifier. For example, in the formula

$$\forall x\, (((x = x) \vee \exists x\, (x \neq x)) \to x = x),$$

only the first three and the last two occurrences of the variable x are within the scope of the first quantifier $\forall x$. The fourth, fifth, and sixth occurrences of x are not within the scope of quantifier $\forall x$ but within the scope of the second quantifier $\exists x$.

3.3 Semantics of First-Order Logic

In the below we work with a first-order language $\mathcal{L}$.

3.3.1 Interpretation of Terms

Definition. Let $t = t(x_1, x_2, \dots, x_n)$ be a term of $\mathcal{L}$. Let $\mathcal{M}$ be an $\mathcal{L}$-structure with underlying set M and let $a_1, a_2, \dots, a_n \in M$. The *interpretation of the term t in the structure $\mathcal{M}$ when the variables $x_1, x_2, \dots, x_n$ are interpreted respectively by the elements $a_1, a_2, \dots, a_n$* is an element of M, denoted by

$$t^{\mathcal{M}}(a_1/x_1, a_2/x_2, \dots, a_n/x_n) \quad \text{or} \quad t^{\mathcal{M}}(a_1, a_2, \dots, a_n)$$

and is defined by induction on the term t as follows:

- if $t = c$ where $c \in \mathcal{L}$ is a constant symbol, then
$$t^{\mathcal{M}}(a_1/x_1, a_2/x_2, \dots, a_n/x_n) = c^{\mathcal{M}}\,;$$
- if $t = x_i$ where $1 \leq i \leq n$, then
$$t^{\mathcal{M}}(a_1/x_1, a_2/x_2, \dots, a_n/x_n) = a_i\,;$$
- if $t = f(t_1, t_2, \dots, t_k)$ where $f \in \mathcal{L}$ is a k-ary function symbol and $t_1, t_2, \dots, t_k$ are terms of $\mathcal{L}$, then
$$t^{\mathcal{M}}(a_1/x_1, \dots, a_n/x_n) = f^{\mathcal{M}}(t_1^{\mathcal{M}}(a_1/x_1, \dots, a_n/x_n), \dots, t_k^{\mathcal{M}}(a_1/x_1, \dots, a_n/x_n))\,.$$

To compute the interpretation of a term in a structure it would be helpful to build the decomposition tree of the term and then follow the definition above to evaluate the interpretations of the nodes of the tree in the given structure from the leaves all the way to the very top.

Given a term $t = t(x_1, x_2, \dots, x_n)$ of $\mathcal{L}$ we see the interpretation of t in $\mathcal{M}$ gives us a function from M^n to M given by the map

$$(a_1, a_2, \dots, a_n) \mapsto t^{\mathcal{M}}(a_1, a_2, \dots, a_n).$$

Example. Consider the first-order language $\mathcal{L} = \{a, c, f, g, h\}$ where a, c are constant symbols, f, g are binary function symbols, h is a unary function symbol. Consider the term

$$t = t(x, y, z) = f(g(f(x, y), g(z, z)), h(x))$$

Take the exponential field of real numbers as an $\mathcal{L}$-structure:

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

Then

$$\begin{aligned} t^{\mathcal{M}}(2/x,\ 4/y,\ 0.5/z) &= f^{\mathcal{M}}(g^{\mathcal{M}}(f^{\mathcal{M}}(2/x,\ 4/y), g^{\mathcal{M}}(0.5/z,\ 0.5/z)), h^{\mathcal{M}}(2/x)) \\ &= f^{\mathcal{M}}(g^{\mathcal{M}}(f^{\mathcal{M}}(2,4), g^{\mathcal{M}}(0.5,0.5)), h^{\mathcal{M}}(2)) \\ &= f^{\mathcal{M}}(g^{\mathcal{M}}((2+4),(0.5\cdot 0.5)), \exp(2)) \\ &= f^{\mathcal{M}}(g^{\mathcal{M}}(6,0.25), \exp(2)) \\ &= f^{\mathcal{M}}((6\cdot 0.25), \exp(2)) \\ &= f^{\mathcal{M}}(1.5, \exp(2)) = 1.5+\exp(2) = 1.5+e^2. \end{aligned}$$

Now consider the term $s(x) = f(g(f(c,c), g(a,c)), h(x))$. Then the interpretation of s in $\mathcal{M}$ is

$$\begin{aligned} s^{\mathcal{M}}(3/x) &= f^{\mathcal{M}}(g^{\mathcal{M}}(f^{\mathcal{M}}(c^{\mathcal{M}}, c^{\mathcal{M}}), g^{\mathcal{M}}(a^{\mathcal{M}}, c^{\mathcal{M}})), h^{\mathcal{M}}(3/x)) \\ &= f^{\mathcal{M}}(g^{\mathcal{M}}(f^{\mathcal{M}}(1,1), g^{\mathcal{M}}(0,1)), h^{\mathcal{M}}(3)) \\ &= f^{\mathcal{M}}(g^{\mathcal{M}}((1+1),(0\cdot 1)), \exp(3)) \\ &= f^{\mathcal{M}}(g^{\mathcal{M}}(2,0), \exp(3)) \\ &= f^{\mathcal{M}}((2\cdot 0), \exp(3)) \\ &= f^{\mathcal{M}}(0, \exp(3)) = 0+\exp(3) = \exp(3) = e^3. \end{aligned}$$

In a similar fashion, show that $s^{\mathcal{M}}(0/x) = 1$. ♠

3.3.2 Satisfaction of Formulas

In the below we work with a first-order language $\mathcal{L}$. Recall that the notation $\varphi = \varphi(x_1, x_2, \ldots, x_n)$ indicates that the free variables in the formula φ are among the variables $x_1, x_2, \ldots, x_n$.

Definition. Suppose that $\varphi(x_1, x_2, \ldots, x_n)$ is a first-order formula of $\mathcal{L}$ and let $\mathcal{M}$ be an $\mathcal{L}$-structure and let $a_1, a_2, \ldots, a_n$ be elements of the domain of $\mathcal{M}$. We write

$$\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n) \quad \text{or} \quad \mathcal{M} \models \varphi(a_1, a_2, \ldots, a_n)$$

for *the formula φ is satisfied in the structure $\mathcal{M}$ when the variables $x_1, x_2, \ldots, x_n$ are interpreted respectively by the elements $a_1, a_2, \ldots, a_n$* (the symbol $\models$ is read 'satisfies'). The satisfaction of a formula in a structure is defined by induction on the formula $\varphi = \varphi(x_1, x_2, \ldots, x_n)$ as follows.

- Suppose that the formula $\varphi(x_1, x_2, \ldots, x_n)$ is an atomic formula, say $\varphi = R(t_1, t_2, \ldots, t_k)$ where $R \in \mathcal{L}$ is a k-ary relation symbol and $t_1, t_2, \ldots, t_k$ are $\mathcal{L}$-terms. So the variables in each term t_i are among $x_1, x_2, \ldots, x_n$. Then we define $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ if and only if

$$\big(t_1^{\mathcal{M}}(a_1/x_1, a_2/x_2, \ldots, a_n/x_n),\ \ldots,\ t_k^{\mathcal{M}}(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)\big) \in R^{\mathcal{M}}.$$

- Suppose that $\varphi = \neg\psi(x_1, x_2, \ldots, x_n)$. Then $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ if and only if $\mathcal{M} \not\models \psi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$. (We write $\mathcal{M} \not\models \psi$ when $\mathcal{M}$ does not satisfy ψ.)

- Suppose $\varphi = (\psi \wedge \theta)$. Then $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ if and only if $\mathcal{M} \models \psi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ and $\mathcal{M} \models \theta(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$.

- Suppose $\varphi = (\psi \vee \theta)$. Then $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ if and only if $\mathcal{M} \models \psi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ or $\mathcal{M} \models \theta(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$.

- Suppose $\varphi = (\psi \to \theta)$. Then $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ if and only if $\mathcal{M} \not\models \psi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ or $\mathcal{M} \models \theta(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$.

- Suppose $\varphi = (\psi \leftrightarrow \theta)$. Then $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ if and only if $\mathcal{M} \models \psi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ and $\mathcal{M} \models \theta(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$;
  or else
  $\mathcal{M} \not\models \psi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ and $\mathcal{M} \not\models \theta(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$.

- Suppose $\varphi = \forall y\, \psi(x_1, x_2, \ldots, x_n, y)$ where $y \notin \{x_1, \ldots, x_n\}$. Then $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ if and only if for every element b in $\mathcal{M}$ we have that $\mathcal{M} \models \psi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n, b/y)$.

- Suppose $\varphi = \exists y\, \psi(x_1, x_2, \ldots, x_n, y)$ where $y \notin \{x_1, \ldots, x_n\}$. Then $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ if and only if there exists at least one element b in $\mathcal{M}$ such that $\mathcal{M} \models \psi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n, b/y)$.

- Suppose $\varphi = \forall x_i\, \psi(x_1, x_2, \ldots, x_n)$ where $1 \leq i \leq n$. Then $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ if and only if for every element b in $\mathcal{M}$ we have that $\mathcal{M} \models \psi(a_1/x_1, \ldots, a_{i-1}/x_{i-1},\ b/x_i,\ a_{i+1}/x_{i+1}, \ldots, a_n/x_n)$.

- Suppose $\varphi = \exists x_i\, \psi(x_1, x_2, \ldots, x_n)$ where $1 \leq i \leq n$. Then $\mathcal{M} \models \varphi(a_1/x_1, \ldots, a_i/x_i, \ldots, a_n/x_n)$ if and only if there is at least one element b in $\mathcal{M}$ such that $\mathcal{M} \models \psi(a_1/x_1, \ldots, a_{i-1}/x_{i-1},\ b/x_i,\ a_{i+1}/x_{i+1}, \ldots, a_n/x_n)$.

There are various ways to read $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$ including the following.

1. The structure $\mathcal{M}$ satisfies the formula φ when the elements $a_1, \ldots, a_n$ interpret respectively the variables $x_1, \ldots, x_n$.

2. The formula φ is true in $\mathcal{M}$ when $a_1, \ldots, a_n$ interpret respectively $x_1, \ldots, x_n$.

3. The formula φ is satisfied in $\mathcal{M}$ by the tuple $(a_1, \ldots, a_n)$.

4. The tuple $(a_1, \ldots, a_n)$ satisfies the formula $\varphi(x_1, \ldots, x_n)$ in the structure $\mathcal{M}$.

When the formula φ is a sentence (i.e. has no free variables) we simply write

$$\mathcal{M} \models \varphi$$

when φ is satisfied in $\mathcal{M}$. For any $\mathcal{L}$-sentence σ and any $\mathcal{L}$-structure $\mathcal{M}$ exactly one of the following occurs: either $\mathcal{M} \models \sigma$ or $\mathcal{M} \not\models \sigma$. For the former we say σ is true in $\mathcal{M}$ and for the latter we say that σ is false in $\mathcal{M}$. When $\mathcal{M} \models \sigma$ we also say that the structure $\mathcal{M}$ is a *model* of σ. Here comes the name 'Model Theory'.

More generally, when φ is a formula whose free variables are exactly $x_1, x_2, \dots, x_n$, we say that φ is satisfied in $\mathcal{M}$ (or φ is true in $\mathcal{M}$) and write $\mathcal{M} \models \varphi$ if the structure $\mathcal{M}$ satisfies the sentence $\forall x_1 \cdots \forall x_n\, \varphi$. This means that the formula $\varphi(x_1, \dots, x_n)$ is true in $\mathcal{M}$ under any interpretation of the variables $x_1, \dots, x_n$, in other words, we have $\mathcal{M} \models \varphi(a_1/x_1, a_2/x_2, \dots, a_n/x_n)$ for all elements $a_1, a_2, \dots, a_n$ in $\mathcal{M}$.

Example. Let $\mathcal{L} = \{g\}$ where g is a binary function symbol. Consider the formula

$$\varphi = \varphi(x) = \forall y\, \forall z\, (g(x, y) = g(x, z) \to y = z).$$

Let $\mathcal{M} = (\mathbb{Z}, \cdot)$. Then $\mathcal{M} \models \varphi(2/x)$ holds. To see this,

$$\begin{aligned}
\mathcal{M} \models \varphi(2/x) \ \text{ iff } & \text{ for any } a \in \mathbb{Z}\text{: } \mathcal{M} \models \forall z\, (g(x, y) = g(x, z) \to y = z) \text{ when } 2/x, a/y \\
\text{ iff } & \text{ for any } a \in \mathbb{Z}, \text{ for any } b \in \mathbb{Z}\text{:} \\
& \quad \mathcal{M} \models (g(x, y) = g(x, z) \to y = z) \text{ when } 2/x, a/y, b/z \\
\text{ iff } & \text{ for any } a, b \in \mathbb{Z}\text{: } \mathcal{M} \not\models (g(x, y) = g(x, z)) \text{ when } 2/x, a/y, b/z \\
& \quad \text{ or } \ \mathcal{M} \models (y = z) \text{ when } 2/x, a/y, b/z \\
\text{ iff } & \text{ for any } a, b \in \mathbb{Z}\text{: } g^{\mathcal{M}}(2, a) \neq g^{\mathcal{M}}(2, b) \ \text{ or } \ (a = b) \\
\text{ iff } & \text{ for any } a, b \in \mathbb{Z}\text{: } (2 \cdot a \neq 2 \cdot b) \ \text{ or } \ (a = b), \ \text{ which is true.}
\end{aligned}$$

In a similar argument one may show that $\mathcal{M} \not\models \varphi(0/x)$. ♠

Lemma 3.3.1. *Suppose that $\mathcal{M}$ is an $\mathcal{L}$-structure and $\psi(x, y_1, \dots, y_k)$ is an $\mathcal{L}$-formula and x does not appear free in ψ. Then for any elements $a, b, c_1, c_2, \dots, c_k$ in $\mathcal{M}$ we have that*

$$\mathcal{M} \models \psi(a/x, c_1/y_1, \dots, c_k/y_k) \text{ if and only if } \mathcal{M} \models \psi(b/x, c_1/y_1, \dots, c_k/y_k).$$

Proof. Exercise. Proceed by induction on first-order formulas. ■

Definition. Let $\varphi(x_1, \dots, x_n)$ be an $\mathcal{L}$-formula and let $\mathcal{M}$ be an $\mathcal{L}$-structure with domain M. Then the set $\varphi(\mathcal{M})$ of all solutions of the formula φ in the structure $\mathcal{M}$ is the set of all n-tuples of elements of M which satisfy φ in $\mathcal{M}$, that is,

$$\varphi(\mathcal{M}) = \big\{(a_1, \dots, a_n) \in M^n \mid \mathcal{M} \models \varphi(a_1, \dots, a_n)\big\}.$$

Example. Let $\mathcal{L} = \{c, f, R\}$ be a first-order language where c is a constant symbol, f is a unary function symbol, and R is a binary relation symbol, and consider the $\mathcal{L}$-structure $\mathcal{M} = (\mathbb{R}, \pi, \cos, \leq)$.

1)	$\varphi(x) = R(c, x)$	$\varphi(\mathcal{M}) = \{a \in \mathbb{R} \mid \pi \leq a\} = [\pi, \infty)$
2)	$\varphi(x) = \exists y\,(f(y) = x)$	$\varphi(\mathcal{M}) = [-1, 1]$
3)	$\varphi(x) = \exists y\,(f(x) = y)$	$\varphi(\mathcal{M}) = \mathbb{R}$
4)	$\varphi(x) = (f(x) = c)$	$\varphi(\mathcal{M}) = \{a \in \mathbb{R} \mid \cos(a) = \pi\} = \emptyset$
5)	$\varphi(x) = \exists y\,(R(c, y) \wedge f(y) = x)$	$\varphi(\mathcal{M}) = [-1, 1]$
6)	$\varphi(x) = \forall y\, R(x, f(y))$	$\varphi(\mathcal{M}) = \{a \in \mathbb{R} \mid a \leq -1\} = (-\infty, -1]$
7)	$\varphi(x) = \forall y\, R(f(x), f(y))$	$\varphi(\mathcal{M}) = \{(2k+1)\pi \mid k \in \mathbb{Z}\}$
8)	$\varphi(x) = \forall x\, \exists y\,(f(y) = x)$	$\varphi(\mathcal{M}) = \emptyset$
9)	$\varphi(x) = \exists y\, \forall z\, R(f(z), y)$	$\varphi(\mathcal{M}) = \mathbb{R}$

Observe that the next-to-last formula is a sentence (closed formula) that is false in $\mathcal{M}$, while the last formula is a true sentence in $\mathcal{M}$. ♠

3.4 First-Order Theories

In mathematics, the word ‘theory’ refers to all the mathematical consequences of a set of axioms. It also refers to the collection of properties shared by a class of structures. By a *first-order theory* we shall mean the set of all consequences of some set of sentences in a first-order language which are called the *axioms* of the theory. We shall make the notion of a ‘consequence’ precise soon. A first-order structure satisfying all the axioms of a theory is called a *model* of the theory. In what follows we will present well-known mathematical theories by describing their axioms.

Equivalence Relations

The theory of *equivalence relations* has the following axioms written in the first-order language $\mathcal{L} = \{R\}$ where R is a binary relation symbol.

1. $\forall x\, R(x,x)$ (R is reflexive)
2. $\forall x\, \forall y\, (R(x,y) \to R(y,x))$ (R is symmetric)
3. $\forall x\, \forall y\, \forall z\, ((R(x,y) \wedge R(y,z)) \to R(x,z))$ (R is transitive)

Let $\mathcal{A} = (A, R^{\mathcal{A}})$ be an $\mathcal{L}$-structure which satisfies the three axioms above, in other words, it is a model of these axioms. Then we get the following properties about the relation $R^{\mathcal{A}} \subseteq A \times A$. Let's denote the three axioms above by φ_1, φ_2, φ_3 respectively.

1. Since $\mathcal{A} \models \varphi_1$, we have for every $a \in A$, the pair $(a,a) \in R^{\mathcal{A}}$.
2. Since $\mathcal{A} \models \varphi_2$, we have for every $a, b \in A$, if $(a,b) \in R^{\mathcal{A}}$, then $(b,a) \in R^{\mathcal{A}}$.
3. Since $\mathcal{A} \models \varphi_3$, we have for every $a, b, c \in A$, if $(a,b) \in R^{\mathcal{A}}$ and $(b,c) \in R^{\mathcal{A}}$, then $(a,c) \in R^{\mathcal{A}}$.

It follows that any structure which is a model of these three axioms is a set together with an equivalence relation defined on the set. A model of these three axioms is simply called an equivalence relation. The intuition behind equivalence relations is that when a pair (a,b) belongs to an equivalence relation it means that the element a is the same as b in a certain way.

Strict Partial Orders

The theory of *strict partial orders* has the following axioms written in the language $\mathcal{L} = \{<\}$ where $<$ is a binary relation symbol. We will write $(x < y)$ for the atomic formula $< (x,y)$.

1. $\forall x\, \neg(x < x)$ (< is irreflexive)

2. $\forall x\, \forall y\, \forall z\, (((x < y) \wedge (y < z)) \to (x < z))$ (< is transitive)

Strict Linear Orders

The theory of *strict linear (total) orders* has the axioms of strict partial orders together with the following axiom.

3. $\forall x\, \forall y\, ((x < y) \vee (x = y) \vee (y < x))$ (< is linear or total)

A model of the first two axioms is called a strict partial order and a model of the three axioms is called a strict linear order. So every linear order is a partial order.

Dense Linear Orders without Endpoints (DLO)

The theory of *dense linear orders without endpoints* has the axioms of strict linear orders together with the following two axioms.

4. $\forall x\, \forall y\, (x < y \to \exists z\, (x < z \wedge z < y))$ (< is dense)

5. $\forall x\, \exists y\, \exists z\, (y < x \wedge x < z)$ (No endpoints)

Non-Strict Partial Orders

The theory of *non-strict partial order* has the following axioms written in the language $\mathcal{L} = \{\leq\}$ where $\leq$ is a binary relation symbol. We will write $(x \leq y)$ for the atomic formula $\leq (x, y)$.

1. $\forall x\, (x \leq x)$ ($\leq$ is reflexive)

2. $\forall x\, \forall y\, (((x \leq y) \wedge (y \leq x)) \to (x = y))$ ($\leq$ is antisymmetric)

3. $\forall x\, \forall y\, \forall z\, (((x \leq y) \wedge (y \leq z)) \to (x \leq z))$ ($\leq$ is transitive)

Non-Strict Linear Orders

The theory of *non-strict linear order* has the axioms of non-strict partial orders together with the following axiom.

4. $\forall x\, \forall y\, ((x \leq y) \vee (y \leq x))$ ($\leq$ is linear or total)

Infinite Sets

The theory of *infinite sets* has infinitely many axioms written in the language of equality $\mathcal{L} = \{=\}$. For each integer $n \geq 2$, we have an axiom φ_n which says there are at least n distinct elements expressed as follows.

$$\begin{aligned} \varphi_2 &= \exists x_1 \, \exists x_2 (x_1 \neq x_2) \\ \varphi_3 &= \exists x_1 \, \exists x_2 \, \exists x_3 (x_1 \neq x_2 \land x_1 \neq x_3 \land x_2 \neq x_3) \end{aligned}$$

In general,

$$\varphi_n = \exists x_1 \, \exists x_2 \ldots \exists x_n \bigwedge_{1 \leq i < j \leq n} x_i \neq x_j$$

Thus the axioms for the theory of infinite sets will be the following set of $\mathcal{L}$-sentences.

$$T = \{\varphi_n \mid n \geq 2\}.$$

Groups

The theory of *groups* has the following axioms expressed in the language $\mathcal{L} = \{e, \circ\}$ where e is a constant symbol and $\circ$ is a binary function symbol. We will write $x \circ y$ for the term $\circ(x, y)$.

1. $\forall x \, \forall y \, \forall z (x \circ (y \circ z) = (x \circ y) \circ z)$ ($\circ$ is associative)
2. $\forall x \, ((x \circ e = x) \land (e \circ x = x))$ (e is an identity element)
3. $\forall x \, \exists y \, ((x \circ y = e) \land (y \circ x = e))$ (Every element has an inverse)

The theory of *abelian groups* has the previous three axioms together with the following additional axiom.

4. $\forall x \, \forall y \, (x \circ y = y \circ x)$ ($\circ$ is commutative)

A model of the first three axioms is called a group and a model of the four axioms is called an abelian group.

Rings

The theory of *rings* has the following axioms expressed in the language $\mathcal{L} = \{0, +, \cdot, -\}$ where 0 is a constant symbol, $+$ and $\cdot$ are binary function symbols, and $-$ is a unary function symbol. We will write $x + y$ for the term $+(x, y)$, and write $x \cdot y$ for the term $\cdot(x, y)$, and write $-x$ for the term $-(x)$.

1. $\forall x \forall y \forall z(x + (y + z) = (x + y) + z)$ (+ is associative)
2. $\forall x\, ((x + 0 = x) \wedge (0 + x = x))$ (0 is an additive identity)
3. $\forall x\, ((x + (-x) = 0) \wedge ((-x) + x = 0))$ (Additive inverses exist)
4. $\forall x \forall y\, (x + y = y + x)$ (+ is commutative)
5. $\forall x \forall y \forall z(x \cdot (y \cdot z) = (x \cdot y) \cdot z)$ ($\cdot$ is associative)
6. $\forall x \forall y \forall z((x \cdot (y + z) = x \cdot y + x \cdot z) \wedge ((y + z) \cdot x = y \cdot x + z \cdot x))$ (Distributivity)

A model of these six axioms is called a ring.

Commutative Rings with Unity

The axioms of the theory of *commutative rings with unity* are expressed in the language $\mathcal{L} = \{0, 1, +, \cdot, -\}$ where 0 and 1 are constant symbol, + and $\cdot$ are binary function symbols, and $-$ is a unary function symbol. They include the axioms of rings plus the following two axioms.

7. $\forall x \forall y\, (x \cdot y = y \cdot x)$ ($\cdot$ is commutative)
8. $\forall x\, ((x \cdot 1 = x) \wedge (1 \cdot x = x))$ (1 is a multiplicative identity called unity)

A model of these eight axioms is called a commutative ring with unity.

Fields

The axioms of the theory of *fields* are expressed in the language $\mathcal{L} = \{0, 1, +, \cdot, -, ^{-1}\}$ where 0 and 1 are constant symbol, + and $\cdot$ are binary function symbols, and $-$ and $^{-1}$ are unary function symbols. They include the axioms of commutative rings with unity plus the following axiom which when satisfied it says that nonzero elements have multiplicative inverses. We write x^{-1} instead of $^{-1}(x)$.

9. $\forall x\, (x \neq 0 \rightarrow ((x \cdot x^{-1} = 1) \wedge (x^{-1} \cdot x = 1)))$

A model of these nine axioms is called a field. Suppose that $\mathcal{M}$ is a field with an underlying set M. This means that $\mathcal{M}$ is an $\mathcal{L}$-structure where $\mathcal{L} = \{0, 1, +, \cdot, -, ^{-1}\}$ satisfying the nine axioms of the theory of fields. Observe that the triple consisting of the set M together with $0^{\mathcal{M}}$ and $+^{\mathcal{M}}$ forms an abelian group, and the triple consisting of the set $M \setminus \{0^{\mathcal{M}}\}$ together with $1^{\mathcal{M}}$ and $\cdot^{\mathcal{M}}$ also forms an abelian group.

Exercise. Write down the axioms for the theory of *ordered fields* in the language $\mathcal{L} = \{0, 1, +, \cdot, -, ^{-1}, <\}$.

3.5 Logical Consequence

Recall that for an $\mathcal{L}$-formula φ whose free variables are exactly $x_1, x_2, \ldots, x_n$, we say that an $\mathcal{L}$-structure $\mathcal{M}$ satisfies φ if $\mathcal{M}$ satisfies the sentence $\forall x_1 \cdots \forall x_n \, \varphi$.

Definition. Fix some first-order language $\mathcal{L}$.

- An $\mathcal{L}$-formula φ is *logically valid* (or *universally valid*) if every $\mathcal{L}$-structure satisfies φ. We write $\models \varphi$ for this.
- Two $\mathcal{L}$-formulas φ and ψ are *logically equivalent* if the formula $(\varphi \leftrightarrow \psi)$ is logically valid. We write $\varphi \equiv \psi$ for this.
- Let T be a set of $\mathcal{L}$-sentences and let $\mathcal{M}$ be an $\mathcal{L}$-structure. We say that $\mathcal{M}$ *satisfies* T if $\mathcal{M}$ satisfies every sentence that belongs to T. In symbols, $\mathcal{M} \models \sigma$ for every $\sigma \in T$. When $\mathcal{M}$ satisfies T, we also say that $\mathcal{M}$ is a *model* of T and write
$$\mathcal{M} \models T.$$
- A set of $\mathcal{L}$-sentences is called *satisfiable* if it has at least one model.
- Given a set of $\mathcal{L}$-sentences T and an $\mathcal{L}$-sentence φ, we say that φ is a *logical consequence* of T, or T *logically implies* φ, if every model of T is also a model of φ. For this we write
$$T \models \varphi.$$
- A *theory* of $\mathcal{L}$ or an $\mathcal{L}$*-theory* is a set T of $\mathcal{L}$-sentences which is satisfiable and closed under logical consequence, i.e. if $T \models \varphi$, then $\varphi \in T$.

  Note. Some authors define a theory to be just a set of sentences; others define a theory to be a satisfiable set of sentences.
- An $\mathcal{L}$-theory is *complete* if for every $\mathcal{L}$-sentence σ, either $\sigma \in T$ or $\neg\sigma \in T$.

Remark. More generally, given a set of $\mathcal{L}$-formulas Γ and an $\mathcal{L}$-formula θ, we say that θ is a *logical consequence* of Γ, or Γ *logically implies* θ, and write $\Gamma \models \theta$, if for every $\mathcal{L}$-structure $\mathcal{M}$ and every sequence $a_0, a_1, a_2, \ldots$ of elements of $\mathcal{M}$ respectively interpreting the variables $x_0, x_1, x_2, \ldots$, if $\mathcal{M} \models \gamma$ with these interpretations for all $\gamma \in \Gamma$, then $\mathcal{M} \models \theta$ with the same interpretations.

We would like to draw the attention of the reader to the point that the symbol '$\models$' is used in two ways. On one hand, the statement $\mathcal{M} \models \varphi$ means that the structure $\mathcal{M}$ satisfies the sentence φ. On the other hand, the statement $T \models \varphi$ means that the sentence φ is a logical consequence of a set T of sentences.

Logically valid formulas correspond to those statements which are always true because of their form, regardless of their interpretations. For example, let R be a binary relation symbol, then the formula $(\forall x \, R(x, y) \to \forall x \, R(x, y))$ is logically valid.

Example. Let $\mathcal{L} = \{P\}$ where P is a unary relation symbol. The sentence

$$(\exists x\, P(x) \to \forall x\, P(x))$$

is not logically valid. To see this, consider the $\mathcal{L}$-structure $\mathcal{M} = (\mathbb{N}, P^{\mathcal{M}})$ where $P^{\mathcal{M}}$ is the set of all even natural numbers. Observe that $\mathcal{M} \models \exists x\, P(x)$ since the there is an element of $\mathcal{M}$ which belongs to the set $P^{\mathcal{M}}$, for instance $4 \in P^{\mathcal{M}}$, however, $\mathcal{M} \not\models \forall x\, P(x)$ since not every element of $\mathcal{M}$ belongs to the set $P^{\mathcal{M}}$, in particular, $5 \notin P^{\mathcal{M}}$. Thus, $\mathcal{M} \not\models (\exists x\, P(x) \to \forall x\, P(x))$, and so $(\exists x\, P(x) \to \forall x\, P(x))$ is not satisfied by every $\mathcal{L}$-structure, so it is not logically valid. ♠

Definition. Let $\mathcal{L}$ be a first-order language. An $\mathcal{L}$-formula φ is called a *tautology* of $\mathcal{L}$ if there exists a tautology α in propositional logic built-up from the propositional variables $p_1, \dots, p_n$ and there exist first-order formulas $\psi_1, \dots, \psi_n$ of $\mathcal{L}$ such that φ is obtained from α by replacing each occurrence of p_i by ψ_i for all $i = 1, 2, \dots, n$. In this case, we say that φ is a *substitution instance* of α.

Example. Let $\mathcal{L} = \{R\}$ where R is a binary relation symbol. The $\mathcal{L}$-formula $(\exists x\, R(x,y) \to \exists x\, R(x,y))$ is a first-order tautology because it can be obtained from the propositional tautology $(p \to p)$ by replacing the propositional variable p with the $\mathcal{L}$-formula $\exists x\, R(x,y)$.

Theorem 3.5.1. *Every first-order tautology is logically valid.*

Proof. Suppose that α is a propositional formula built-up from the propositional variables $p_1, p_2, \dots, p_n$, and suppose that $\psi_1, \psi_2, \dots, \psi_n$ are first-order formulas of a language $\mathcal{L}$. Let φ be the $\mathcal{L}$-formula obtained by substituting every ψ_i for p_i in α. Let $\mathcal{M}$ be an $\mathcal{L}$-structure and define the truth assignment $\delta : \{p_1, p_2, \dots, p_n\} \to \{0,1\}$ as follows:

$$\delta[p_i] = \begin{cases} 1 & \text{if } \mathcal{M} \models \psi_i; \\ 0 & \text{if } \mathcal{M} \not\models \psi_i. \end{cases}$$

By induction on the propositional formula α we can show the following claim:

$$\mathcal{M} \models \varphi \text{ if and only if } \delta[\alpha] = 1.$$

Now suppose that φ is a tautology of $\mathcal{L}$ obtained from a propositional tautology α. Let $\mathcal{M}$ be any $\mathcal{L}$-structure and let δ be the truth assignment defined above. Since α is a tautology, we get that $\delta[\alpha] = 1$ and thus, by the claim, $\mathcal{M} \models \varphi$. Therefore, φ is satisfied in any $\mathcal{L}$-structure, and so φ is logically valid. ■

Definition. A first-order formula which has no quantifiers is called a *quantifier-free formula.* A formula is in *prenex normal form* (PNF) if it is quantifier-free or has the form

$$Q_1 x_1\, Q_2 x_2\, \dots\, Q_n x_n\, \theta$$

where each Q_i is a quantifier ($\forall$ or $\exists$), and θ is a quantifier-free formula. The string of the quantifiers $Q_1 x_1\, Q_2 x_2\, \dots\, Q_n x_n$ is called the *prefix* of the formula.

For example,

$$\exists x\, \forall y\, (P(y) \to \neg R(x,y))$$

is in prenex normal form. However, the formula

$$\exists x\, (\forall y\, P(y) \to \neg R(x,y))$$

is not in prenex normal form.

Theorem 3.5.2. *Any first-order formula is logically equivalent to a formula in prenex normal form.*

Chapter 4

Gödel's Completeness Theorem

The completeness theorem of first-order logic for countable languages was proved by Kurt Gödel in his doctoral dissertation in 1930 at the University of Vienna. The compactness theorem for countable languages was given as a corollary of Gödel's work. In 1936 in Russia, Anatoly Mal'cev proved the compactness theorem for uncountable languages. His proof used Skolem functions and the compactness theorem for propositional logic. The compactness theorem is regarded as one of the most important results in logic.

4.1 Substitution of Terms for Variables

Definition. Fix a first-order language $\mathcal{L}$. Let φ be an $\mathcal{L}$-formula, t be an $\mathcal{L}$-term, and x be a variable.

- By $\varphi(t/x)$ or $\varphi(t)$ we denote the formula obtained by replacing every free occurrence of the variable x in φ by the term t.
- The term t is said to be *freely substitutable* for x in φ if none of the variables in t becomes bound in $\varphi(t/x)$.

Thus, a term t is freely substitutable for x in φ if x has no free occurrences in φ within the scope of a quantifier $\forall y$ or $\exists y$ where y is a variable occurring in t. In simple words, it means that t can be substituted for every free occurrence of x in φ without introducing any interactions between the variables in t and the quantifiers in φ. Observe that in the case when the variable x is not free in φ, the formula $\varphi(t)$ is the same as the original formula φ, and we would still say that t is freely substitutable for x in φ in this case.

Example. Let $\mathcal{L} = \{f, R\}$ where f is a binary function symbol and R is a binary relation symbol. Let w, x, y, z be variables. Consider the formula

$$\varphi = (\exists x\,\exists w(R(x,\dot{y}) \vee \forall z\,(w = z)) \to \forall y\,R(y,\dot{x})).$$

Let t be the term $f(w, z)$. Substitute t for free x in φ to obtain

$$\varphi(t/x) = (\exists x\, \exists w(R(x, y) \vee \forall z\, (w = z)) \to \forall y\, R(y, f(w, z))).$$

Moreover, t is freely substitutable for x in φ since the variables w, z of t were not bound by the quantifier $\forall y$ in $\varphi(t/x)$.

However, when we substitute the term t for free y in φ we obtain the formula

$$\varphi(t/y) = (\exists x\, \exists w(R(x, f(w, z)) \vee \forall z\, (w = z)) \to \forall y\, R(y, x))$$

where t is not freely substitutable for y in φ since now the variable w of the term t falls within the scope of the quantifier $\exists w$ and so this w of t becomes bound in $\varphi(t/y)$. ♠

Theorem 4.1.1. *Suppose that φ is an $\mathcal{L}$-formula and t is an $\mathcal{L}$-term freely substitutable for the variable x in φ. Then the formula*

$$(\forall x\, \varphi \to \varphi(t/x))$$

is logically valid.

The next example explains why we need the term t to be freely substitutable for x in φ in the hypothesis of the theorem above.

Example. Let $\mathcal{L} = \{g\}$ where g is a unary function symbol. Consider the $\mathcal{L}$-formula

$$\varphi = \exists x(g(x) = y).$$

The universal quantification of φ with respect to y is the formula

$$\forall y\, \varphi = \forall y\, \exists x(g(x) = y)$$

which says that the interpretation of the function symbol g is a surjective function. Now consider the term $t = x$, and substitute t for free y in φ to get the formula

$$\varphi(t/y) = \exists x(g(x) = x).$$

The new formula $\varphi(t/y)$ says that g has a fixed-point. Observe that t is not freely substitutable for y in φ. Now consider the $\mathcal{L}$-structure $\mathcal{M} = (\mathbb{Z}, g^{\mathcal{M}})$ where $g^{\mathcal{M}}$ is the successor function, that is, $g^{\mathcal{M}} : \mathbb{Z} \to \mathbb{Z}$ given by $g^{\mathcal{M}}(n) = n+1$ for every $n \in \mathbb{Z}$. Since the successor function is surjective, it follows that $\mathcal{M} \models \forall y\, \varphi$. However, $\mathcal{M} \not\models \varphi(t/y)$, that is, $\mathcal{M} \not\models \exists x(g(x) = x)$ since the successor function has no fixed-points. Therefore,

$$\mathcal{M} \not\models (\forall y\, \varphi \to \varphi(t/y)).$$

Thus, $(\forall y\, \varphi \to \varphi(t/y))$ is not logically valid. ♠

4.2 A Proof System for First-Order Logic

To each first-order language $\mathcal{L}$ we introduce a formal system $\mathcal{S}_{\mathcal{L}}$ with the following axioms and deduction rules.

Axioms of $\mathcal{S}_{\mathcal{L}}$

We have the following axiom schemes where φ, ψ, θ are first-order formulas of $\mathcal{L}$.

(Ax 1) $\big(\varphi \to (\psi \to \varphi)\big)$

(Ax 2) $\big((\varphi \to (\psi \to \theta)) \to ((\varphi \to \psi) \to (\varphi \to \theta))\big)$

(Ax 3) $\big((\neg\varphi \to \neg\psi) \to (\psi \to \varphi)\big)$

(Ax 4) $(\forall x\, \varphi \to \varphi(t/x))$ where t is a term freely substitutable for x in φ

(Ax 5) $(\forall x\, (\varphi \to \psi) \to (\varphi \to \forall x\, \psi))$ where x is not a free variable of φ

(Ax 6) $\forall x\, (x = x)$

(Ax 7) $(x = y \to (\varphi \to \varphi'))$ where φ is an atomic formula and φ' is obtained from φ by replacing some (not necessarily all) occurrences of x in φ by y

Deduction Rules of $\mathcal{S}_{\mathcal{L}}$

Modus Ponens. From formulas φ and $(\varphi \to \psi)$, derive the formula ψ.

$$\text{(MP)} \quad \frac{\varphi,\ (\varphi \to \psi)}{\psi}$$

Generalisation. From a formula φ, derive the formula $\forall x\, \varphi$.

$$\text{(Gen)} \quad \frac{\varphi}{\forall x\, \varphi}$$

Remark. Axiom 4 tells us how to eliminate a universal quantifier. Observe that a formula of the form $(\forall x\, \varphi \to \varphi)$ where x is not a free variable of φ is an instance of Axiom 4. The Generalisation rule tells us how to introduce a universal quantifier. Axiom 5 describes how $\forall$ and $\to$ interact with each other. Axiom 6 and Axiom 7 are axioms describing the behaviour of the equality symbol. Axiom 7 says that if two objects are equal, then any property of the first is also a property of the second.

In a similar fashion as was done in propositional logic we now define a formal proof in first-order logic.

Definition. Let Γ be a set of first-order formulas and let ψ be a first-order formula. A *derivation* (or a *proof*) of ψ from Γ within the system $\mathcal{S}_{\mathcal{L}}$ is a finite sequence of first-order formulas

$$\varphi_1, \varphi_2, \dots, \varphi_n,$$

where the last formula φ_n in the sequence is the formula ψ and where each formula φ_k in the sequence satisfies one of the following:

(i) $\varphi_k \in \Gamma$;

(ii) φ_k is one of the axioms of the system $\mathcal{S}_{\mathcal{L}}$;

(iii) φ_k is deduced by MP, i.e. there are formulas φ_i and φ_j in the sequence such that $i < k$ and $j < k$, and $\varphi_j = (\varphi_i \to \varphi_k)$.

(iv) φ_k is deduced by the Generalisation rule, i.e. $\varphi_k = \forall x\, \varphi_i$ where $i < k$.

Definition. We say that ψ is *derivable* (or *provable*) from Γ if there exists a derivation of ψ from Γ within the system $\mathcal{S}_{\mathcal{L}}$. We write

$$\Gamma \vdash \psi$$

when ψ is derivable from Γ. When $\emptyset \vdash \psi$ we write $\vdash \psi$ and say that ψ is a *theorem* of the system $\mathcal{S}_{\mathcal{L}}$. We write $\varphi \vdash \psi$ for $\{\varphi\} \vdash \psi$.

Example. Let φ be an $\mathcal{L}$-formula where a variable y does not appear in φ. Let $\varphi(y/x)$ be the formula obtained by replacing every free occurrence of the variable x in φ by the variable y. Notice that y is freely substitutable for x in φ. Show that

$$\forall x\, \varphi \vdash \forall y\, \varphi(y/x).$$

1.	$\forall x\, \varphi$	Assumption
2.	$(\forall x\, \varphi \to \varphi(y/x))$	Ax 4
3.	$\varphi(y/x)$	MP 1, 2
4.	$\forall y\, \varphi(y/x)$	Gen 3

For example, take φ to be $(R(x,z) \wedge \exists x\, R(x,x))$ where R is a binary relation symbol, then we get that $\forall x\, (R(x,z) \wedge \exists x\, R(x,x)) \vdash \forall y\, (R(y,z) \wedge \exists x\, R(x,x))$. ♠

Example. Let t be an $\mathcal{L}$-term. Then the atomic formula $(t = t)$ is a theorem of $\mathcal{S}_{\mathcal{L}}$. In symbols,

$$\vdash (t = t)$$

1.	$\forall x\, (x = x)$	Ax 6
2.	$(\forall x\, (x = x) \to (t = t))$	Ax 4
3.	$(t = t)$	MP 1, 2

♠

Example. Let x, y be variables. Then

$$(x = y) \vdash (y = x)$$

1.	$\forall x\,(x = x)$	Ax 6
2.	$(\forall x\,(x = x) \to (x = x))$	Ax 4
3.	$(x = x)$	MP 1, 2
4.	$(x = y)$	Assumption
5.	$((x = y) \to ((x = x) \to (y = x)))$	Ax 7
6.	$((x = x) \to (y = x))$	MP 4, 5
7.	$(y = x)$	MP 3, 6

♠

Theorem 4.2.1. *Suppose that ψ is a tautology of a first-order language $\mathcal{L}$. Then ψ is a theorem of the system $\mathcal{S}_{\mathcal{L}}$, in symbols, $\vdash \psi$.*

Proof. Let ψ be a tautology of $\mathcal{L}$. Then there is a tautology α in propositional logic built-up from propositional variables $p_1, p_2, \ldots, p_n$ together with $\mathcal{L}$-formulas $\theta_1, \theta_2, \ldots, \theta_n$ such that ψ is obtained from α by replacing every p_i with θ_i for every $i = 1, 2, \ldots, n$. Since α is a tautology, i.e. $\models \alpha$, it follows by the completeness theorem of Propositional Logic that $\vdash_{\mathcal{S}} \alpha$ within the system $\mathcal{S}$. This derivation of α uses only Ax 1, Ax 2, Ax 3, and MP of the system $\mathcal{S}$. Next, we transform this derivation to a derivation of ψ within the system $\mathcal{S}_{\mathcal{L}}$ by replacing every propositional variable p_i with the $\mathcal{L}$-formula θ_i throughout the derivation. The result is a derivation in $\mathcal{S}_{\mathcal{L}}$ since Ax 1, Ax 2, Ax 3, and MP are common to both proof systems $\mathcal{S}$ and $\mathcal{S}_{\mathcal{L}}$. Therefore, $\vdash_{\mathcal{S}_{\mathcal{L}}} \psi$ as desired. ■

Example. Here is an application of the proof above. Let φ be any $\mathcal{L}$-formula. The formula $(\forall\varphi \to \forall\varphi)$ is a tautology of $\mathcal{L}$ obtained from the propositional tautology $(p \to p)$. The derivation below was obtained from the derivation of $(p \to p)$ in $\mathcal{S}$ by replacing every occurrence of p with $\forall\varphi$.

1.	$\big(\forall\varphi \to (\forall\varphi \to \forall\varphi)\big)$	Ax 1
2.	$\big(\forall\varphi \to ((\forall\varphi \to \forall\varphi) \to \forall\varphi)\big)$	Ax 1
3.	$\big((\forall\varphi \to ((\forall\varphi \to \forall\varphi) \to \forall\varphi)) \to ((\forall\varphi \to (\forall\varphi \to \forall\varphi)) \to (\forall\varphi \to \forall\varphi))\big)$	Ax 2
4.	$((\forall\varphi \to (\forall\varphi \to \forall\varphi)) \to (\forall\varphi \to \forall\varphi))$	MP 2, 3
5.	$(\forall\varphi \to \forall\varphi)$	MP 1, 4

Thus $(\forall\varphi \to \forall\varphi)$ is a theorem of $\mathcal{S}_{\mathcal{L}}$.

Ax 7 is restricted to atomic formulas, the next theorem says that we can derive in the system $\mathcal{S}_{\mathcal{L}}$ instances of Ax 7 that cover all formulas.

Theorem 4.2.2. *Let $\mathcal{L}$ be a first-order language and let φ be any $\mathcal{L}$-formula. Denote by φ' the formula obtained from φ by replacing some (not necessarily all) of the free occurrences of x in φ by y, provided that y is freely substitutable for these occurrences of x. Then*

$$\vdash (x = y \to (\varphi \to \varphi')).$$

We even can show a more general form of the previous theorem.

Corollary 4.2.3. *Let $\mathcal{L}$ be a first-order language and let φ be any $\mathcal{L}$-formula and t be any $\mathcal{L}$-term. Denote by φ'' the formula obtained from φ by replacing some (not necessarily all) of the free occurrences of x in φ by the term t, provided that t is freely substitutable for these occurrences of x. Then*

$$\vdash (x = t \to (\varphi \to \varphi'')).$$

Proof. Let y be a variable not occurring in φ, so y is freely substitutable for x in φ. Let φ' be the formula obtained from φ by replacing the free occurrences of x chosen when forming φ'' by y. We proceed as follows.

1.	$(x = y \to (\varphi \to \varphi'))$	Theorem 4.2.2
2.	$\forall y\,(x = y \to (\varphi \to \varphi'))$	Gen 1
3.	$(\forall y\,(x = y \to (\varphi \to \varphi')) \to (x = t \to (\varphi \to \varphi'')))$	Ax 4 (t is freely sub. for y)
4.	$(x = t \to (\varphi \to \varphi''))$	MP 2, 3

■

4.3 The Soundness Theorem

We aim to show that the system $\mathcal{S}_\mathcal{L}$ is sound, that is, the existence of a derivation of ψ from Γ in $\mathcal{S}_\mathcal{L}$ implies that ψ is a logical consequence of Γ. Consequently, all theorems of $\mathcal{S}_\mathcal{L}$ are logically valid formulas. As was done in the soundness theorem of propositional logic, the first step towards establishing the first-order version of the theorem is to show that all axioms of $\mathcal{S}_\mathcal{L}$ are logically valid.

Theorem 4.3.1. *All of the seven axioms of the system $\mathcal{S}_\mathcal{L}$ are logically valid.*

Proof. Ax 1, Ax 2, and Ax 3 are first-order tautologies. To see this, Ax 1 is obtained from the tautology $\big(p \to (q \to p)\big)$, Ax 2 is obtained from the tautology $\big((p \to (q \to r)) \to ((p \to q) \to (p \to r))\big)$, and Ax 3 is obtained from the tautology $\big((\neg p \to \neg q) \to (q \to p)\big)$. Check Lemma 2.3.1 to see that these propositional formulas are indeed tautologies. By Theorem 3.5.1 we know that every first-order tautology is logically valid, and so all instances of Ax 1, Ax 2, and Ax 3 of $\mathcal{S}_\mathcal{L}$ are logically valid first-order formulas.

By Theorem 4.1.1, Ax 4 is logically valid.

Ax 5 is logically valid: exercise.

We now show that Ax 6 is logically valid. To this end, let $\mathcal{M}$ be any $\mathcal{L}$-structure with domain M.

$$\begin{aligned}\mathcal{M} \models \forall x\,(x = x) \ &\text{iff}\ \text{ for any } a \in M\text{: } \mathcal{M} \models (x = x) \text{ when } a/x\\ &\text{iff}\ \text{ for any } a \in M\text{: } a = a\\ &\text{iff}\ \text{ for any } a \in M\text{: } (a, a) \in {=^{\mathcal{M}}}, \quad \text{which is true.}\end{aligned}$$

Thus, every $\mathcal{L}$-structure satisfies Ax 6, and so it is logically valid.

Finally, we show that Ax 7 is logically valid. Ax 7 is the formula $(x = y \to (\varphi \to \varphi'))$ where φ is an atomic formula and φ' is obtained from φ by replacing some (not necessarily all) occurrences of x in φ by y. As φ is atomic, it has the form $R(t_1, t_2, \ldots, t_n)$ where R is an n-ary relation symbol and $t_1, t_2, \ldots, t_n$ are $\mathcal{L}$-terms in the variables, say, $x, y, z_1, z_2, \ldots, z_k$. Let t' be the term obtained from a term t by replacing some (not necessarily all) of the occurrences of x in t by y. It follows that $\varphi' = R(t_1', t_2', \ldots, t_n')$. We need to show that Ax 7 is satisfied in every $\mathcal{L}$-structure under any interpretations of the variables appearing in Ax 7. So let $\mathcal{M}$ be any $\mathcal{L}$-structure with domain M and let $a, b, c_1, c_2, \ldots, c_k$ be elements in M interpreting respectively the variables $x, y, z_1, z_2, \ldots, z_k$. In order to show that $\mathcal{M} \models$ Ax 7 when a/x, b/y, $c_1/z_1, \ldots, c_k/z_k$, we assume that $\mathcal{M} \models (x = y)$ and $\mathcal{M} \models \varphi$ under these interpretations, and aim to show that $\mathcal{M} \models \varphi'$ under the same interpretations. By induction on terms we can show the following claim.

Claim. Suppose that $t = t(x, y, z_1, \ldots, z_n)$ is an $\mathcal{L}$-term. Under the interpretations a/x, b/y, $c_1/z_1, \ldots, c_k/z_k$, if $\mathcal{M} \models (x = y)$, then $t^{\mathcal{M}} = t'^{\mathcal{M}}$.

By our assumption $\mathcal{M} \models (x = y)$ and the claim, it follows that $t_i^{\mathcal{M}} = t_i'^{\mathcal{M}}$ for each $i = 1, 2, \ldots, n$. We now proceed as follows where the satisfaction is computed under the interpretations a/x, b/y, $c_1/z_1, \ldots, c_k/z_k$.

$$\begin{aligned} \mathcal{M} \models \varphi \text{ iff } & \mathcal{M} \models R(t_1, t_2, \ldots, t_n) \\ \text{iff } & \mathcal{M} \models (t_1^{\mathcal{M}}, t_2^{\mathcal{M}}, \ldots, t_n^{\mathcal{M}}) \in R^{\mathcal{M}} \\ \text{iff } & \mathcal{M} \models (t_1'^{\mathcal{M}}, t_2'^{\mathcal{M}}, \ldots, t_n'^{\mathcal{M}}) \in R^{\mathcal{M}} \\ \text{iff } & \mathcal{M} \models R(t_1', t_2', \ldots, t_n') \\ \text{iff } & \mathcal{M} \models \varphi'. \end{aligned}$$

Since $\mathcal{M} \models \varphi$, it follows that $\mathcal{M} \models \varphi'$ as well. Thus, we have shown that for any structure $\mathcal{M}$ whenever $\mathcal{M} \models (x = y)$ and $\mathcal{M} \models \varphi$, it must be that $\mathcal{M} \models \varphi'$. Therefore, $\mathcal{M} \models (x = y \to (\varphi \to \varphi'))$ showing that Ax 7 is logically valid. ■

Lemma 4.3.2. *Let θ be an $\mathcal{L}$-formula whose free variables are exactly $y_1, \ldots, y_k$ and let $\mathcal{M}$ be an $\mathcal{L}$-structure. Then*

$$\mathcal{M} \models \theta \textit{ if and only if } \mathcal{M} \models \forall x\, \theta$$

for any variable x.

Proof. We first leave it for the reader to show the special case when the formula is a sentence. That is, if σ is an $\mathcal{L}$-sentence, then $\mathcal{M} \models \sigma$ if and only if $\mathcal{M} \models \forall x\, \sigma$. We then proceed as follows for an $\mathcal{L}$-formula θ whose free variables are exactly $y_1, \ldots, y_k$.

$$\begin{aligned} \mathcal{M} \models \theta \ \text{ iff } \ & \mathcal{M} \models \forall y_1 \cdots \forall y_k\, \theta \\ \text{iff } \ & \mathcal{M} \models \forall x\, \forall y_1 \cdots \forall y_k\, \theta \\ \text{iff } \ & \mathcal{M} \models \forall y_1 \cdots \forall y_k \forall x\, \theta \\ \text{iff } \ & \mathcal{M} \models \forall x\, \theta. \end{aligned}$$

■

Theorem 4.3.3 (Soundness Theorem for $\mathcal{S}_{\mathcal{L}}$)**.** *Let T be a set of first-order sentences and let ψ be a first-order formula. If T proves ψ, then T logically implies ψ. In symbols,*

$$\textit{If } T \vdash \psi, \textit{ then } T \models \psi.$$

Proof. We will show by mathematical induction on the length n of the derivation the following:

If $\alpha_1, \alpha_2, \ldots, \alpha_n$ is a derivation from T in the system $\mathcal{S}_{\mathcal{L}}$ and $\mathcal{M}$ is an $\mathcal{L}$-structure that satisfies T, then $\mathcal{M}$ satisfies α_n.

Base case. Suppose that $n = 1$ (one-step derivation) and let $\mathcal{M}$ be an $\mathcal{L}$-structure which satisfies T. We need to show that $\mathcal{M} \models \alpha_1$. Observe that α_1 being the first formula of a derivation either belongs to T or is an axiom of $\mathcal{S}_{\mathcal{L}}$. For the former case, as $\mathcal{M}$ satisfies every formula in T we get $\mathcal{M} \models \alpha_1$. For the latter case, by Theorem 4.3.1, we know that axioms of $\mathcal{S}_{\mathcal{L}}$ are logically valid, and so satisfied in every $\mathcal{L}$-structure.

Induction step. Suppose that result holds for all derivations of length strictly less than n. That is, if $\beta_1, \beta_2, \dots, \beta_k$ is a derivation from T and $k < n$ and $\mathcal{M} \models T$, then $\mathcal{M} \models \beta_k$. Consider a derivation $\alpha_1, \alpha_2, \dots, \alpha_n$ from T and let $\mathcal{M}$ be a structure which satisfies T. We need to show that $\mathcal{M} \models \alpha_n$. The formula α_n being in a derivation either belongs to T, is an axiom of $\mathcal{S}_{\mathcal{L}}$, or deduced by Modus Ponens or by Generalisation rule. If α_n is in T or an axiom, then $\mathcal{M}$ satisfies α_n as discussed in the base case.

Otherwise, α_n may have been deduced in the derivation by Modus Ponens. So there are previous formulas α_i and α_j in the derivation such that $i < n$ and $j < n$ and $\alpha_j = (\alpha_i \to \alpha_n)$. By induction hypothesis, we have that $\mathcal{M} \models \alpha_i$ and $\mathcal{M} \models \alpha_j$. Since $\mathcal{M} \models \alpha_j$, i.e. $\mathcal{M} \models (\alpha_i \to \alpha_n)$, it follows by definition of satisfaction that either $\mathcal{M} \not\models \alpha_i$ or $\mathcal{M} \models \alpha_n$. Since $\mathcal{M} \models \alpha_i$, it must be that $\mathcal{M} \models \alpha_n$ as desired.

We are left with α_n being inferred by an application of the Generalisation rule. Therefore, for some previous formula α_i where $i < n$ we have that $\alpha_n = \forall x \, \alpha_i$. By induction hypothesis, we have that $\mathcal{M} \models \alpha_i$. By Lemma 4.3.2, we obtain that $\mathcal{M} \models \forall x \, \alpha_i$, that is, $\mathcal{M} \models \alpha_n$ as desired. This completes the induction step and establishes the theorem. ■

Corollary 4.3.4. *Every theorem of the system $\mathcal{S}_{\mathcal{L}}$ is logically valid. That is,*

$$\text{if } \vdash \varphi, \text{ then } \models \varphi$$

for every $\mathcal{L}$-formula φ.

As in propositional logic we have the converse of the Deduction Theorem as stated below.

Lemma 4.3.5. *Let Γ be a set of $\mathcal{L}$-formulas, and φ, ψ be $\mathcal{L}$-formulas. Then, if $\Gamma \vdash (\varphi \to \psi)$, then $\Gamma \cup \{\varphi\} \vdash \psi$.*

We now head towards establishing a restricted version of the Deduction Theorem for the system $\mathcal{S}_{\mathcal{L}}$. The Deduction Theorem in its full generality as we know from propositional logic does not hold for $\mathcal{S}_{\mathcal{L}}$ as illustrated by the next example.

Example. Let φ be any first-order formula. Then it is clear that $\varphi \vdash \forall x \, \varphi$ (by a two-step derivation: the first is φ, an assumption, and the second by an application of Gen). If the Deduction Theorem in its general form were true for $\mathcal{S}_{\mathcal{L}}$ we would expect to have $\vdash (\varphi \to \forall x\,\varphi)$, and so $(\varphi \to \forall x\,\varphi)$ is a theorem of $\mathcal{S}_{\mathcal{L}}$. By the Soundness Theorem, if follows that $(\varphi \to \forall x\,\varphi)$ is logically valid, i.e true in all $\mathcal{L}$-structures. However, we will show that this is not the case by giving a structure where it is not satisfied. The problem arises when we apply the Generalisation rule with respect to a variable which occurs free in φ.

Let $\mathcal{L} = \{P\}$ where P is a unary relation symbol and let $\varphi = P(x)$. Consider the $\mathcal{L}$-structure $\mathcal{N} = (\mathbb{N}, A)$ where $A = P^{\mathcal{N}} = \{0,1,2\}$. We will show that $\mathcal{N} \not\models (P(x) \to \forall x\,P(x))$. Observe that x is the only free variable in $(P(x) \to \forall x\,P(x))$. Thus,

$$\begin{aligned}
\mathcal{N} \models (P(x) \to \forall x\,P(x)) \;\; &\text{iff} \;\; \mathcal{N} \models \forall x\,(P(x) \to \forall x\,P(x)) \\
&\text{iff} \;\; \text{for any } a \in \mathbb{N}\text{: } \mathcal{N} \models (P(x) \to \forall x\,P(x)) \text{ when } a/x \\
&\text{iff} \;\; \text{for any } a \in \mathbb{N}\text{: } \mathcal{N} \not\models P(x) \text{ when } a/x \text{ or } \mathcal{N} \models \forall x\,P(x) \\
&\text{iff} \;\; \text{for any } a \in \mathbb{N}\text{: } (a \notin P^{\mathcal{N}} \quad \text{or} \\
&\qquad\qquad\qquad \text{for all } b \in \mathbb{N}\text{: } \mathcal{N} \models P(x) \text{ when } b/x) \\
&\text{iff} \;\; \text{for any } a \in \mathbb{N}\text{: } (a \notin P^{\mathcal{N}} \;\text{ or }\; \text{for all } b \in \mathbb{N}\text{: } b \in P^{\mathcal{N}}) \\
&\text{iff} \;\; \text{for any } a \in \mathbb{N}\text{: } \;(a \notin A \;\text{ or }\; \text{for all } b \in \mathbb{N}\text{: } b \in A).
\end{aligned}$$

The very last statement says that either every natural number does not belong to A or every natural number belongs to A, which is false. Thus, $\mathcal{N} \not\models (P(x) \to \forall x\,P(x))$, showing that $(P(x) \to \forall x\,P(x))$ is not logically valid. ♠

Theorem 4.3.6 (Deduction Theorem for $\mathcal{S}_{\mathcal{L}}$). *Let $\mathcal{L}$ be a first-order language. For any set Γ of $\mathcal{L}$-formulas and any $\mathcal{L}$-formula φ and $\mathcal{L}$-formula ψ,*

if $\Gamma \cup \{\varphi\} \vdash \psi$ and the derivation contains no application of Generalisation rule involving a variable which occurs free in φ, then $\Gamma \vdash (\varphi \to \psi)$.

Proof. Let Γ be a set of first-order formulas, and φ be a first-order formula. We will prove the theorem by mathematical induction on the length of the derivation. More precisely, we will prove by induction on n the following statement:

If $\alpha_1, \alpha_2, \ldots, \alpha_n$ is a derivation in $\mathcal{S}_{\mathcal{L}}$ from $\Gamma \cup \{\varphi\}$, then $\Gamma \vdash (\varphi \to \alpha_n)$.

The proof is very similar to the proof of the Deduction Theorem of $\mathcal{S}$ in propositional logic, Theorem 2.2.2, with one more part taking care of the Generalisation rule in the induction step.

Let $n > 1$ and suppose the result holds for *all* derivations of length strictly less than n. We will show that the result holds for derivations of length n, so let

$$\alpha_1, \;\alpha_2, \;\ldots, \;\alpha_n$$

be a derivation from $\Gamma \cup \{\varphi\}$ in the system $\mathcal{S}_{\mathcal{L}}$ that contains no application of Generalisation rule involving a variable which occurs free in φ. To complete the induction step, we need to show that $\Gamma \vdash (\varphi \to \alpha_n)$ in the case where α_n was inferred by Generalisation rule, i.e. $\alpha_n = \forall x\, \alpha_i$ where $i < n$. It follows that x is not a free variable of φ, as it is involved in an application of Generalisation rule in the derivation of α_n from $\Gamma \cup \{\varphi\}$. By induction hypothesis, we obtain that $\Gamma \vdash (\varphi \to \alpha_i)$. Let $\theta_1, \theta_2, \ldots, \theta_m$ be a derivation of $(\varphi \to \alpha_i)$ from Γ in $\mathcal{S}_{\mathcal{L}}$. We now construct the following derivation in $\mathcal{S}_{\mathcal{L}}$:

1.	θ_1	$*$
2.	θ_2	$*$
$\vdots$	$\vdots$	$\vdots$
$m.$	$(\varphi \to \alpha_i)$	$*$
$m+1.$	$\forall x\, (\varphi \to \alpha_i)$	Gen m
$m+2.$	$(\forall x\, (\varphi \to \alpha_i) \to (\varphi \to \forall x\, \alpha_i))$	Ax 5
$m+3.$	$(\varphi \to \forall x\, \alpha_i)$	MP $m+1,\ m+2$

Since the variable x is not free in φ, the use of Ax 5 in the derivation above is correct. Thus, we constructed a derivation of $(\varphi \to \alpha_n)$ from Γ, and so $\Gamma \vdash (\varphi \to \alpha_n)$ as desired. This completes the induction step. ■

Example. In $\mathcal{S}_{\mathcal{L}}$ we will use $\exists x$ to abbreviate $\neg\forall x\neg$. Show that for any first-order formulas α and β we have that

$$\vdash (\forall x(\alpha \to \beta) \to (\exists x\alpha \to \exists x\beta)).$$

We will establish this in two stages. First we will show that

$$\{\forall x(\alpha \to \beta), \forall x\neg\beta\} \vdash \forall x\neg\alpha.$$

1.	$\forall x(\alpha \to \beta)$	Assumption
2.	$\forall x\neg\beta$	Assumption
3.	$(\forall x(\alpha \to \beta) \to (\alpha \to \beta))$	Ax 4
4.	$(\alpha \to \beta)$	MP 1, 3
5.	$((\alpha \to \beta) \to (\neg\beta \to \neg\alpha))$	Theorem of $\mathcal{S}_{\mathcal{L}}$
6.	$(\neg\beta \to \neg\alpha)$	MP 4, 5
7.	$(\forall x\neg\beta \to \neg\beta)$	Ax 4
8.	$\neg\beta$	MP 2, 7
9.	$\neg\alpha$	MP 6, 8
10.	$\forall x\neg\alpha$	Gen 9

Since x doe not occur free in $\forall x\neg\beta$, we deduce by the deduction theorem that

$$\forall x(\alpha \to \beta) \vdash (\forall x\neg\beta \to \forall x\neg\alpha).$$

Next we build on this derivation the following derivation.

1.	$\forall x(\alpha \to \beta)$	Assumption
2.	$(\forall x \neg \beta \to \forall x \neg \alpha)$	Derived above
3.	$((\forall x \neg \beta \to \forall x \neg \alpha) \to (\neg \forall x \neg \alpha \to \neg \forall x \neg \beta))$	Theorem of $\mathcal{S}_{\mathcal{L}}$
4.	$(\neg \forall x \neg \alpha \to \neg \forall x \neg \beta)$	MP 1, 3

This shows that

$$\forall x(\alpha \to \beta) \vdash (\neg \forall x \neg \alpha \to \neg \forall x \neg \beta).$$

That is,

$$\forall x(\alpha \to \beta) \vdash (\exists x \alpha \to \exists x \beta).$$

Since x is not free in $\forall x(\alpha \to \beta)$ we can apply the deduction theorem again to get the desired result,

$$\vdash (\forall x(\alpha \to \beta) \to (\exists x \alpha \to \exists x \beta)).$$

♠

4.4 The Completeness Theorem

Let $\mathcal{L}$ be a first-order language. The completeness theorem states that for any set T of $\mathcal{L}$-sentences and any $\mathcal{L}$-sentence ψ, if $T \models \psi$, then $T \vdash \psi$. This means that the axioms and deduction rules of our proof system $\mathcal{S}_{\mathcal{L}}$ are powerful enough to capture truth in first-order logic. As we mentioned earlier, the completeness theorem for countable languages in first-order logic was first proved by the Austrian mathematician and logician Kurt Gödel in 1930. The proof that we will present here is based on one produced by the American Logician Leon Henkin in 1949. The method extends the approach we adopted in establishing the completeness theorem for propositional logic where first we showed that every consistent set of propositional formulas is satisfiable as follows.

- Start with a consistent set Δ of formulas.
- Extend Δ to a complete set of formulas Σ.
- Use Σ to define a truth assignment δ by setting $\delta[p] = 1$ if and only if $\Sigma \vdash p$ for every propositional variable p.
- Then showed that δ satisfies Σ, and so it satisfies its subset Δ as well.
- Thus, Δ is satisfiable.

Definition. A set T of $\mathcal{L}$-sentences is said to be *inconsistent* if there exists an $\mathcal{L}$-sentence θ such that $T \vdash \theta$ and $T \vdash \neg\theta$. And T is *consistent* if it is not inconsistent.

In a similar fashion, towards showing the completeness theorem for first-order logic, we will show first that every consistent set T of $\mathcal{L}$-sentences has a model. Once we have this result we proceed to show the completeness theorem as follows.

Theorem 4.4.1 (Gödel's Completeness Theorem for $\mathcal{S}_{\mathcal{L}}$). *For any set of $\mathcal{L}$-sentences T and any $\mathcal{L}$-sentence ψ,*

$$\text{if } T \models \psi, \text{ then } T \vdash \psi.$$

Proof. Suppose that $T \models \psi$, so every model of T is a model of ψ. It follows that the set $T \cup \{\neg\psi\}$ has no model. By Theorem 4.4.2, we get that the set $T \cup \{\neg\psi\}$ is inconsistent. Thus, there exists an $\mathcal{L}$-formula θ such that $T \cup \{\neg\psi\} \vdash \theta$ and $T \cup \{\neg\psi\} \vdash \neg\theta$. By the deduction theorem, it follows that, $T \vdash (\neg\psi \to \theta)$ and $T \vdash (\neg\psi \to \neg\theta)$.

Observe that the formula $((\neg\psi \to \theta) \to ((\neg\psi \to \neg\theta) \to \psi))$ is a first-order tautology obtained from the propositional tautology $((\neg p \to q) \to ((\neg p \to \neg q) \to p))$ and so it is a theorem of $\mathcal{S}_{\mathcal{L}}$ by Theorem 4.2.1. Using what we already have we now compose the following derivation from T.

1.	$(\neg\psi \to \theta)$	Derived from T
2.	$(\neg\psi \to \neg\theta)$	Derived from T
3.	$((\neg\psi \to \theta) \to ((\neg\psi \to \neg\theta) \to \psi))$	Theorem of $\mathcal{S}_{\mathcal{L}}$
4.	$((\neg\psi \to \neg\theta) \to \psi)$	MP 1, 3
5.	ψ	MP 2, 4

Therefore, $T \vdash \psi$ as desired. ■

We are left with showing that every consistent set of sentences has a model.

Definition. A set T of $\mathcal{L}$-sentences is called *complete* if T is consistent and for each $\mathcal{L}$-sentence φ, either $T \vdash \varphi$ or $T \vdash \neg\varphi$.

Definition. A set Δ of $\mathcal{L}$-sentences has the *witnessing property* if for any $\mathcal{L}$-formula $\varphi(x)$ with free variable x, whenever the sentence $\exists x\, \varphi \in \Delta$, then $\varphi(c) \in \Delta$ for some constant symbol $c \in \mathcal{L}$. Here $\varphi(c) = \varphi(c/x)$ the formula obtained from φ be replacing every free occurrence of x by c.

Theorem 4.4.2. *Every consistent set of $\mathcal{L}$-sentences has a model.*

Proof. Let $\mathcal{L}$ be a first-order countable language and let Σ be a consistent set of $\mathcal{L}$-sentences. We expand the language $\mathcal{L}$ by adding countably many new constant symbols c_i where $i \in \mathbb{N}$ to form a new countable first-order language $\hat{\mathcal{L}}$.

$$\hat{\mathcal{L}} = \mathcal{L} \cup \{c_0, c_1, c_2, c_3, \ldots\}$$

The remaining part of the proof will be split in three stages.

Stage I. We form a set $\hat{\Sigma}$ of $\hat{\mathcal{L}}$-sentences satisfying the following four properties.

1. $\Sigma \subseteq \hat{\Sigma}$.
2. $\hat{\Sigma}$ is consistent.
3. $\hat{\Sigma}$ is complete.
4. $\hat{\Sigma}$ has the witnessing property.

As there are countably many $\hat{\mathcal{L}}$-sentences we enumerate them as

$$\varphi_0, \varphi_1, \varphi_2, \ldots$$

Towards building $\hat{\Sigma}$, we will inductively build a chain

$$\Sigma_0 \subseteq \Sigma_1 \subseteq \Sigma_2 \subseteq \cdots$$

of consistent sets of $\hat{\mathcal{L}}$-sentences. We start by setting $\Sigma_0 = \Sigma$ which is consistent. Now suppose that Σ_n has been constructed and it is consistent. We will next face one of the following two cases.

Case (i). $\Sigma_n \vdash \varphi_n$.
Here, there are two subcases. If φ_n is of the form $\exists x\, \psi(x)$ where $\psi(x)$ is an $\hat{\mathcal{L}}$-formula with free variable x, then put

$$\Sigma_{n+1} = \Sigma_n \cup \{\varphi_n,\ \psi(c_i)\}$$

for the first new constant symbol c_i not appearing in Σ_n and not appearing in φ_n. Otherwise, if φ_n is not of this form, then put

$$\Sigma_{n+1} = \Sigma_n \cup \{\varphi_n\}.$$

Case (ii). $\Sigma_n \nvdash \varphi_n$. Then put

$$\Sigma_{n+1} = \Sigma_n \cup \{\neg\varphi_n\}.$$

One can check (as we did in propositional logic) that the set Σ_{n+1} is still consistent. We now form the desired set

$$\hat{\Sigma} = \bigcup_{n\in\mathbb{N}} \Sigma_n.$$

Let us check that $\hat{\Sigma}$ has the wanted four properties. First, $\Sigma = \Sigma_0 \subseteq \hat{\Sigma}$. Second, $\hat{\Sigma}$ is consistent since every Σ_n is consistent. Third, $\hat{\Sigma}$ is complete. To see this, let θ be any $\hat{\mathcal{L}}$-sentence, so $\theta = \varphi_n$ for some $n \in \mathbb{N}$. By the construction at step $n+1$, we see that either $\varphi_n \in \Sigma_{n+1}$ or $\neg\varphi_n \in \Sigma_{n+1}$. Finally, we show that $\hat{\Sigma}$ has the witnessing property. Towards this end, suppose that $\psi(x)$ is some $\hat{\mathcal{L}}$-formula with free variable x and the sentence $\exists x\, \psi \in \hat{\Sigma}$. Since $\exists x\, \psi$ is an $\hat{\mathcal{L}}$-sentence, there is some $n \in \mathbb{N}$ such that $\varphi_n = \exists x\, \psi$. As $\hat{\Sigma}$ is consistent, it must be that $\neg\varphi_n \notin \hat{\Sigma}$, meaning that Case (i) had to hold at step $n+1$, and thus $\Sigma_{n+1} = \Sigma_n \cup \{\exists x\, \psi,\ \psi(c)\}$ for some new constant symbol $c \in \hat{\mathcal{L}}$. Thus $\psi(c) \in \hat{\Sigma}$ as needed. So $\hat{\Sigma}$ has the witnessing property.

Stage II. We build an $\hat{\mathcal{L}}$-structure $\hat{\mathcal{M}}$.

Let $C = \{c_i \mid i \in \mathbb{N}\}$ be the set of the new constant symbols. We define a relation $\sim$ on C by declaring for any $c_i, c_j \in C$ the following

$$c_i \sim c_j \text{ if and only if the sentence } (c_i = c_j) \in \hat{\Sigma}.$$

One can show that $\sim$ is an equivalence relation on C. Let $\tilde{c_i}$ denote the equivalence class of c_i, that is,

$$\tilde{c_i} = \{c_j \in C \mid c_j \sim c_i\}.$$

We are now ready to define the structure $\hat{\mathcal{M}}$. The underlying set of $\hat{\mathcal{M}}$ is the set $\{\tilde{c_i} \mid i \in \mathbb{N}\}$ of all equivalence classes. We now interpret the symbols of $\hat{\mathcal{L}}$ as follows.

- If $c \in \hat{\mathcal{L}}$ is any constant symbol, then there is some new constant symbol c_i such that the sentence $(c = c_i) \in \hat{\Sigma}$. Declare $c^{\hat{\mathcal{M}}}$ to be $\tilde{c_i}$.

- For any k-ary function symbol f in $\hat{\mathcal{L}}$ and any $\tilde{c}_{i_1}, \tilde{c}_{i_2}, \ldots, \tilde{c}_{i_k}$ elements from $\hat{\mathcal{M}}$ we define
$$f^{\hat{\mathcal{M}}}(\tilde{c}_{i_1}, \tilde{c}_{i_2}, \ldots, \tilde{c}_{i_k}) = \tilde{c}_j$$
where $\tilde{c}_j$ is the unique equivalence class such that the sentence $(f(c_{i_1}, c_{i_2}, \ldots, c_{i_k}) = c_j)$ belongs to $\hat{\Sigma}$.

- For any k-ary relation symbol R in $\hat{\mathcal{L}}$ and any $\tilde{c}_{i_1}, \tilde{c}_{i_2}, \ldots, \tilde{c}_{i_k}$ elements from $\hat{\mathcal{M}}$ we define
$$(\tilde{c}_{i_1}, \tilde{c}_{i_2}, \ldots, \tilde{c}_{i_k}) \in R^{\hat{\mathcal{M}}} \text{ if and only if } R(c_{i_1}, c_{i_2}, \ldots, c_{i_k}) \in \hat{\Sigma}.$$

Stage III. We show that $\hat{\mathcal{M}}$ is a model of $\hat{\Sigma}$.

To show that $\hat{\mathcal{M}} \models \hat{\Sigma}$ we will show by induction on $\hat{\mathcal{L}}$-formulas $\varphi(x_1, x_2, \ldots, x_n)$ that

$$\hat{\mathcal{M}} \models \varphi(\tilde{c}_{i_1}, \tilde{c}_{i_2}, \ldots, \tilde{c}_{i_n}) \text{ if and only if } \varphi(c_{i_1}, c_{i_2}, \ldots, c_{i_n}) \in \hat{\Sigma}. \qquad (\dagger)$$

for any new constant symbols $c_{i_1}, c_{i_2}, \ldots, c_{i_n}$.

For the base case, $\varphi(x_1, \ldots, x_n)$ is an atomic formula, and here the statement (†) holds by definition of the structure $\hat{\mathcal{M}}$. We leave it to the reader to show that φ satisfies (†) when φ is built using a propositional connective from two simpler $\hat{\mathcal{L}}$-formulas satisfying (†).

Now suppose that $\varphi(x_1, \ldots, x_n) = \exists y\, \psi(x_1, \ldots, x_n, y)$ and that $\psi(x_1, \ldots, x_n, y)$ satisfies (†). We need to show that φ satisfies (†). We first show the forward direction of (†).

$$\begin{aligned}
\hat{\mathcal{M}} \models \varphi(\tilde{c}_{i_1}, \tilde{c}_{i_2}, \ldots, \tilde{c}_{i_n}) &\Rightarrow \hat{\mathcal{M}} \models \exists y\, \psi(\tilde{c}_{i_1}, \tilde{c}_{i_2}, \ldots, \tilde{c}_{i_n}, y) \\
&\Rightarrow \text{ there is } \tilde{c} \in \hat{\mathcal{M}} \text{ we have: } \hat{\mathcal{M}} \models \psi(\tilde{c}_{i_1}, \tilde{c}_{i_2}, \ldots, \tilde{c}_{i_n}, \tilde{c}) \\
&\overset{\text{IH}}{\Longrightarrow} \psi(c_{i_1}, c_{i_2}, \ldots, c_{i_n}, c) \in \hat{\Sigma} \\
&\Rightarrow \exists y\, \psi(c_{i_1}, c_{i_2}, \ldots, c_{i_n}, y) \in \hat{\Sigma} \\
&\Rightarrow \varphi(c_{i_1}, c_{i_2}, \ldots, c_{i_n}) \in \hat{\Sigma}.
\end{aligned}$$

For the converse of (†), we suppose that $\varphi(c_{i_1}, \ldots, c_{i_n}) \in \hat{\Sigma}$, and this means that $\exists y\, \psi(c_{i_1}, \ldots, c_{i_n}, y) \in \hat{\Sigma}$. Since $\hat{\Sigma}$ has the witnessing property, there is a new constant symbol c such that $\psi(c_{i_1}, \ldots, c_{i_n}, c) \in \hat{\Sigma}$. By induction hypothesis, we get that $\hat{\mathcal{M}} \models \psi(\tilde{c}_{i_1}, \ldots, \tilde{c}_{i_n}, \tilde{c})$, thus $\hat{\mathcal{M}} \models \exists y \psi(\tilde{c}_{i_1}, \ldots, \tilde{c}_{i_n}, y)$, and so $\hat{\mathcal{M}} \models \varphi(\tilde{c}_{i_1}, \ldots, \tilde{c}_{i_n})$.

It remains to show that φ satisfies (†) when $\varphi(x_1, \ldots, x_n) = \forall y\, \psi(x_1, \ldots, x_n, y)$ for some formula $\psi(x_1, \ldots, x_n, y)$ satisfying (†). This is left as an exercise.

Finally, since $\Sigma \subseteq \hat{\Sigma}$, we also get that $\hat{\mathcal{M}} \models \Sigma$. Now let $\mathcal{M}$ be the 'reduct' of $\hat{M}$ to the language $\mathcal{L}$, that is, $\mathcal{M}$ is an $\mathcal{L}$-structure with the same underlying set as $\hat{\mathcal{M}}$ and the interpretation of a symbol in $\mathcal{L}$ is the same interpretation given to it in $\hat{\mathcal{M}}$. Then $\mathcal{M}$ is a model of Σ as desired. ■

At this stage we have all the tools needed to establish the compactness theorem for first-order logic.

Theorem 4.4.3. *Let Σ be a set of $\mathcal{L}$-sentences of a first-order language $\mathcal{L}$. If every finite subset of Σ has a model, then Σ has a model.*

Proof. Suppose that every finite subset of Σ has a model, i.e. Σ is finitely satisfiable. For the sake of a contradiction, suppose further that Σ does not have a model. By Theorem 4.4.2, Σ must be inconsistent. So there is an $\mathcal{L}$-sentence θ such that $\Sigma \vdash \theta$ and $\Sigma \vdash \neg\theta$. Let $\Delta \subseteq \Sigma$ be the set of all assumptions from Σ used these two derivations. As derivations are of finite length Δ has finitely many sentences. Clearly, $\Delta \vdash \theta$ and $\Delta \vdash \neg\theta$. By soundness theorem we get that $\Delta \models \theta$ and $\Delta \models \neg\theta$. Since Δ is a finite subset of Σ we know that Δ has a model, say $\mathcal{M} \models \Delta$. It follows that $\mathcal{M} \models \theta$ and $\mathcal{M} \models \neg\theta$, a contradiction. Thus Σ must have a model. ■

Corollary 4.4.4. *For any set of $\mathcal{L}$-sentences Σ and any $\mathcal{L}$-sentence ψ, if $\Sigma \models \psi$, then there exists a finite subset $\Delta \subseteq \Sigma$ such that $\Delta \models \psi$.*

Proof. Suppose that $\Sigma \models \psi$ meaning that every model of Σ is also a model of ψ. This implies that $\Sigma \cup \{\neg\psi\}$ had no model. By compactness theorem there is a finite subset $\Delta \subseteq \Sigma \cup \{\neg\psi\}$ that has no model. Let $\Delta_0 = \Delta \setminus \{\neg\psi\}$ and notice that $\Delta_0 \subseteq \Sigma$. Now as $\Delta = \Delta_0 \cup \{\neg\psi\}$ has no model, every model of Δ_0 must satisfy ψ, hence $\Delta_0 \models \psi$ as desired. ■

Theorem 4.4.5. *Suppose that Σ is a set of $\mathcal{L}$-sentences where for every $k \in \mathbb{N}$, there exists $m \in \mathbb{N}$ with $m > k$ and a finite model of Σ of cardinality m (i.e. Σ has models of arbitrarily large finite cardinality). Then Σ has an infinite model.*

Proof. For $n \geq 2$, let φ_n be the $\mathcal{L}$-sentence expressing the existence of at least n distinct elements. For example, φ_3 is $\exists x \exists y \exists z\, (x \neq y \wedge x \neq z \wedge y \neq z)$. Now consider the set of $\mathcal{L}$-sentences

$$\Gamma = \Sigma \cup \{\varphi_n \mid n \geq 2\}.$$

We claim that every finite subset of Γ has a model. Take a finite subset $\Delta \subseteq \Gamma$. Then $\Delta = \Delta_0 \cup \{\varphi_{i_1}, \varphi_{i_2}, \ldots, \varphi_{i_k}\}$ where $\Delta_0 \subseteq \Sigma$ and $2 \leq i_1 < i_2 < \cdots < i_k$. By our hypothesis we know that Σ has a model $\mathcal{A}$ of cardinality larger than i_k. Thus, $\mathcal{A} \models \Delta$. By the compactness theorem, we obtain that Γ has a model $\mathcal{M}$. So $\mathcal{M} \models \Sigma$ and $\mathcal{M} \models \varphi_n$ for every $n \geq 2$ implying that $\mathcal{M}$ has infinitely many elements in its domain. Hence $\mathcal{M}$ is an infinite model of Σ. ■

Chapter 5

Set Theory

The German mathematician Georg Cantor is considered to be the founder of modern set theory. After paradoxes started to float on the surface of naive set theory such as Russell's paradox, Cantor's paradox, and Burali-Forti paradox, mathematicians proposed several axiomatic systems in the early twentieth century to study sets. One of the best known and most studied systems is the Zermelo-Fraenkel Set Theory.

5.1 Zermelo-Fraenkel Set Theory

Our aim is to axiomatize our understanding of sets. We want to see how set theory can be considered as a first-order theory. What first-order language shall we use to describe sets? What axioms shall we choose to reflect the behaviour of sets as we know in everyday mathematics?

All the objects we deal with will be sets. In particular, members of sets will also be sets. In some cases we will deal with collection of sets which are too large to be sets, such a collection is called a *class*. As usual, we will write $x \in y$ when a set x is a member of a set y.

We will express these axioms, called *Zermelo-Fraenkel Axioms* (ZF), in the first-order language of set theory $\mathcal{L} = \{ \in \}$ where $\in$ is a binary relation symbol. Of course $\mathcal{L}$ contains the equality symbol $=$ as well. An $\mathcal{L}$-structure would be $\mathcal{V} = (V, \in)$ where $\mathcal{V}$ is called the universe of sets. Any element of V is called a set and the domain V itself is a set in the naive sense. The variables $x, y, z, \ldots$ represent elements of $\mathcal{V}$ so they stand for sets. Two elements in V are related by the relation $\in$ if the first is a "member" of the second.

Empty Set Axiom. There exists an empty set.

$$\exists x \, \forall y \, \neg(y \in x)$$

Extensionality Axiom. Two sets are equal if and only if they have the same members.

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

Pairing Axiom. If x and y are sets, then there is a set containing exactly them. That is, whenever x and y are sets, then the unordered pair $\{x, y\}$ is also a set.

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

Union Axiom. If x is a set, then the collection of all members of members of x is a set (denoted by $\bigcup x$).

$$\forall x\,\exists y\,\forall z\,(z \in y \leftrightarrow \exists w(z \in w \land w \in x))$$

Separation Axiom (or Comprehension Axiom). If x is a set, then the collection of all members of x satisfying some first-order property is also a set.

For any $\mathcal{L}$-formula $\varphi(z)$ in one free variable z the following $\mathcal{L}$-sentence is an axiom.

$$\forall x\,\exists y\,\forall z\,(z \in y \leftrightarrow (z \in x \land \varphi(z)))$$

Power Set Axiom. The power set (collection of all subsets) of any set exists.

$$\forall x\,\exists y\,\forall z\,(z \in y \leftrightarrow \forall w(w \in z \to w \in x))$$

Replacement Axiom. If x is a set and f is a function, then the collection of all images of members of x under f is a set.

For any $\mathcal{L}$-formula $\varphi(u, z)$ in free variables u, z the following $\mathcal{L}$-sentence is an axiom.

$$\forall x\Big(\forall u\forall z\forall w\big((u \in x\land\varphi(u,z)\land\varphi(u,w)) \to z = w\big) \to \exists y\forall z\big(z \in y \leftrightarrow \exists u\,(u \in x\land\varphi(u,z))\big)\Big)$$

Infinity Axiom. There is an infinite set.

$$\exists x\,\big(\emptyset \in x \land \forall y\,(y \in x \to y \cup \{y\} \in x)\big)$$

Foundation Axiom. Every nonempty set x contains an element y such that $x \cap y = \emptyset$.

$$\forall x\,\big(\exists u\,(u \in x) \to \exists y\,(y \in x \land \neg\exists z\,(z \in y \land z \in x))\big)$$

All the axioms above are the axioms of Zermelo-Fraenkel Set Theory (**ZF**). The **ZF** axioms without the Axiom of Replacement is called Zermelo Set Theory (**Z**). Adding another axiom, called the Axiom of Choice, to **ZF** axioms will give us *Zermelo-Fraenkel Set Theory with the Axiom of Choice* abbreviated as **ZFC**.

Axiom of Choice (AC). For any set x whose members are all nonempty sets, there exists a function f defined on x, called a *choice function*, that maps each member y of x to a member of y.

$$\forall x\,(\neg(\emptyset \in x) \to \exists f(\mathrm{Func}(f) \land \mathrm{dom}(f) = x \land \mathrm{Im}(f) = \bigcup x \land \forall y(y \in x \to f(y) \in y)))$$

Remark. A function is a set f of ordered pairs such that whenever $(x, y) \in f$ and $(x, z) \in f$, then $y = z$. We write $f(x) = y$ when the pair $(x, y) \in f$. The domain of a function f is the set $\mathrm{dom}(f) = \{x \mid \exists y\, (x, y) \in f\}$ and the image of f is the set $\mathrm{Im}(f) = \{y \mid \exists x\, (x, y) \in f\}$.
ZF proves that a choice function exists for finite sets, however, for certain infinite sets we need **AC**.

Here are immediate consequences of **ZF** axioms.

- The empty set is unique. It is denoted by $\emptyset$.
  By axioms of Empty Set and Extensionality.

- If x is a set, then $\{x\}$ is also a set.
  By axiom of Pairing.

- If x and y are sets, then so is $x \cup y$.
  By axioms of Pairing and Union.

- If x and y are sets, then so is $x \cap y$.
  By axiom of Separation.

- If x is a set, then $x \cup \{x\}$ is a set, called the *successor* of x.

- If x and y are sets, we define the ordered pair (x, y) to be $\{\{x\}, \{x, y\}\}$, and it is called *Kuratowski pair.*

- If x and y are sets, then the ordered pair (x, y) is a set.
  By three applications of the Pairing axiom.

- Let x, y, u, v be sets. Then $(x, y) = (u, v)$ if and only if $x = u$ and $y = v$.

Lemma 5.1.1. *If x and y are sets, then the collection of all ordered pairs (u, v) where $u \in x$ and $v \in y$ is a set called the cartesian product of x and y and is denoted by $x \times y$.*

Proof. Let x and y be sets (members of a universe of sets $\mathcal{V}$). By Pairing axiom, $\{x, y\}$ is a set. By Union axiom, $x \cup y$ is a set. By Power Set axiom applied twice, $\mathcal{P}(\mathcal{P}(x \cup y))$ is a set. So if $u \in x$ and $v \in y$, then $\{\{u\}, \{u, v\}\}$, namely the ordered pair (u, v), is a member of $\mathcal{P}(\mathcal{P}(x \cup y))$. Finally, by Separation axiom the collection of members of $\mathcal{P}(\mathcal{P}(x \cup y))$ of the form (u, v) where $u \in x$ and $v \in y$ is a set. ■

The Axiom of Foundation gives the following result.

Lemma 5.1.2. *No set is a member of itself.*

Proof. Let x be a set. By Pairing axiom, $\{x\}$ is a set. By Foundation axiom, there is $y \in \{x\}$ such that $y \cap \{x\} = \emptyset$. But $y = x$, and so $x \cap \{x\} = \emptyset$, and hence $x \notin x$. ■

Theorem 5.1.3. *The following are equivalent in the theory **ZF**.*

*(i) Axiom of Choice (**AC**).*

*(ii) Well-Ordering Principle (**WOP**): For any set x there is a well-ordering of x.*

(iii) Zorn's Lemma: Any nonempty partially ordered set in which every chain (totally ordered subset) has an upper bound contains a maximal element. (See next section.)

(iv) Tychonoff's Theorem: The product of any family of compact topological spaces is compact with respect to the product topology.

(v) Every vector space has a basis.

Let us prove one of the implications in the theorem above.

Theorem 5.1.4. *In **ZF**, the Well-Ordering Principle implies the Axiom of Choice.*

Proof. Let x be any set where every one of its members is nonempty. By Union Axiom, let $X = \cup x$, so X is the set of all members of members of x. By **WOP**, there is a well-ordering relation on X. Let z be any member of x. Then every member of z is in X, and so z is a nonempty subset of X. Since X is well-ordered and z is a nonempty subset of X, there is a least element s of z. Consider the function f which maps every $z \in x$ to the least element s of z when z is seen as a subset of the well-ordered set X. So $f : x \to X$ where for any $z \in x$ we have that $f(z)$ is the least element of z. Finally, it remains to show that this function, i.e. the collection of all such ordered pairs (z, s), is a set. Towards this we need to apply the Separation Axiom to show that the function f is a subset of the cartesian product $x \times X$. ■

5.2 Well-Orderings

Definition. A *partial order* is a binary relation $<$ on a set X which is irreflexive and transitive. That is,

- For all $x \in X$ we have $(x \not< x)$. (Irreflexivity)
- For all $x, y, z \in X$, if $x < y$ and $y < z$, then $x < z$. (Transitivity)

We read $x < y$ as 'x precedes y' or as 'x is less than y'. We call the pair $(X, <)$ a *partially ordered set* or *poset.*

Exercise. Let $<$ be a partial order on X. Define a new relation $\leq$ on X be setting $x \leq y$ if and only if $x < y$ or $x = y$ for any $x, y \in X$. Show that the new relation $\leq$ is reflexive, antisymmetric, and transitive.

Definition. Let $(X, <)$ be a partially ordered set. An element $a \in X$ is called *minimal* if there is no element $x \in X$ such that $x < a$. An element $b \in X$ is called *maximal* if there is no element $x \in X$ such that $b < x$.

Definition. Let $(X, <)$ be a partially ordered set. Two elements $x, y \in X$ are called *comparable* if $x < y$ or $x = y$ or $y < x$, otherwise, they are called *incomparable.*

A poset where every pair of elements are comparable has a special name.

Definition. A *total* (or *linear*) *order* is a partial order $<$ on a set X which further satisfies that for all $x, y \in X$ either $x < y$ or $x = y$ or $y < x$. In this case, we say that $(X, <)$ is a *totally* (or *linearly*) *ordered set.*

Example.

- Let $X = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Define $m < n$ if and only if $m \neq n$ and $m \mid n$ for every $m, n \in X$. Then $(X, <)$ is a poset which is not a total order. Let $<$ be the usual order relation of the integers, then $(X, <)$ is a totally ordered set. The pair 2 and 6 are comparable while the pair 4 and 6 are incomparable. The elements $8, 9, 10$ are maximal elements. The elements $2, 3, 5$ are minimal. The element 7 is both minimal and maximal. The element 4 is neither minimal nor maximal.
- Let $A = \{a, b, c\}$ and let $S = \mathcal{P}(A)$. Define $x < y$ if and only if $x \subsetneq y$ for every $x, y \in S$. Then $(S, <)$ is a poset which is not a total order. ♠

Definition. A *well-order* is a total order $(X, <)$ where every nonempty subset of X has a least element, that is, for any subset $Y \subseteq X$, if $Y \neq \emptyset$, then there is $y_0 \in Y$ such that for all $y \in Y (y_0 \leq y)$.

Example.

- $(\mathbb{N}, <)$ is a well-ordered set. Let's use $\longmapsto$ to denote $(\mathbb{N}, <)$.

- $(\mathbb{N}, <)$ followed by a point i.e. $\longmapsto \cdot$ is well ordered.

- Similarly, $\longmapsto \cdot\,\cdot$ and $\longmapsto \cdot\,\cdot\,\cdot$ are well-ordered. Also, $(\mathbb{N}, <)$ followed by another $(\mathbb{N}, <)$ i.e. $\longmapsto\longmapsto$ is well-ordered, and so is $\longmapsto\longmapsto \cdot$ and so on.

- $(\mathbb{Z}, <)$ is not a well-ordered set as $\mathbb{Z}$ itself is a subset with no least element.

- $([0, 1], <)$ is not a well-ordered set since the subset $(0, 1)$ has no least element, also the subset $\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\}$ has no least element, and many others. ♠

Lemma 5.2.1. *Any subset Y of a well-ordered set $(X, <)$ is well-ordered by the restriction of $<$ to Y. The restriction $<^Y$ is $<^X \cap (Y \times Y)$.*

Lemma 5.2.2. *Any finite totally ordered set is well-ordered.*

Proof. Suppose that $(X, \prec)$ is a totally ordered set and X is finite. Towards a contradiction, suppose that $(X, \prec)$ is not well-ordered. Then there exists a nonempty subset $Y \subseteq X$ which has no least element. As Y is nonempty, pick an element $y_0 \in Y$. As Y has no least element, there is $y_1 \in Y$ with $y_1 \prec y_0$. Again, as y_1 is not a least element of Y, there is $y_2 \in Y$ with $y_2 \prec y_1$. Continuing in this fashion we obtain a sequence $\cdots \prec y_3 \prec y_2 \prec y_1 \prec y_0$ of elements of Y. We claim that $y_j \neq y_i$ whenever $j > i$. To see this notice as $y_j \prec \cdots \prec y_{i+1} \prec y_i$ it follows by transitivity of $\prec$ that $y_j \prec y_i$, and by irreflexivity we must have that $y_j \neq y_i$. Therefore, there are infinitely many distinct elements in X, but X is finite, so we get a contradiction. Thus, $(X, \prec)$ is well-ordered. ■

Definition. Let $(X, <)$ and $(Y, \prec)$ be partially-ordered sets. An *isomorphism* between them is a bijection $f : X \to Y$ such that for all $x, z \in X$ we have that

$$x < z \text{ if and only if } f(x) \prec f(z).$$

If there is at least one isomorphism between $(X, <)$ and $(Y, \prec)$ we say they are *isomorphic* and we write $(X, <) \cong (Y, \prec)$.

Example.

- Let $A = \{1, 4, 9\}$ and $B = \{2, 3, 7\}$. Then $(A, <) \cong (B, <)$ via the bijection $1 \mapsto 2$, $4 \mapsto 3$, $9 \mapsto 7$.

- $([0, 1], <) \cong ([0, 2], <)$ via the bijection $x \mapsto 2x$ for every $x \in [0, 1]$.

- $([0, 1], <) \ncong ([0, 1), <)$. One needs to show here that it is impossible for an isomorphism to exist. ♠

Theorem 5.2.3. *Let $(X, <)$ and $(Y, \prec)$ be well-ordered sets. If $(X, <) \cong (Y, \prec)$, then the isomorphism between them is unique.*

Proof. Let $f : X \to Y$ and $g : X \to Y$ be two isomorphisms between $(X, <)$ and $(Y, \prec)$. Towards a contradiction assume that $f \neq g$. Consider the set of points where the functions f and g disagree:

$$A = \{x \in X \mid f(x) \neq g(x)\}.$$

Now $A \subseteq X$ and as $f \neq g$ we get $A \neq \emptyset$. Since $(X, <)$ is well-ordered, the subset A has a least element, say $a \in A$. So a is the smallest element of X where $f(a) \neq g(a)$ meaning that f and g agree at every point less than a. Let $f(a) = v$ and $g(a) = w$. Since $v, w \in Y$, $v \neq w$, and $(Y, <)$ is totally ordered either $v < w$ or $w < v$. Without loss of generality, suppose that $v < w$. Since g is surjective there exists $x \in X$ such that $g(x) = v$. Since $(X, <)$ is totally ordered either $x < a$, $x = a$, or $a < x$. If $x < a$, then f and g agree on x and so $f(x) = g(x) = v$, but then $f(x) = f(a)$ contradicting the injectivity of f. If $x = a$, then g assigns to a two different images v and w contradicting that g is a function. We are left with the case $a < x$. If $a < x$, then $g(a) < g(x)$ since g is an isomorphism, hence $w < v$. We also know that $v < w$, and so we get $w < w$ by transitivity of the order. This contradicts that $<$ is irreflexive. Therefore, it must be that $f(x) = g(x)$ for every $x \in X$, i.e. $f = g$. ■

Definition. Let $(X, <)$ be a totally ordered set.

- An *initial segment* of $(X, <)$ is a subset $I \subseteq X$ such that for any $a \in I$ and $x \in X$, if $x < a$, then $x \in I$. We also say that I is *closed downwards* for this.
- Let $x \in X$. Then the *initial segment of* $(X, <)$ *determined by* x is

$$I_x = \{y \in X \mid y < x\}.$$

Exercise. Show that I_x is an initial segment of $(X, <)$, i.e. show that I_x is closed downwards.

Lemma 5.2.4. *Let I be a proper initial segment of a well-ordered set $(X, <)$. Then there is a unique $x \in X$ such that $I = I_x$.*

Proof. Let $(X, <)$ be a well-ordered set and let I be a proper initial segment of X, so $I \neq X$. Consider the set $A = X \setminus I$ which is nonempty since I is a proper subset of X. By well-ordering, A contains a least element call it a. We claim that $I = I_a$. Choose an arbitrary element $y \in I$. Clearly, $y \neq a$ because $y \in I$ and a belongs to the complement of I. Suppose for contradiction that $a < y$. But then as I is closed downwards we must have $a \in I$, a contradiction. By total ordering, we must have that $y < a$ and so $y \in I_a$ showing that $I \subseteq I_a$. Next take an element $x \in I_a$, and so $x < a$. If $x \notin I$, then $x \in A$ but this contradicts that a is the least element of A, thus $x \in I$ and so $I_a \subseteq I$. Therefore, $I = I_a$. We leave to the reader to show that a is unique. ■

Theorem 5.2.5. *Suppose that $(X,<)$ and $(Y,\prec)$ are well-ordered sets. Then exactly one of the following holds:*

(i) $(X,<)\cong(Y,\prec)$

(ii) $(X,<)\cong(I_y,\prec)$ *where I_y is a proper initial segment of $(Y,\prec)$.*

(iii) $(Y,\prec)\cong(I_x,<)$ *where I_x is a proper initial segment of $(X,<)$.*

Moreover in each of these cases, the isomorphism is unique.

Proof. Start with well-ordered sets $(X,<)$ and $(Y,\prec)$. We will define a partial function f from X to Y, so a function $f:A\to Y$ where $A\subseteq X$, as follows: for $x\in X$ and $y\in Y$ we declare $f(x)=y$ if $I_x\cong I_y$, in other words, if the initial segment of $(X,<)$ determined by x is isomorphic to the initial segment of $(Y,\prec)$ determined by y. We leave to the reader to verify the following facts:

- f is well-defined.
- The domain I of f is an initial segment of $(X,<)$.
- The image J of f is an initial segment of $(Y,\prec)$.
- $f:I\to J$ is an isomorphism between $(I,<)$ and $(J,\prec)$.

We next claim that $I=X$ or $J=Y$. Suppose it is not the case, then both I and J are proper initial segments of X and Y, respectively. It follows by Lemma 5.2.4 that I is the set of predecessors of some $a\in X$ and J is the set of predecessors of some $b\in Y$, that is, $I=I_a$ and $J=I_b$. Since $f:I\to J$ is an isomorphism we get by the definition of f that $f(a)=b$ which implies that $a\in\operatorname{dom}(f)=I=I_a$. This contradicts that $a\notin I_a$. Therefore, it must be $I=X$ or $J=Y$. If $I=X$ and $J=Y$, then $(X,<)\cong(Y,\prec)$. If $I=X$ and $J\neq Y$, then $(X,<)$ is isomorphic to a proper initial segment J of Y. And finally, if $I\neq X$ and $J=Y$, then $(Y,\prec)$ is isomorphic to a proper initial segment I of X.

The uniqueness of f follows from Theorem 5.2.3. ■

5.3 Ordinals and Cardinals

The isomorphism relation forms an equivalence relation on the 'class' of all well-ordered sets. Georg Cantor thought of an ordinal as an equivalence class of well-ordered sets under the relation of being isomorphic. John von Neumann elegantly chose a special representative from such an equivalence class to be an ordinal.

Definition. A set X is called *transitive* if whenever $z \in y$ and $y \in X$, then $z \in X$.

A set is transitive if and only if every member of a member is a member if and only if every member is a subset. The set $\{\, \emptyset,\ \{\emptyset\},\ \{\emptyset, \{\emptyset\}\}\,\}$ is transitive.

Definition. A set X is an *ordinal* if X is transitive and $(X, \in)$ is a well-ordered set.

The membership relation $\in$ on a set X is the set $\{(a, b) \in X \times X \mid a \in b\}$. When $(X, \in)$ is a partially ordered set, then the membership relation $\in$ orders the members of X. So for any $a \in X$ and $b \in X$, if $a \in b$, then we think of this as the element a precedes b and in this case we may write $a < b$ in place of $a \in b$. We also write $a \leq b$ when $a = b$ or $a \in b$.

The first examples of ordinals are the natural numbers. We may view a natural number as an ordinal as follows.

- Define 0 to be $\emptyset$, and
- define $n + 1$ to be $n \cup \{n\}$.

So 1 is $\{\emptyset\}$, and 2 is $\{\emptyset, \{\emptyset\}\}$ and 3 is $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, and so on. All these sets are ordinals. For instance to see that the set 3 is well-ordered by the membership relation $\in$ observe that

$$\emptyset \in \{\emptyset\} \in \{\emptyset, \{\emptyset\}\} \text{ and } \emptyset \in \{\emptyset, \{\emptyset\}\}.$$

So we may write $\emptyset < \{\emptyset\} < \{\emptyset, \{\emptyset\}\}$ or simply $0 < 1 < 2$. Hence the pair $(3, \in)$ is a well-ordered set. The set 3 is also transitive and so 3 is an ordinal.

Let α and β be ordinals. We will write $\alpha \cong \beta$ when the well-ordered sets $(\alpha, \in)$ and $(\beta, \in)$ are isomorphic.

Theorem 5.3.1. *The following hold.*

(i) Any member of an ordinal is an ordinal.

(ii) Any member α of an ordinal β is the set of predecessors of α in β, i.e. $\alpha = \{x \in \beta \mid x \in \alpha\}$.

(iii) For any ordinals α and β, if $\alpha \cong \beta$, then $\alpha = \beta$.

(iv) For any ordinals α and β, exactly one of the following holds:
$\alpha = \beta$ or $\alpha \in \beta$ or $\beta \in \alpha$.

(v) Any nonempty set of ordinals has a least element with respect to the membership relation $\in$.

(vi) Any transitive set of ordinals is an ordinal.

Proof. (i) Let α be an ordinal and let $x \in \alpha$. Since α is transitive, we get $x \subseteq \alpha$. As $(\alpha, \in)$ is well-ordered and $x \subseteq \alpha$ the restriction of the relation $\in$ to x is a well-ordering of x and so $(x, \in)$ is a well-ordered set. Next we need to show that x is a transitive set. So suppose that $z \in y$ and $y \in x$. Since $x \in \alpha$ and α is a transitive set, $y \in \alpha$. But $z \in y$, and again by transitivity of α, we must have $z \in \alpha$. We now have that x, y, z are members of α. But $\in$ is a transitive relation on the members of α, and since $z \in y$ and $y \in x$ we get that $z \in x$, hence x is a transitive set. Thus x is an ordinal.

(ii) Let β be an ordinal and let $\alpha \in \beta$. Consider the set $A = \{x \in \beta \mid x \in \alpha\}$ of all predecessors of α in β. Clearly, $A \subseteq \alpha$. Now let $x \in \alpha$. Since β is transitive and $\alpha \in \beta$, it follows that $x \in \beta$, and so $x \in A$. Thus $\alpha \subseteq A$. This shows that $\alpha = A$.

(iii) Suppose that α and β are ordinal in which $\alpha \cong \beta$. Let $f : \alpha \to \beta$ be the unique isomorphism. We claim that $f(x) = x$ for all $x \in \alpha$ i.e. f is the identity map. Towards a contradiction, assume that f is not the identity map and let $A = \{x \in \alpha \mid f(x) \neq x\}$. Since A is a nonempty subset of a well-ordered set α it has a least element, say $a \in A$. So $f(x) = x$ for all $x \in a$. Let $f(a) = b$ where $a \neq b$. Using (ii) and that f is a bijection preserving the order $\in$ and also it is the identity map on members of a we get the following:

$$\begin{aligned} f(a) = b &= \{y \in \beta \mid y \in b\} \\ &= \{f(x) \mid x \in \alpha \wedge x \in a\} \\ &= \{x \mid x \in \alpha \wedge x \in a\} \\ &= \{x \in \alpha \mid x \in a\} = a. \end{aligned}$$

Thus $b = a$, a contradiction. Therefore $f : \alpha \to \beta$ is a bijection and $f(x) = x$ for all $x \in \alpha$. It follows that

$$\beta = f(\alpha) = \{f(x) \mid x \in \alpha\} = \{x \mid x \in \alpha\} = \alpha$$

as desired.

(iv) Let α and β be ordinals. By Theorem 5.2.5 we have three cases. First case: α and β are isomorphic and so by (iii) we get that $\alpha = \beta$. Second case: α is isomorphic to a proper initial segment of β which by *(ii)* must be a member y of β which itself is an ordinal. So $\alpha \cong y$ and $y \in \beta$. By *(iii)*, we get $\alpha = y$ and so $\alpha \in \beta$. Third case: β is isomorphic to a proper initial segment of α, but then in a similar fashion we get that $\beta \in \alpha$.

(v) Let S be a nonempty set of ordinals. Choose an ordinal $\alpha \in S$. If $\alpha \cap S = \emptyset$, then α is a $\in$-least element of S. Otherwise, $\alpha \cap S$ is a nonempty subset of α and so has a $\in$-least element, say β. Then β is a $\in$-least element of S.

(vi) Let X be a transitive set of ordinals. By (iv) and (v) we get that $(X, \in)$ is a well-ordering. So X is an ordinal. ∎

Theorem 5.3.2. *Any well-ordered set is isomorphic to a unique ordinal* $(\alpha, \in)$.

Proof. Let $(X, <)$ be a well-ordered set. Let $a \in X$ and recall that I_a is the initial segment of X determined by a. If $(I_a, <) \cong (\alpha, \in)$ for some ordinal α, then such α is unique by Theorem 5.3.1(iii). Let $A \subseteq X$ be the set of all elements a of X such that $(I_a, <) \cong (\alpha, \in)$ for some ordinal α. We now define a function f on A where for any $a \in A$ we declare $f(a) = \alpha$ where α is the unique ordinal such that $(I_a, <) \cong (\alpha, \in)$. Let S be the image of f, that is, $S = \{f(a) \mid a \in A\}$ and notice that S is a set of ordinals. One can show that

- A is an initial segment of X.
- S is a transitive set and so S is an ordinal itself by Theorem 5.3.1(vi).
- $f : (A, <) \to (S, \in)$ is an isomorphism.

Now for the sake of contradiction suppose that A is a proper initial segment, and so $A \neq X$. Then $A = I_x$ for some $x \in X$. Thus f is an isomorphism between I_x and the ordinal S, and so $x \in A$. But $A = I_x = \{y \in X \mid y < x\}$ and so $x < x$ contradicting that $<$ is irreflexive. Therefore, it must be that $A = X$, and so $(X, <)$ is isomorphic to the ordinal S. Finally, by Theorem 5.3.1(iii) such an ordinal is unique. ∎

Definition. Let α be an ordinal. The *successor* of α is the set $\alpha \cup \{\alpha\}$ and we denote it by $\alpha + 1$.

Lemma 5.3.3.

(i) If α is an ordinal, then its successor $\alpha + 1$ is an ordinal as well.

(ii) If S is a set of ordinals, then $\bigcup_{\alpha \in S} \alpha$ is an ordinal.

The membership relation $\in$ is a well-ordering relation on the class of all ordinals, that is, it is a total order where every nonempty class of ordinals has a least element with respect to the relation $\in$. Let's check this. First, $\alpha \notin \alpha$ for any ordinal α as no set is a member of it self, so $\in$ is an irreflexive relation. Second, suppose that

α, β, γ are ordinals where $\alpha \in \beta$ and $\beta \in \gamma$. Since is γ is a transitive set, we get that $\alpha \in \gamma$, so $\in$ is a transitive relation on the class of all ordinals. Third, by Theorem 5.3.1(iv) we get that $\in$ is a total relation on the class of ordinals. Thus, $\in$ is a total order on the ordinals. It remains to show it is also a well-order.

Notation. When α and β are ordinals we write $\alpha < \beta$ in place of $\alpha \in \beta$. Similary, we write $\alpha \leq \beta$ for $\alpha \in \beta$ or $\alpha = \beta$.

Definition. A *successor ordinal* is an ordinal β such that $\beta \neq 0$ and there exists an ordinal α such that $\beta = \alpha + 1$.

Definition. A *natural number* is an ordinal α such that for every $\beta \leq \alpha$, either $\beta = 0$ or β is a successor ordinal.

Observe that the sets $0, 1, 2, \ldots$ as were defined above are natural numbers.

Lemma 5.3.4. *There is a set which contains precisely all the natural numbers. We call this set ω. Moreover, ω is an ordinal.*

Proof. Let y be a set as in the Axiom of infinity. So $0 \in y$ and if $x \in y$, then $x + 1 \in y$. We claim that every natural number is a member of y. For the sake of contradiction, suppose not. Let α be a natural number which is not in y. By Theorem 5.3.1(v) we may take α to be the least natural number not in y. Clearly, $\alpha \neq 0$ and so α must be a successor ordinal. Hence there is an ordinal x such that $\alpha = x + 1$. Since $x < \alpha$ (i.e. $x \in \alpha$) and by the choice of α, the ordinal $x \in y$. But then by the property of y, we get that $x + 1 \in y$ and so $\alpha \in y$, a contradiction. Thus y contains all natural numbers. Finally, by the Separation Axiom, the collection of all members of y satisfying the property of being a natural number is a set, denoted by ω.

As any member of natural number is a natural number, the set ω is a transitive set. Thus by Theorem 5.3.1(vi) we get that ω is itself an ordinal. ■

Observe that ω is not a successor ordinal since if so then ω will be the successor of a natural number and so a natural number itself, and so it will be a member of itself contradicting that $\in$ is irreflexive. Nonzero ordinals which are not successor ordinals are called *limit ordinals*. The ordinal ω is the least limit ordinal. Here is the beginning of the list of ordinals:

$$0,\ 1,\ 2,\ 3,\ \ldots,\ \omega,\ \omega+1,\ \omega+2,\ \ldots,\ \omega+\omega,\ \ldots,\ \omega^2\ \ldots,\ \omega^3,\ \ldots,\ \omega^\omega,\ \ldots,\ \omega^{\omega^\omega},\ \ldots\ldots$$

We now introduce a special type of ordinals called cardinals.

Definition. For any sets X and Y, we say that X is *equinumerous* to Y and write $X \sim Y$ if there exists at least one bijective function $f : X \to Y$.

Theorem 5.3.5 (Schröder–Bernstein Theorem)**.** *Let X and Y be sets. If there exists an injective function from X to Y, and an injective function from Y to X, then X and Y are equinumerous.*

The reader needs to check that $\sim$ is an equivalence relation on sets. An equivalence class of this relation contains all sets of the same 'size'. For example, $\omega \sim \omega + 1$, so both ω and $\omega + 1$ belong to the same equivalence class. We aim to single out one special member from every such equivalence class.

Working in **ZFC**, by the well-ordering principle we know that every set X admits a well-ordering $\prec$. By Theorem 5.3.2, there exists an ordinal α such that $(X, \prec) \cong (\alpha, \in)$, and therefore $X \sim \alpha$. Thus every set is equinumerous to some ordinal.

Definition. Let X be any set. We define the *cardinality* of X, denoted by $|X|$, to be the least ordinal α that is equinumerous to X.

So one may think of cardinality as a mapping which assigns to every set an ordinal. Moreover, since the identity mapping is a bijective function we have that $|\alpha| \leq \alpha$ for any ordinal α. Those ordinals which match their own cardinality are special.

Definition. A *cardinal* is an ordinal κ such that $|\kappa| = \kappa$.

Lemma 5.3.6. *Every natural number is a cardinal.*

Proof. We will show by induction on the natural numbers that for any natural number n, there is no injective function from n into some m where $m < n$. Clearly, this holds for 0 as there is no ordinal less than 0. Now let $n \in \omega$ and assume that the property holds for n. We need to show it holds for $n+1 = n \cup \{n\}$. Suppose not, hence there is some natural number $k < n+1$ and an injective function $f : n+1 \to k$. Since k cannot be 0, it must be a successor ordinal and so $k = m+1$ for some natural number m. Consider the following function $h : n \to m$ defined as follows for every $i \in n$:

$$h(i) = \begin{cases} f(i) & \text{if } f(i) < m; \\ f(n) & \text{if } f(i) = m. \end{cases}$$

One can check that h is an injective function from n to m, but $m < k$ and $k < n+1$ and so $m < n$, contradicting the induction hypothesis. Thus, there is no injective function from $n+1$ into a smaller natural number. It follows that there is no bijection from a natural number n into a smaller natural number m, and hence the smallest ordinal equinumerous to n is n itself, i.e. $|n| = n$. Therefore, every natural number is a cardinal. ■

Corollary 5.3.7. *The ordinal ω is a cardinal. (It is the first infinite cardinal.)*

Proof. We need to show that cardinality of ω is ω, in symbols, $|\omega| = \omega$. For the sake of contradiction, suppose that there is a bijection f from ω to some natural number

n. Since $n+1 \in \omega$ and so $n+1 \subseteq \omega$, the restriction of f to $n+1$ induces an injective function from $n+1$ to n contradicting the previous lemma. Thus there exists no bijection from ω to a smaller ordinal, and so the least ordinal equinumerous to ω is ω itself and so $|\omega| = \omega$ showing that ω is a cardinal. ■

Remark. When the ordinal ω is viewed as a cardinal we denote it by $\aleph_0$. The first cardinal that is bigger than $\aleph_0$ is called $\aleph_1$.

The theory of **ZFC** is incomplete, that is, there is a sentence in the language of set theory which is not a logical consequence of **ZFC** nor its negation is. An example of such a sentence is the *Continuum Hypothesis* (**CH**) which was proposed by Georg Cantor in 1878 and it says that the cardinality of the power set of $\aleph_0$ is $\aleph_1$, in symbols, **CH** states that $|\mathcal{P}(\aleph_0)| = \aleph_1$ or $2^{\aleph_0} = \aleph_1$. Kurt Gödel proved that $\mathbf{ZFC} \not\models \neg\mathbf{CH}$ (there is a model of **ZFC** which satisfies **CH**), and Paul Cohen introduced the technique of *forcing* to prove that $\mathbf{ZFC} \not\models \mathbf{CH}$ (there is a model of **ZFC** which satisfies the negation of **CH**). Consequently, we say that **CH** is independent of **ZFC**. Cohen received the Fields medal in 1966 for his work on the continuum hypothesis.