DANIEL GONZALEZ CEDRE

DISCRETE MATHEMATICS

UNIVERSITY OF NOTRE DAME

January 28, 2024

These notes are intended for students of CSE 20110 Discrete Mathematics at the University of Notre Dame.

https://daniel-gonzalez-cedre.github.io

Contents

Notation

SYNTAX		SEMANTICS
$\top$	*"True."*	A *true* sentence; a tautology.
$\bot$	*"False."*	A *false* sentence; a contradiction.
$x \coloneqq y$	*"x is, by definition, y."*	The name x has been assigned to the object referenced by y.
$x = y$	*"p equals q."*	p and q refer to the same object.
$p \equiv q$ $p \Leftrightarrow q$	*"p is equivalent to q."* *"p if, and only if, q."*	The sentence p is logically equivalent to the sentence q.
$p \vdash q$ $p \Rightarrow q$	*"p proves q."* *"p implies q."*	By assuming the sentence p, we can prove the sentence q.
$\varnothing$	*"The empty set."*	The set containing no elements.
$\{a, b, c\}$	*"The set containing a, b, and c."*	The object whose elements are a, b, c.
$\{x \mid \varphi(x)\}$	*"The set of all x such that $\varphi(x)$."*	The object whose elements consist of all possible objects x for which the sentence $\varphi(x)$ is *true*.
$\{x \in \mathcal{A} \mid \varphi(x)\}$	*"The set of all x in $\mathcal{A}$ such that $\varphi(x)$."*	The object consisting of all x from $\mathcal{A}$ for which the sentence $\varphi(x)$ is *true*.
$f : \mathcal{A} \to \mathcal{B}$	*"f is a function from $\mathcal{A}$ to $\mathcal{B}$."*	A function named f with domain $\mathcal{A}$ and codomain $\mathcal{B}$.
$f(x)$	*"f of x."*	The output of f on the input x, where $x \in \mathcal{A}$ and $f(x) \in \mathcal{B}$.
$\mathbb{N}$	*"enn"*	The set of natural numbers.
$\mathbb{Z}$	*"zee"*	The set of integers.
$\mathbb{Q}$	*"queue"*	The set of rational numbers.
$\mathbb{R}$	*"arr"*	The set of real numbers.

Table 1: An overview of some important notation. Note that some expressions, like $p \equiv q$ and $p \vdash q$, have more than one equivalent notation. The middle column gives some common ways of *reading* each notation in English. The last column provides the *meaning* of each expression.

COLOR	INTERPRETATION
	Emphasis
	Pronunciation
	Definition
	External link
	Internal link

Table 2: Color legend.

MARK	MEANING
定義	***definition***
直覺	***idea***
公理	***axiom***
引理	***lemma***
定理	***theorem***
推論	***corollary***
演算法	***algorithm***

Table 3: Notation for organizing topics.

GLYPH	NAME	IPA	GLYPH	NAME	IPA
$A\ \alpha$	*alpha*	[a]	$N\ \nu$	*nu*	[n]
$B\ \beta$	*beta*	[v]	$\Xi\ \xi$	*xi*	[ks]
$\Gamma\ \gamma$	*gamma*	[ɣ]	$O\ o$	*omicron*	[o]
$\Delta\ \delta$	*delta*	[ð]	$\Pi\ \pi$	*pi*	[p]
$E\ \varepsilon$	*epsilon*	[e]	$P\ \rho$	*rho*	[r]
$Z\ \zeta$	*zeta*	[z]	$\Sigma\ \sigma$	*sigma*	[s]
$H\ \eta$	*eta*	[ɛː]	$T\ \tau$	*tau*	[t]
$\Theta\ \theta$	*theta*	[θ]	$\Upsilon\ \upsilon$	*upsilon*	[yː]
$I\ \iota$	*iota*	[iː]	$\Phi\ \varphi$	*phi*	[f]
$K\ \kappa$	*kappa*	[k]	$X\ \chi$	*chi*	[k^h]
$\Lambda\ \lambda$	*lambda*	[l]	$\Psi\ \psi$	*psi*	[ps]
$M\ \mu$	*mu*	[m]	$\Omega\ \omega$	*omega*	[ɔː]

Table 4: The Greek alphabet.

Logic

0

Language

> *"No language is justly studied merely as an aid to other purposes. It will in fact better serve other purposes, philological or historical, when it is studied for love, for itself."*
>
> – J. R. R. Tolkien

We communicate our thoughts to others with the use of language. This is worth reflecting on. You are probably reading this because you have some interest in computation, mathematics, logic, or are incurably bored; the goal of these notes is—in part—to provide the mathematical background necessary to study these fields at a higher level. This is particularly true for aspiring *computer scientists,* who may have some misconceptions about their field because of its misleading name,[1] and who may not be aware that the field properly and historically falls under the grand umbrella of *mathematics.*

This ambitious undertaking must therefore involve engaging with the tumultuous and violent history of mathematics. Although modern computer science is now richly interdisciplinary, the field was born during a particularly turbulent period in the late 19th and early 20th centuries AD[2] agitated by an existential crisis in mathematics: a crisis caused by our flagrant use of language. Here's a short summary.

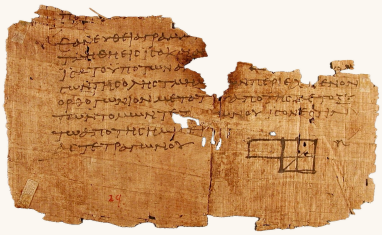

Figure 1: A fragment of book 2 from Euclid's *Elements* taken from the Oxyrhynchus papyri, dated ca. 100 AD.

[1] It's *not* about computers *nor* is it science.

[2] We will see later that its roots span at least to the time of Euclid in 300 BC.

0.1 A Brief History of...

The serious study of rhetoric—the art of argumentation and persuasion—as a subject in its own right dates back to at least the 5th century BC.[3] Around the 3rd century BC, Euclid's 13 books of the *Elements* heralded the birth of geometry, algorithmic computation, and the first theory of numbers,[4] where he *proved* certain statements followed from a list of *axiomatic* assumptions. This was a great achievement, establishing mathematical *proof* as a form of *argumentation* that logically deduces conclusions from a list of common assumptions. The contemporaneous Greek philosopher Theophrastus further pushed the envelope by describing the *form* of these arguments and establishing their validity.

[3] The time of the ancient Greek sophists, who were notably opposed by Socrates, Plato, and Aristotle.

[4] The only evidence of algorithms before this time—for multiplying, factoring, and finding square roots—dates back to Egypt and Babylon before 1600 BC.

The ancient Greeks laid the foundation for the two instrumental aspects of mathematical thought: *abstraction* and *argumentation*. Euclid abstracted what were thought to be the fundamental truths of geometry into a list of 12 *axioms*[1] so that, instead of thinking about *that* particular wall or *that* particular stick or *that* particular roof, he could make statements and observations about *quadrilaterals*, and *lines*, and *triangles* in general. These axioms were meant to encode the *universal truths* of geometry: the nature of what it fundamentally means to construct and measure distances, angles, and (simple) shapes. The last of these axioms would quickly become infamous.

axiom

[1] An *axiom* is a statement that we assume is true without justification nor proof.

Axiom (Parallel Postulate).
If two straight lines meet a third straight line making two interior angles that are each less than right angles, then the two lines—if they were to be extended—must intersect on that side where the interior angles are.
公理

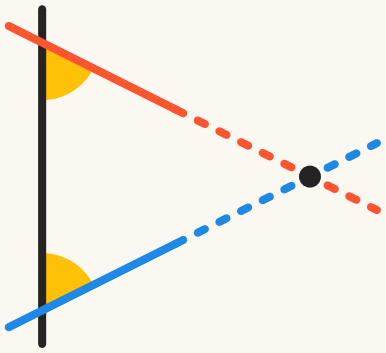

Figure 2: The parallel postulate says that any two lines and that make acute interior angles and with a third line must intersect at a point.

If you stop to think for a moment, this postulate says something very obvious. Assuming all of Euclid's other axioms, there are a few *equivalent* ways to restate the parallel postulate:

1. For any line L and point P not on L, there is exactly one line parallel to L passing through P.
2. The sum of interior angles in any triangle is 180 degrees.
3. A right triangle with side lengths A, B, C satisfies $A^2 + B^2 = C^2$.

You'll recognize this third statement as the Pythagorean *theorem*,[2] which is not merely an assumption![3] For the next 2000 years, the mathematical community was haunted by the thought that it was possible to *prove* the parallel postulate using the other axioms. It seemed like the rest of the axioms did such a perfectly good job of characterizing geometry that the parallel postulate *must necessarily* follow from the other axioms.

[2] A *theorem* is a statement that has a proof.

[3] The first two are called Playfair's axiom and the triangle postulate respectively.

However, between 1810–1832 AD, no less than *three* papers on *hyperbolic* geometry were published, and by 1854 Bernhardt Riemann had developed a theory of *Riemannian* geometry on manifolds. These were all different examples of consistent models of geometry that *denied* the parallel postulate! These ideas were intensely contested: many mathematicians and natural philosophers of the time refused to accept the notion that geometry could be non-Euclidean because *it went against their intuitive notion of how geometry should behave.*

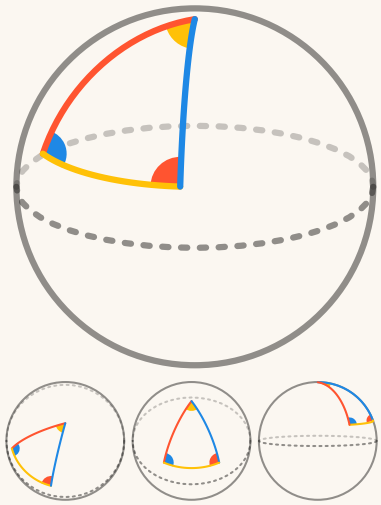

Figure 3: Four views of the same triangle whose angles sum to 270 degrees. Notice how the notions of *straight* and *parallel* differ on the surface of a sphere.

This whole ordeal was only foreshadowing what would come at the turn of the century. In 1874, Georg Cantor would make a series of discoveries[4] surrounding the nature of infinity so fundamentally opposed to common mathematical thought that he would be antagonized and

[4] We will study these later.

ostracized for decades, causing him to suffer serious depressive crises. Once again, mathematicians' *intuitive notions of how infinity should behave* were shown to be wrong. Cantor's discoveries sparked not only a civil war within the mathematical community but also a concerted effort by many mathematicians and logicians in the early 20th century to *fix* mathematics by establishing it on a firm *foundation*.

The cause of all this turmoil was, fundamentally, a *lack of precision and rigor* in the way people would communicate mathematical ideas and arguments. What does it *mean* for a line to be straight, or for two straight lines to be parallel? What does it *mean* to have two lines, or to have infinitely many lines? What *is* infinity? Is infinity a number? What *are* numbers? How do we *know* we are saying anything *true* at all?

If we hope to answer any of these questions, we must first develop a language for *precise* mathematical communication. This necessarily begins with a systematic deconstruction and analysis of *language* itself.

0.2 Syntax and Semantics

Languages encode ideas into sequences of symbols.[1] These symbols represent objects, ideas, actions, and concepts. The *meaning* behind a particular cluster of symbols is called its *semantics*. The *form* the language takes, dictated by its *grammatical rules* for composing symbols into valid sentences, is called its *syntax*. We refer to objects by giving them names. A *variable* is a symbol[2] that stands in place for an object that has not been determined yet.[3] We can assign name to a *particular* object with the $:=$ symbol. We call these the *terms* of an expression.

semantics

syntax

variable

term

[1] For our purposes, we will focus only on written—as opposed to spoken or signed—languages.

[2] We typically denote variables using single Latin or Greek letters, though there are no strict universal rules. Some common examples are listed below.

- $a, b, c, i, j, k, \ell, m, n, p, q, u, v, w, x, y, z$
- $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{G}, \mathcal{H}, \mathcal{M}, \mathcal{N}, \mathcal{R}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$
- $\alpha, \beta, \gamma, \delta, \epsilon, \eta, \theta, \lambda, \mu, \pi, \sigma, \tau, \varphi, \psi, \omega$

[3] A variable does not *necessarily* refer one particular object, or even any object at all.

Definition* 0.1 *(Sentences).

sentence

atomic

A *sentence* is the expression of a complete thought or idea in accordance with the syntactic and grammatical rules of a given language. A statement is called *atomic* if it can't be broken down into smaller semantic components in any way that obeys the language's syntax and grammar.

1. A *declarative* sentence is one that describes something. They typically consist of a *subject* being described and a *predicate* property it has.
2. An *interrogative* sentence asks a non-rhetorical question.
3. An *imperative* sentence heralds a command or request.

"Oft hope is born when all is forlorn."

"What has it got in its pocketses?"

"Keep your forked tongue behind your teeth."

Mathematical practice principally involves *making and justifying observations about mathematical objects*.[4] As such, we are only really interested in crafting *declarative* sentences—sentences that describe *terms*. We will systematically deconstruct and analyse these kinds of sentences, extract their *logical essence*, and build up a new language.

[4] We leave the problem of *what* a mathematical object actually *is* for later.

0.3 A Recurring Theme

Before going any further, we should make a brief detour to discuss a topic that lies at the *heart* of computing, logic, and the 20th century foundational crisis in mathematics: *recursion*. In a very strong sense, what we *mean* when we say that some *thing* is *computable* is that there is a *recursive procedure* that produces that *thing*.

Idea (Church-Turing Thesis).
We say something is *computable* if it is expressible as a *general recursive process*, is a *term in the λ-calculus*, or could described by a *Turing machine*.

範例

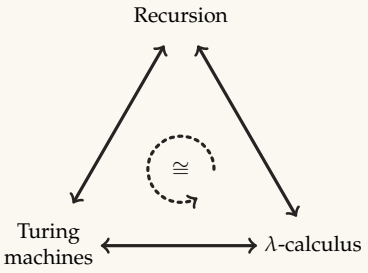

Figure 4: The Church-Turing thesis states that these three concepts—which are all *formally* equivalent—correspond with our *informal* notion of computability. In modern times, many people now take this as a definition for computability.

[1] We leave Turing machines and the λ-calculus for a future time.

Actually, the three concepts described above are all *equivalent* to each other. It should then be no surprise that *recursion* (and its twin *induction*) will play a central role in our studies, so we will take this brief moment to quickly describe the fundamental idea at behind recursion.[1]

First, an example: how do we *compute* the sum of a list of n numbers?

$$3+5+9+2$$

With some hard work and determination access to the internet, we can see that $3+5+9+2=19$, but *how* did we get that answer? At the most basic level, we started by taking two of the numbers, 3 and 5 say, computing their sum $3+5=8$, and adding this intermediate result to another number from the list, 9 say, to get $8+9=17$, and adding that again to yet another element of the list—in this case, only 2 remains—to finally arrive at $17+2=19$.

$$\begin{aligned} 3+5+9+2 &= 3+5+9+2 && (1)\\ &= 8+9+2 && (2)\\ &= 8+9+2 && (3)\\ &= 17+2 && (4)\\ &= 17+2 && (5)\\ &= 19 && (6) \end{aligned}$$

This might seem so obvious it physically hurts, but let's analyse what we just did more closely. Suppose we have a list of n arbitrary numbers.[2]

$$x_0 + x_1 + x_2 + \cdots + x_{n-2} + x_{n-1}$$

Once again, we begin by taking the first two numbers and computing $x_0 + x_1$, then adding *this* result to x_2, then adding *that* result to x_3, then adding *that* result to x_4, and so on until we reach the end of the list. So, in order to compute $x_0 + x_1 + x_2 + \ldots x_{n-2} + x_{n-1}$, we *first* need to compute $x_0 + x_1 + x_2 + \ldots x_{n-2}$ and then add that result to x_{n-1}.

[2] Notice that, being the sophisticates we are, we start counting at 0, so that a list of n numbers will be indexed starting at 0 and ending at $n-1$.

But wait, isn't $x_0 + x_1 + x_2 + \ldots x_{n-2}$ also the sum of a list? It is, it's just that the list has one less element! So *how do we compute the sum of elements in a list*? We first *compute the sum of elements in a list*, and then add one more element to that result. So, it seems like in order to do what we want, we need to already know how to do what we want; the key here is that we only need to know how to sum the elements of a *smaller* list in order to get the result we want for the *larger* list. As long as we can *eventually* get a result for one of these *"smaller"* sums, we will be able to build up a solution to our original problem by passing this result *"back up"* the chain of computation. Back to our first example.

$$
\begin{aligned}
3+5+9+2 &= 3+5+9+2 && (1)\\
&= 3+5+9+2 && (2)\\
&= 3+5+9+2 && (3)\\
&= 8+9+2 && (4)\\
&= 17+2 && (5)\\
&= 19 && (6)
\end{aligned}
$$

This paragraph describes the *recurrence.*

Steps (1) through (3) continually decompose the given list into sublists on the left until we have no more lists we can break up. Each one of these lists is a smaller version of the original problem, and we compute the sums of these smaller lists by breaking them down and computing *their* sublists' sums, recombining these results at the end.

This paragraph encounters the *basis.*

This now brings us to an important point: *we can't decompose* 3 *any further,* because this list only has one element in it. Do we know what the sum of all numbers in a list with one element is? Of course we do: it's just *that* number. Now we can return this result *back up* to the 5 that was waiting to be added to it, and when we add them together, we can return *that* result back to the 9 that was waiting, and then return *that* result to the 2 that was waiting, finally letting us conclude that the sum over the whole list is 19. The *recurrence relation* below summarizes this.[1]

recurrence relation

$$
\texttt{sum}(x_0, x_1, \ldots x_{n-1}) = \begin{cases} 0 & \text{if } n = 0 \\ \texttt{sum}(x_0, x_1, \ldots x_{n-2}) + x_{n-1} & \text{if } n \geqslant 1 \end{cases}
$$

[1] Notice that this is actually written slightly differently than the procedure we've just described; think about *how* this is different and whether or not it actually computes the same result as the procedure we were just analysing.

We've exposed here a *recurrence* and a *basis*—the two key components underlying recursion (and, later, induction). The *recurrent* part of this procedure explains how to express a problem in terms of *"smaller"* instances of the *same problem,* describing how to combine the solutions to those subproblems into a solution for the original problem. Obviously, though, if you just keep decomposing problem into subproblems forever, you'll never be able to actually generate an answer to anything. Eventually, you need to *stop* and actually say what the answer to something is. The *basis*, does exactly this by providing explicit answers to the *smallest* versions of the problem.[2]

recurrence

basis

[2] This is sometimes called the *base case.*

1

Zeroth-Order Logic

"The limits of my language means the limits of my world."
– Ludwig Wittgenstein

Figure 1.1: Illustration by Rockwell Kent from "Moby Dick: or, The Whale."

As we saw in the previous chapter, sentences can be broadly classified based on the kind of information they convey—their *functional role* in language. How do we begin deconstructing the descriptive fragment of our language? Naturally, we can think to classify the descriptive sentences by asking the fundamental question: *is this description true?*

1.1 *Truth Values*

Let's consider the following declarative sentence.

"Ahab is a captain." (1.1)

Here we have a descriptive sentence about the term *Ahab*—a man and thus an object of our discourse—asserting he *is a captain.* In the context of Herman Melville's *Moby Dick,* this is an accurate description. Referring to the above sentence as $\sigma_{1.1}$, we would then say $\sigma_{1.1}$ is *true*. We introduce the symbol $\top$ to denote these kinds of sentences.

true
$\top$

"Ishmael is a whale." (1.2)

The above sentence, however, which we will name $\sigma_{1.2}$, immediately furrows the brow and strikes at the heart of our conscience. We know from the story that Ishmael is a sailor, and thus human, and therefore *not* a whale! We should then want to say that $\sigma_{1.2}$ is *false*, reserving the symbol $\bot$ for sentences of this kind.

false
$\bot$

The symbols $\top$ and $\bot$ are also sometimes called *"top"* and *"bot"* respectively.

The attributes *true* and *false* that we are attaching to these sentences are what we call *truth values*, and they are *the* essential component of the kinds of sentences we want to express. Sentences that are *true* all exhibit a quality that makes them similar to each other but dissimilar to *false* sentences, regardless what the actual sentences themselves *mean*

truth value

semantically. What we've just done is *abstract* the fundamental concept of truth value from descriptive sentences. This abstraction allows us to notice that *all true sentences are essentially the same as each other,* at least from the perspective of their truth values, with the same applying to *false* sentences. On the other hand, *true* and *false* sentences are complete opposites. This relationship inspires our first definition below.

Definition 1.1 (Propositional Equivalence).
equivalence
$\equiv$
We say that two sentences φ and ψ are *equivalent* when they have the same truth value. We denote this by writing $\varphi \equiv \psi$.[1]

[1] *"φ is (logically) equivalent to ψ."*

Axiom (Equivalence is an Equivalence Relation).
We will take the following three properties to be *true* for any sentences φ, ψ, and ξ that are carriers of truth values.

1. $\varphi \equiv \varphi$. *reflexivity*
2. *If* $\varphi \equiv \psi$, *then* $\psi \equiv \varphi$. *symmetry*
3. *If* $\varphi \equiv \psi$ *and* $\psi \equiv \xi$, *then* $\varphi \equiv \xi$. *transitivity*

公理

With this new definition, we can formalize our observations from the preceding paragraph as $\sigma_{1.1} \equiv \top$ and $\sigma_{1.2} \equiv \bot$ as well as $\sigma_{1.1} \not\equiv \sigma_{1.2}$. Notice that each of these three expressions is a complete sentence describing properties[2] held by some objects.[3] In fact, these statements were themselves *true* declarative sentences. Now, let's ponder the following sentence, which we will call $\sigma_{1.3}$.

[2] being (or not) logically equivalent

[3] the sentences $\sigma_{1.1}$ and $\sigma_{1.2}$

"Colorless green ideas sleep furiously." (1.3)

Like the previous examples, this is a grammatically correct, declarative sentence, but what does this sentence *mean?* Is it *true?* Is it *false?* Taking the normal English definitions for each of the words in this sentence, it doesn't seem to make any sense. We then clearly can't call it an accurate description of anything, so it can't possibly be *true.* Does that mean it must be *false?* Well, if we assume it is *false,* then what about the following sentence?

"Colorless green ideas *do not* sleep furiously." (1.4)

This one, which we will call $\sigma_{1.4}$, seems to be saying the opposite of whatever $\sigma_{1.3}$ was saying, so if the other one is *false,* then this one must be *true.* The question then becomes: what is $\sigma_{1.4}$ accurately describing? *This* sentence seems to make just as little sense as the original! This should lead us to conclude that $\sigma_{1.3}$ could not have been *false* either, so that sentence *has no truth value!* We call expressions like this *nonsensical* because they *carry no semantic meaning.*
nonsense

Let's now analyse the following statement, which we will call $\sigma_{1.5}$.

"This sentence is *false*." (1.5)

Expressed a little more *formally*, this is the sentence—named $\sigma_{1.5}$—that says $\sigma_{1.5} \equiv \bot$. This certainly doesn't seem like nonsense; it says something clear about a well-understood object. So, what is the truth value of this sentence? We can try reasoning about this like we did before by examining the two possible truth values the $\sigma_{1.5}$ can take.

First, let's assume $\sigma_{1.5}$ is *true*, which we write formally as $\sigma_{1.5} \equiv \top$. By definition, this would imply $\sigma_{1.5}$ is an accurate description of some object, so we should believe what the sentence says about that object. In this case, the object is $\sigma_{1.5}$ and the description is that $\sigma_{1.5} \equiv \bot$. This *contradicts* our initial assumption! ↯ Therefore, $\sigma_{1.5}$ is *not true!*[1]

[1] We conclude this because this is the opposite of our initial assumption, which lead us to a contradiction.

That rules out one truth value. What happens then if we assume $\sigma_{1.5}$ is *false?* Again, we can write this formally as $\sigma_{1.5} \equiv \bot$. By definition, this implies we should *reject* what $\sigma_{1.5}$ is asserting, leaving us with $\sigma_{1.5} \not\equiv \bot$. As before, a *contradiction* emerges! ↯ Therefore, $\sigma_{1.5}$ is *not false* either!

From this simple analysis, we can see that $\sigma_{1.5}$ *does not have a truth value!* Sentences that *contradict themselves* like this are called *paradoxes*.[2] In the preceding analysis, we relied on the idea that $\top$ and $\bot$ are opposed to each other, so that the same sentence can't meaningfully be both $\top$ and $\bot$ at the same time. This should be intuitive based on our natural understanding and usage of the words *true* and *false*, but we will make it a point to *formally* introduce this idea now.

paradox

[2] The word *paradox* is unfortunately overload and context-dependent. When referring to specific sentences, we will use it to specifically mean a self-contradictory sentence such as $\sigma_{1.5}$, but it is also commonly used in some contexts to refer to situations that are simply *unintuitive* rather than outright contradictory.

Axiom (Principle of Bivalence).
Sentences expressing truth values are either *true* or *false* but not both.
公理

What this analysis has hopefully shown us is that *not every* well-formed, declarative sentence expresses a truth value. In order for a sentence to express a truth value, it must satisfy the following three properties.

1. The sentence must be grammatically well-formed.
2. The sentence must be declarative.
3. The sentence must be semantically meaningful.

These are the kinds of statements are *eligible to carry a truth value*—the ones for which *it would make sense* to say they are either *true* or *false*—so they will form the foundation of our new language. We will eventually call these *propositions*, but beware that this is not (yet) a *formal* definition of what a proposition is. First, we need to get a better sense of *what* propositions are linguistically and *how* they are formed.

1.2 *Logical Connectives*

The examples of sentences we've seen so far have all been *atomic*—meaning they can't be broken down into simpler sentences that themselves are complete thoughts—but we can obviously express thoughts that are more than merely atomic. These *compounded* propositions are formed by taking smaller propositional sentences and *connecting* them together based on what our intended meaning is.

p	q	$\neg p$	$p \wedge q$	$p \vee q$	$p \to q$	$p \leftrightarrow q$
$\top$	$\top$	$\bot$	$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$	$\bot$	$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\top$	$\bot$	$\top$	$\top$	$\bot$
$\bot$	$\bot$	$\top$	$\bot$	$\bot$	$\top$	$\top$

Table 1.1: A truth table summarizing the basic connectives of classical logic. The two left-most columns represent the *input* values of the propositions p and q. The remaining columns describe the *output* of each expression given the corresponding inputs on each row.

Each of these different ways of connecting sentences together suggests a different way of *transforming* between truth values by combining the truth values of the component propositions into a truth value for the compound expression.

logical connective

In this section, we will uncover these different transformations—which we will call *logical connectives*—and encode them using *truth tables*, which specify the output truth values for every combination of inputs.

Negations

Suppose we encountered the following sentence, which we call $\sigma_{1.6}$.

"Espresso is not delicious." (1.6)

Immediately, the moral observer will realize the offensive absurdity of this sentence, compelled by the force of conscience to declare $\sigma_{1.6} \equiv \bot$! With this, we could simply carry on with our day; however, pausing to think for a moment, we can see that $\sigma_{1.6}$ is intimately related to the following (much more pleasant) sentence, which we call $\sigma_{1.7}$.

"Espresso is delicious." (1.7)

This sentence is *clearly true,* letting us sigh $\sigma_{1.7} \equiv \top$ in relief. Not only that, it is the saying exactly the opposite of what $\sigma_{1.6}$ asserted! We call propositions like these *negations* of each other. This is our first example of a *transformation* of truth value: the negation of a proposition is another proposition with the opposite truth value. To denote this formally, we introduce the $\neg$ symbol, allowing us to write $\sigma_{1.6} \equiv \neg\sigma_{1.7}$.

negation

$\neg$

Table 1.2: Truth table for negations.

p	$\neg p$
$\top$	$\bot$
$\bot$	$\top$

We can now think of $\neg$ formally as a *unary function* that operates on truth values.[1] This function works by mapping $\neg\top$ to $\bot$ and by

[1] A function is *unary* if it takes only one input argument. We will study functions in more detail later.

mapping $\neg\bot$ to $\top$. This gives us a way of abstracting negations at the level of truth values, so that we can formally define what it means to *negate* a proposition. We provide this definition now in table 1.2, where the left-most column represents the inputs[1] to $\neg$ and the right-most column shows the truth values of the resulting output expression.[2]

[1] ... shown with white backgrounds ...

[2] ... shown with colored backgrounds ...

Conjunctions & Disjunctions

But we can obviously connect two (and sometimes more) sentences together to create larger sentences in English. For example,

"Espresso is delicious, and it nourishes the soul." (1.8)

This sentence is composed of two smaller atomic sentences, namely *"espresso is delicious"* and *"espresso nourishes the soul,"* which we know are both independently *true.* Connecting them together with the word *"and"* should then, based on the way this word works in English, produce another *true* sentence. Conversely, if either of the subexpressions had been *false,* the compound result should also be *false.* This *binary* connective is called the logical *conjunction*, and we denote it using the $\wedge$ symbol. It is defined in table 1.3.

conjunction

$\wedge$

Table 1.3: Truth table for logical conjunctions and disjunctions.

p	q	$p \wedge q$	$p \vee q$
$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$	$\top$
$\bot$	$\top$	$\bot$	$\top$
$\bot$	$\bot$	$\bot$	$\bot$

There are several distinct ways this connective can appear in English that are nonetheless equivalent. Some examples are listed below.

"Espresso is delicious, *and* it nourishes the soul."
"Espresso is delicious *and* soul-nourishing."
"Espresso is delicious, *but* it nourishes the soul."
"Espresso is delicious, *yet* nourishing to the soul."
"Espresso is delicious; *further*, it nourishes the soul."
"*Although* espresso is delicious, it *also* nourishes the soul."

Table 1.4: These sentences are all logically equivalent to $\sigma_{1.8}$, though this list is obviously not exhaustive.

disjunction

$\vee$

The conjunction has a *logical dual* called the *disjunction*, defined in table 1.3 using the $\vee$ symbol and exemplified by the following sentence.

"Espresso is delicious, or it nourishes the soul." (1.9)

We all these two connectives *dual* to each other because negating all of the inputs to one of them is equivalent to negating the output of the other. Specifically, the expression $\neg p \vee \neg q$ is equivalent to $\neg(p \wedge q)$ whenever p and q are propositions.

Definition 1.2 (Logical Duality).

logical duality

We say two logical connectives f and g are *logically dual* if negating the inputs of f is always logically equivalent to negating the output of g. Equivalently, we can say f is *logically dual* to g if applying f after $\neg$ always gives the same result as applying $\neg$ after g on any given inputs.

Conjunctions and disjunctions are just one example of a dual connective pair. In fact, every logical connective is dual to some other connective![1] For now, we present this result about $\wedge$ and $\vee$ *without proof;* we will *prove* this statement when we discuss theorem 1.5 in a short while.

[1] Why might this be? Think about this.

Conditional Statements

We turn our attention now to sentence $\sigma_{1.10}$ below.

"If espresso nourishes the soul, then I will drink it." (1.10)

This is a *conditional* sentence, composed of two subclauses called the *antecedent* and the *consequent*.[2] When we use this sort of linguistic construction, we mean to say that *if* the premise happens, *then* the conclusion must also happen. Said another way: the conclusion must occur *whenever* the premise is satisfied. Notice *we are not asserting anything* about the antecedent or consequent individually! We are only establishing a *relationship* where the consequent occurs *every time* that the premise is satisfied. We call this the *material implication*, denoted by the $\to$ symbol and defined in table 1.5.

implication
$\to$

Table 1.5: Truth table for conditionals.

p	q	$p \to q$	$p \leftrightarrow q$
$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$	$\bot$
$\bot$	$\top$	$\top$	$\bot$
$\bot$	$\bot$	$\top$	$\top$

[2] Synonyms for *antecedent* & *consequent*.

protasis	apodosis
sufficient	necessary
premise	inference
assumption	conclusion
supposition	deduction
implicant	implicand
hypothesis	thesis

$$p_{1.10} := \text{"Espresso nourishes the soul."}$$
$$q_{1.10} := \text{"I will drink espresso."}$$

The antecedent and consequent for $\sigma_{1.10}$ are defined above. With these definitions, we can now write $\sigma_{1.10} \equiv p_{1.10} \to q_{1.10}$ and observe that $\sigma_{1.10}$ simply says: *if* $p_{1.10} \equiv \top$, *then* $q_{1.10} \equiv \top$. Importantly, this is *the only thing* that $\sigma_{1.10}$ is asserting! This sentence *is not saying* that if $p_{1.10} \equiv \bot$, then $q_{1.10} \equiv \bot$. In fact, if the premise is *false,* then $\sigma_{1.10}$ says *nothing* about whether or not $q_{1.10}$ is *true* or *false.*

To make this concrete, suppose I told you the following.

"If you make an $\mathcal{A}$ in this class, then I will eat my shoe." (1.11)

If you do happen to make an $\mathcal{A}$ in this class, then I'll be forced to physically eat my shoe in order to keep up my end of the bargain; in that case, the sentence was *true.*[3] On the other hand, if you make an $\mathcal{B}$ instead, then I can go home with both shoes and conscience intact; in this case, the sentence was also *true.*[4] However, what if you make the $\mathcal{B}$ but I decide to eat my shoe anyways? Did I lie? No; just because you failed to make an $\mathcal{A}$ doesn't mean I *can't* eat my shoe! All I said was that I definitely would if you made an $\mathcal{A}$.[5] That sentence is only a lie when you *do* make an $\mathcal{A}$ in the class, but I refuse to eat my shoe, since I really am breaking my promise then.[6]

[3] $\top \to \top \equiv \top$

[4] $\bot \to \bot \equiv \top$

[5] $\bot \to \top \equiv \top$

[6] $\top \to \bot \equiv \bot$

In table 1.6, we list several ways of verbalising $p \to q$ in English. Since this connective can be worded in so many unintuitive ways; careful attention must be paid to phrases involving conditionals.

"I will drink espresso *if* it nourishes the soul."
"Espresso nourishes the soul *only if* I drink it."
"It is *sufficient* that espresso nourish the soul for me to drink it."
"It is *necessary* that I drink espresso for it to nourish the soul."
"I will drink espresso *unless* it doesn't nourish the soul."

Table 1.6: These sentences are *all* logically equivalent to $\sigma_{1.10}$. Pay close attention to *grammar* of each sentence, and make special note of *where* the connectives appear.

biconditional ↔

Finally, the *material equivalence*,[1] also called the *biconditional* and written $p \leftrightarrow q$, is *true* exactly when p and q have the same truth value and is *false* otherwise. With these connectives all defined, we are now ready to formally introduce the *recursive definition* of a proposition.

[1] This is often written *"if and only if"* in English, abbreviated *iff* .

A Formal Proposition

Definition* 1.3 *(Proposition).

proposition

We say that λ is a *proposition iff* λ satisfies the following recurrence.

1. $\lambda = \top$ or $\lambda = \bot$.
2. $\lambda = \neg(\varphi)$, where φ is a proposition.
3. $\lambda = (\varphi) \wedge (\psi)$ where φ and ψ are propositions.
4. $\lambda = (\varphi) \vee (\psi)$, where φ and ψ are propositions.
5. $\lambda = (\varphi) \to (\psi)$ where φ and ψ are propositions.
6. $\lambda = (\varphi) \leftrightarrow (\psi)$, where φ and ψ are propositions.

Notice the use of *equality* $=$ rather than *equivalence* $\equiv$ throughout this definition. In each statement here, we are saying that the statement λ is *equal* to the expression on the right-hand side of the $=$ symbol, meaning *they are the same sentence written in the same way.* This gives a *syntactic* definition of what a proposition is.

The use of parentheses in this definition is to avoid issues with order of operations; in situations where the meaning is clear, we can *carefully* drop parentheses.

This definition works by first establishing as our *basis* that $\top$ and $\bot$ are propositions in (1). We then, in (2) through (6), specify larger propositions *recursively* by composing together smaller, already-existing propositions using logical connectives. This then lets us verify that statements like $((\neg\top) \wedge (\bot \wedge \top)) \to \top$ are indeed propositions.

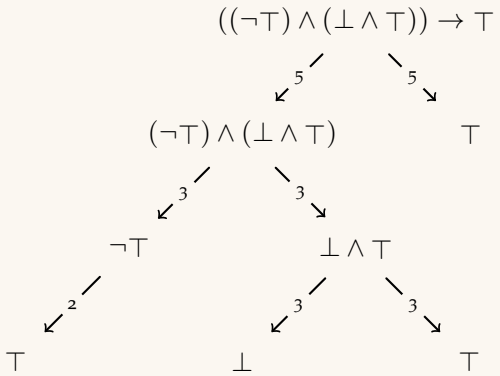

Figure 1.2: In this example, we have dropped some unambiguous parentheses for clarity. Notice, however, that some parentheses *cannot* be dropped: for example, those around the premise of the $\to$ conditional, and those separating the arguments of the two $\wedge$ conjunctions. If those parentheses had been placed like $((\neg\top) \wedge \bot) \wedge \top$ instead, we would have parsed $\not\curlywedge$ instead of $\not\curlyvee$ as in the figure.

Alternatively, think of this as *inductive bootstrapping*.[2] Beginning with $\top$ and $\bot$ from (1) as our initial instances of propositions, we then

[2] *"Pulling itself up by the bootstraps."*

build larger propositions like $\neg\bot$ and $\top \wedge \bot$, which fall into (2) and (3) respectively. We can then take those expressions, conjunct them again using (2), and place an implication between that result and $\top$ using (5) to arrive at our final expression $((\neg\top) \wedge (\bot)) \to \top$. By taking basis expressions and connecting them together according to the rules laid out in the definition, we *computed* a way of building the final expression in a way that satisfies the definition, verifying that it is a proposition.

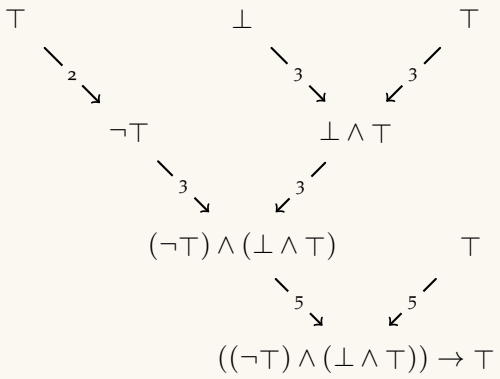

Figure 1.3: The inductive way of building up the expression, as contrasted with the recursive way of tearing down the expression in the previous figure.

Definition 1.4 (Propositional Formula).

propositional formula

A *propositional formula* is an expression that evaluates as a proposition when all of its *variables* are themselves replaced by propositions.

Logical Equivalence

The astute reader may have noticed that some expressions are logically equivalent to each other even if they look different when written out.

p	q	$\neg(p \wedge q)$	$\neg p \vee \neg q$	$p \to q$	$\neg q \to \neg p$
$\top$	$\top$	$\bot$	$\bot$	$\top$	$\top$
$\top$	$\bot$	$\top$	$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\top$	$\top$	$\top$	$\top$
$\bot$	$\bot$	$\top$	$\top$	$\top$	$\top$

Table 1.7: A truth table verifying two equivalences. First, that $\neg(p \wedge q)$ and $\neg p \vee \neg q$ are equivalent as predicted by DeMorgan. Second, that $p \to q$ is equivalent to its *contrapositive* $\neg q \to \neg p$.

For example, it's clear that $p \leftrightarrow q \equiv (p \to q) \wedge (q \to p)$, as the name *"if and only if"* would suggest. We saw another example of an equivalence when we examined the duality of $\wedge$ and $\vee$, illustrated in table 1.7. We can see that statements like these are logically equivalent because the output truth values are always the same whenever we assign the same input truth values to the variables in these expressions. In their joint truth table, the output columns for the two expressions are identical. *Equivalent propositions are essentially the same when we view them through the lens of truth values.*[1]

[1] The idea of blurring the lines between objects that are *essentially the same* according to some salient characteristics is a fundamental idea in mathematics that shows up basically everywhere. This is, fundamentally, *why* abstractions are useful and interesting: we abstract in order to draw equivalences between things we previously thought of as distinct.

Following this idea means having to construct a joint truth table whenever we want to check whether or not two formulæ are equivalent. Although it would be a straightforward to automate, doing all of our work by hand would be *extremely* tedious. If we are given two propositions $\varphi(p_1, p_2, \ldots p_n)$ and $\psi(p_1, p_2, \ldots p_n)$ consisting of the same variables, then answering $\varphi(p_1, p_2, \ldots p_n) \stackrel{?}{\equiv} \psi(p_1, p_2, \ldots p_n)$ requires computing truth values for φ and ψ with *all possible combinations* of truth assignments to $p_1, p_2, \ldots p_n$ and checking that they match.

Now, p_1 can either be $\top$ or $\bot$. For each of these truth values, we then have check both truth values p_2 can take. Then, for each of those, we need to check the two truth values for p_3, and so on until we reach p_n. Each particular assignment of truth values to all of the propositional variables corresponds to one row in our truth table.

If $n = 1$, so our propositions each involve one variable, this means we only need two rows in our truth table to exhaust the entire search space: one row if the variable is $\top$, and one row if it's $\bot$. However, with each new variable we introduce, we *double* the size of our search space because this new variable comes with *two new possible truth values* that we need to check *for each* of the rows we've already computed. We summarize this phenomenon with the following *recurrence relation.*[1]

$$\texttt{rows}(n) = \begin{cases} 1 & \text{if } n = 0 \\ 2 & \text{if } n = 1 \\ 2 \cdot \texttt{rows}(n-1) & \text{if } n \geqslant 2 \end{cases} \tag{1.12}$$

[1] The degenerate case of $n = 0$, when neither expression has any propositional variables, would just require one row in our truth table since each proposition only has one, unchanging truth value.

This shows us that answering the equivalence question for propositional formulæ of n variables involves computing a truth table with 2^n rows. Obviously, *this doesn't scale;* it quickly becomes infeasible to even *allocate enough space* for our output columns, much less actually compute and check these outputs. The thinking man's alternative is to instead *prove* that the two expressions are equivalent, constructing a formal, logical argument that derives $\varphi(p_1, p_2, \ldots p_n) \equiv \psi(p_1, p_2, \ldots p_n)$ from assumptions—called *axioms*—using *rules of inference.*

proof

axiom

Logical Nonequivalence

Showing that two propositional expressions are *not* equivalent is computationally easier than showing that they *are*. Checking that two propositional formulæ are equivalent involves either writing proof or computing *every row* of an exponentially sized truth table. However, checking that two formulæ are *not* equivalent requires *just one example* of a truth assignment on which the propositions disagree. Instead of an entire truth table, all we need is a *single row*.

p	q	$\neg(p \wedge q)$	$\neg p \wedge \neg q$
$\top$	$\top$	$\bot$	$\bot$
$\top$	$\bot$	$\top$	$\bot$
$\bot$	$\top$	$\top$	$\bot$
$\bot$	$\bot$	$\top$	$\top$

Table 1.8: A truth table showing negations do not distribute over conjunctions.

For example, to show that $p \to q \not\equiv q \to p$, all we have to do is let $p := \top$ and $q := \bot$. We can then observe that $p \to q \equiv \top \to \bot \equiv \bot$. Meanwhile, $q \to p \equiv \bot \to \top \equiv \top$. Thus, we conclude $p \to q \not\equiv q \to p$.

Definition 1.5 (Logical Equivalence & Nonequivalence).
Let φ and ψ be propositional formulæ both consisting of the *same* variables $p_1, \ldots p_n$. We say that φ is *equivalent* to ψ if *every* assignment of truth values to the variables of φ and ψ produces the same truth value. In this case, we write $\varphi \equiv \psi$.

logical equivalence

$\equiv$

We say that φ is *not equivalent* to ψ if *there is* an assignment of truth values to the formulæ's variables that makes the truth values of φ and ψ different. In this case, we write $\varphi \not\equiv \psi$.

logical non equivalence

$\not\equiv$

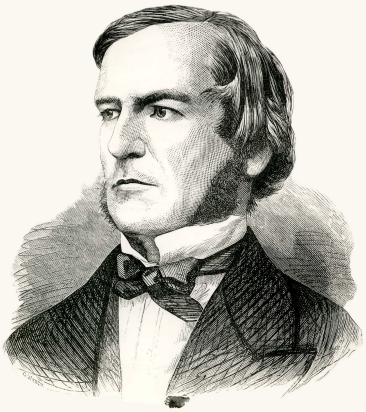

Figure 1.4: George Boole, a largely self-taught mathematician, logician, and philosopher, first described the eponymous *Boolean algebra* in his 1854 monograph *The Laws of Thought.*

1.3 The Propositional Logic

Axioms & Proofs

The axioms of propositional logic encode the *foundational assumptions* we are making about the nature of truth-value-based reasoning. We take these truths to be self-evident *without justification.*

IDENTITY	$\top \wedge p \equiv p$	$\bot \vee p \equiv p$
COMPLEMENT	$\neg p \wedge p \equiv \bot$	$\neg p \vee p \equiv \top$
COMMUTATIVITY	$p \wedge q \equiv q \wedge p$	$p \vee q \equiv q \vee p$
ASSOCIATIVITY	$p \wedge (q \wedge r) \equiv (p \wedge q) \wedge r$	$p \vee (q \vee r) \equiv (p \vee q) \vee r$
DISTRIBUTIVITY	$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$	$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$
CONDITIONAL DISINTEGRATION		$p \to q \equiv \neg p \vee q$
BICONDITIONAL DISINTEGRATION	$p \leftrightarrow q \equiv (p \to q) \wedge (q \to p)$	

Table 1.9: The axioms of classical logic. The first five specify a *Boolean algebra*; notice that each of these first five axioms has a conjunctive fragment (left) and a *dual* disjunctive fragment (right).

Each of the statements in this table is a logical equivalence establishing that the two expressions are *interchangeable in all contexts.* We could verify each of these by constructing the appropriate truth table; however, the attitude we will take is that each statement in the table simply *is true a priori,* without any need for verification. Instead, they will form the *basis* upon which we build proofs of *other* statements.

The *complement* axiom in the second row of table 1.9 shows us two important facts about the negation of any proposition. If we take a proposition p and conjunct it with its negation $\neg p$, that axiom tells us that we get $\bot$; dually, disjuncting p with its negation gives us $\top$. Is this behavior *characteristic* of $\neg p$? The following theorem tells us *yes,* that any proposition that *behaves like* the negation of p must be *indistinguishable* from $\neg p$ through the lens of truth values! With that said, let's try to prove our first *theorem*.[1]

theorem

[1] A *theorem* is a provable proposition.

Theorem 1.1 (Uniqueness of Complements).
For any propositions p and q, if $p \wedge q \equiv \bot$ and $p \vee q \equiv \top$, then $\neg p \equiv q$.

定理

Proof. Let p and q be arbitrary propositions.[2] Assume $p \wedge q \equiv \bot$ and $p \vee q \equiv \top$.[3] We will prove $\neg p \equiv q$ by showing that $\neg p$ and q are both equivalent to the same expression. First, observe the following.

[2] Since we need to prove this statement *for any two propositions* p and q, we introduce two *arbitrary* propositions at the beginning of our proof.

[3] These assumptions are warranted because they are the *premise* of the *conditional* statement we are proving.

$$
\begin{aligned}
\neg p &\equiv \top \wedge \neg p && \text{by } \textit{identity} \\
&\equiv \neg p \wedge \top && \text{by } \textit{commutativity} \\
&\equiv \neg p \wedge (p \vee q) && \text{because we assumed } p \vee q \equiv \top \\
&\equiv (\neg p \wedge p) \vee (\neg p \wedge q) && \text{by } \textit{distributivity} \\
&\equiv \bot \vee (\neg p \wedge q) && \text{by } \textit{complement} \\
&\equiv \neg p \wedge q && \text{by } \textit{identity}
\end{aligned}
$$

As a result, $\neg p \equiv \neg p \wedge q$. Similarly, we can now observe the following.

$$
\begin{aligned}
q &\equiv \top \wedge q && \text{by } \textit{identity} \\
&\equiv q \wedge \top && \text{by } \textit{commutativity} \\
&\equiv q \wedge (p \vee \neg p) && \text{by } \textit{complement} \\
&\equiv (q \wedge p) \vee (q \wedge \neg p) && \text{by } \textit{distributivity} \\
&\equiv (p \wedge q) \vee (\neg p \wedge q) && \text{by } \textit{commutativity} \\
&\equiv \bot \vee (\neg p \wedge q) && \text{because we assumed } p \wedge q \equiv \bot \\
&\equiv \neg p \wedge q && \text{by } \textit{identity}
\end{aligned}
$$

This gives us $q \equiv \neg p \wedge q$. Therefore, we conclude $\neg p \equiv \neg p \wedge q \equiv q$.

Q.E.D.

Q.E.D. stands for *Quod Erat Demonstrandum,* which is Latin for "what was to be shown has been demonstrated," after the Greek Ὅπερ ἔδει δεῖξαι. This is called a *tombstone,* and it is a traditional way of denoting the end of a proof. Modern authors might use □ or ■ instead.

Notice how *every* statement in the proof above is written with *purpose,* and much of the proof is inspired by *the form of the theorem* we are trying to prove. Let's analyze what just happened. Before we begin writing the proof, we first read the theorem focussing on two things: the *form* of the statement, and *what* the statement says.

First and foremost, this theorem says something about *any propositions.* We have two options for proving something is true about every single

proposition: we can check all of them individually, or we can show that the thing we are trying to prove is an *inherent quality of being a proposition.* The former approach is clearly unworkable whenever we have infinitely many—or even just a large amount of—things to check, as we do here. Instead, we will take the later approach: by taking an *arbitrary* proposition and *making no assumptions, imposing no constraints,* then any argument we make about this particular proposition will also apply to any other proposition we encounter.[1] The first sentence of the proof introduces these two arbitrary propositions.

Now that we know we are proving something *universal* about propositions, we keep reading the theorem and see that it's a statement of the form *"if ___, then ___."* This is a *conditional* statement, and the most straight-forward way to show a conditional statement is *true* is to *demonstrate the conclusion is fulfilled whenever the premise is true.* Thus, we can *assume* the premise of the conditional is *true,* and our task them is to derive the conclusion. The second sentence of our proof assumes the premise, which happens to be a conjunction of two statements.

Up to this point, everything we've done has been determined solely by the *form* of the theorem we are trying to prove. Now, our task is to take what we have and show the conclusion.[2] What follows next is a sequence of logical statements, each of which is *justified,*[3] which ends at the conclusion we wanted. *How* you decide to craft this sequence of statements—what statements to make in what order, what proof techniques to use, what intuition inspired your approach—is entirely dependent on *your style* as long as all of the logic is clear, all of the logical rules are followed, and all of the justification is correct.

Proof-writing is an art form in much the same way building a musical instrument is. When a luthier makes a guitar, the process is guided by the particular luthier's traditions, experiences, style, and tastes; so long as the final product is truly a guitar that sounds and plays like a guitar should, the luthier has complete liberty. While two master luthiers might take radically different approaches that lead to guitars with unique aesthetic qualities, they will nonetheless produce two functioning guitars and preference of one over the other will be a matter of judgement and taste. This is much the same when it comes to writing proofs; the analogue to programming should be clear.

Since we proved theorem 1.1, we can now use this result in the future when proving more complicated statements. For example, it should be easy to see intuitively that $\top \equiv \neg\bot$ and $\bot \equiv \neg\top$, based on the way we use the words *true* and *false* in natural language and how $\top$ and $\bot$ are meant to correspond to those truth values. We can now prove this as a *corollary*—a simple consequence—of theorem 1.1.

[1] As an example, suppose we wanted to prove that the square of any positive number is also positive. We obviously can't check all of the positive numbers one-by-one. Instead, we can take an *arbitrary* number x such that $x > 0$, and then argue that $x^2 > 0$. If we do this successfully, then we can take *any* particular number, such as 5, substitute it for x in our argument, and obtain a proof that $5^2 > 0$. However, if we couldn't have written our *original argument* in terms of 5; this would have meant imposing the *additional constraint* that $x = 5$, preventing our argument from generalizing to *all* positive numbers.

[2] If our conclusion were a longer, compound statement, we would continue breaking the problem down *recursively* until we were left with something *atomic*.

[3] ...either by a *definition*, an *axiom*, an *assumption* we've made, or a prior *theorem* we've proven ...

Figure 1.5: Examples of three distinct bracing styles for the classical guitar.

***Corollary* 1.1.**
$\top \equiv \neg\bot$ and $\bot \equiv \neg\top$.

推論

Proof. Observe that $\bot \wedge \top \equiv \bot$ by the *identity* axiom. Similarly, we have that $\bot \vee \top \equiv \top \vee \bot \equiv \top$ by *commutativity* and the *identity* axiom again. So, we can apply theorem 1.1[1] and conclude $\top \equiv \neg\bot$. Similarly, we can observe that $\top \wedge \bot \equiv \bot \wedge \top \equiv \bot$ by *commutativity* and *identity*, and $\top \vee \bot \equiv \bot$ by the *identity* axiom. Thus, $\bot \equiv \neg\top$ by theorem 1.1.

Q.E.D.

[1] We can invoke the theorem here because we have just *proven* the premises of the theorem are *true* for the particular propositions we are looking at (in this case, $p \coloneqq \bot$ and $q \coloneqq \top$). That means, having satisfied the premises, we get to assert the conclusion, justified by that theorem.

A proof gives us more than just a formal verification of a statement. It tells us that the statement is a *necessary consequence* of the axioms we assumed in setting up our logical system, and every instance of a proof gives us insight into *why* that's the case. These past two proofs show us that we didn't have to explicitly *define* or *assume* $\top$ to be the opposite of $\bot$ because this is a face satisfied by *any* instance of a Boolean algebra.

Let's prove another simple, but useful, theorem.

***Corollary* 1.2.**
For any propositions p and q, if $p \equiv q$, then $\neg p \equiv \neg q$.

推論

Proof. Let p and q be propositions such that $p \equiv q$ and observe.

$$\begin{aligned} q \wedge \neg p &\equiv p \wedge \neg p && \text{because we assumed } p \equiv q \\ &\equiv \bot && \text{by } \textit{complement} \end{aligned}$$

We can do a very similar thing in the disjunctive case.

$$\begin{aligned} q \vee \neg p &\equiv p \vee \neg p && \text{because we assumed } p \equiv q \\ &\equiv \top && \text{by } \textit{complement} \end{aligned}$$

Therefore, applying theorem 1.1, we conclude that $\neg p \equiv \neg q$.

Q.E.D.

***Corollary* 1.3.**
For any propositions p, q, r, s such that $p \equiv q$ and $r \equiv s$, the following.

$$\begin{aligned} p \wedge r &\equiv q \wedge s \\ p \vee r &\equiv q \vee s \\ p \to r &\equiv q \to s \\ p \leftrightarrow r &\equiv q \leftrightarrow s \end{aligned}$$

推論

We include corollary 1.3 above just for completeness, so that some of the basic properties of $\equiv$ are codified somewhere; their proofs are not particularly interesting. We are now ready to tackle the proof of a claim you probably find so obvious as to not even be worth mentioning—but it *is* worth mentioning how much groundwork we had to lay in order to prove this simple fact!

Theorem 1.2 (Double Negation).
For any proposition p, we have that $p \equiv \neg\neg p$.

定理

Proof. Let p be a proposition. We will show that $p \equiv \neg\neg p$ by showing that p acts like the negation of $\neg p$. Observe that $\neg p \wedge p \equiv p \wedge \neg p \equiv \bot$ by *communtativity* and the *complement* axiom. We can similarly see $\neg p \vee p \equiv p \vee \neg p \equiv \top$ by *commutativity* and *complement*. Therefore, we can conclude that $p \equiv \neg(\neg p)$ by theorem 1.1.

Q.E.D.

Our goal for the rest of this section will be to prove *De Morgan's laws*. This will finally establish the establish dual relationship between conjunctions and disjunctions. For this, we first need a few more tools.

Theorem 1.3 (Idempotency).
For any proposition p, we have $p \wedge p \equiv p$ and $p \vee p \equiv p$.

定理

Proof. Let p be a proposition. For the conjunctive statement, observe.

$$\begin{aligned}
p \wedge p &\equiv (p \wedge p) \vee \bot && \text{by } \textit{identity} \\
&\equiv (p \wedge p) \vee (p \wedge \neg p) && \text{by } \textit{complement} \\
&\equiv p \wedge (p \vee \neg p) && \text{by } \textit{distributivity} \\
&\equiv p \wedge \top && \text{by } \textit{complement} \\
&\equiv p && \text{by } \textit{identity}
\end{aligned}$$

An analogous chain of reasoning takes us through the disjunctive case.

$$\begin{aligned}
p \vee p &\equiv (p \vee p) \wedge \top && \text{by } \textit{identity} \\
&\equiv (p \vee p) \wedge (p \vee \neg p) && \text{by } \textit{complement} \\
&\equiv p \vee (p \wedge \neg p) && \text{by } \textit{distributivity} \\
&\equiv p \vee \bot && \text{by } \textit{complement} \\
&\equiv p && \text{by } \textit{identity}
\end{aligned}$$

Therefore, we have $p \wedge p \equiv p$ and $p \vee p \equiv p$ as desired.

Q.E.D.

Theorem 1.4 (Domination).
For any proposition p, we have $p \vee \top \equiv \top$ and $p \wedge \bot \equiv \bot$.

定理

Proof. Let p be a proposition and observe.

$$\begin{aligned} p \vee \top &\equiv p \vee (p \vee \neg p) && \text{by } \textit{complement} \\ &\equiv (p \vee p) \vee \neg p && \text{by } \textit{associativity} \\ &\equiv p \vee \neg p && \text{by } \textit{idempotency} \\ &\equiv \top && \text{by } \textit{complement} \end{aligned}$$

The dual case works out similarly.

$$\begin{aligned} p \wedge \bot &\equiv p \wedge (p \wedge \neg p) && \text{by } \textit{complement} \\ &\equiv (p \wedge p) \wedge \neg p && \text{by } \textit{complement} \\ &\equiv p \wedge \neg p && \text{by } \textit{idempotency} \\ &\equiv \bot && \text{by } \textit{complement} \end{aligned}$$

We therefore conclude $p \vee \top \equiv \top$ and $p \wedge \bot \equiv \bot$.

Q.E.D.

We are now ready to prove *De Morgan's laws.*

Theorem 1.5 (De Morgan's Laws).
If p, q are propositions, $\neg(p \wedge q) \equiv \neg p \vee \neg q$ and $\neg(p \vee q) \equiv \neg p \wedge \neg q$.

定理

Proof. Let p and q be propositions. We will prove the first half of this theorem by showing $\neg p \vee \neg q$ acts like the negation of $p \wedge q$, satisfying the premise of *uniqueness of negations.* We then apply the theorem to obtain $\neg(p \wedge q) \equiv \neg p \vee \neg q$. We will leave the other half—proving $\neg(p \vee q) \equiv \neg p \wedge \neg q$—as an exercise to the reader.

First, the conjunctive branch.

$$\begin{aligned} (p \wedge q) \wedge (\neg p \vee \neg q) &\equiv p \wedge (q \wedge (\neg p \vee \neg q)) && \text{by } \textit{associativity} \\ &\equiv p \wedge ((q \wedge \neg p) \vee (q \wedge \neg q)) && \text{by } \textit{distributivity} \\ &\equiv p \wedge ((q \wedge \neg p) \vee \bot) && \text{by } \textit{complement} \\ &\equiv p \wedge (q \wedge \neg p) && \text{by } \textit{identity} \\ &\equiv p \wedge (\neg p \wedge q) && \text{by } \textit{commutativity} \\ &\equiv (p \wedge \neg p) \wedge q && \text{by } \textit{associativity} \\ &\equiv \bot \wedge q && \text{by } \textit{complement} \\ &\equiv \bot && \text{by } \textit{domination} \end{aligned}$$

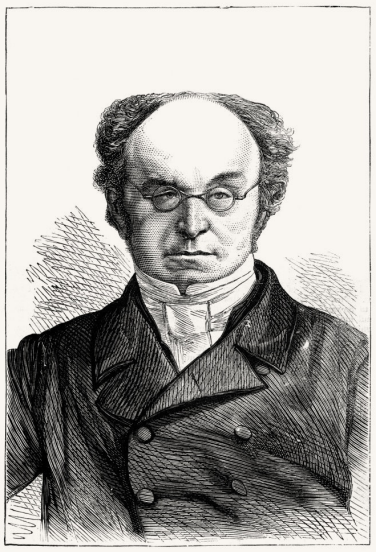

Figure 1.6: Augustus De Morgan, after whom these laws are named, is also notable for his work on logical quantification and mathematical induction.

We have now worked out that $(p \wedge q) \wedge (\neg p \vee \neg q) \equiv \bot$. We will show $(p \wedge q) \vee (\neg p \vee \neg q) \equiv \top$ in the disjunctive branch below analogously.

$$
\begin{aligned}
(p \wedge q) \vee (\neg p \vee \neg q) &\equiv ((p \wedge q) \vee \neg p) \vee \neg q && \text{by } \textit{associativity} \\
&\equiv (\neg p \vee (p \wedge q)) \vee \neg q && \text{by } \textit{commutativity} \\
&\equiv ((\neg p \vee p) \wedge (\neg p \vee q)) \vee \neg q && \text{by } \textit{distributivity} \\
&\equiv ((p \vee \neg p) \wedge (\neg p \vee q)) \vee \neg q && \text{by } \textit{commutativity} \\
&\equiv (\top \wedge (\neg p \vee q)) \vee \neg q && \text{by } \textit{complement} \\
&\equiv (\neg p \vee q) \vee \neg q && \text{by } \textit{identity} \\
&\equiv \neg p \vee (q \vee \neg q) && \text{by } \textit{associativity} \\
&\equiv \neg p \vee \top && \text{by } \textit{complement} \\
&\equiv \top && \text{by } \textit{domination}
\end{aligned}
$$

Therefore, by theorem 1.1, we conclude $\neg(p \wedge q) \equiv \neg p \vee \neg q$ as desired.

Q.E.D.

Rules of Inference

THE DEDUCTION RULE	$(p \vdash q) \vdash (p \to q)$	If, by assuming p, we can prove q, then we can write $p \to q$.
MODUS PONENS	$p, (p \to q) \vdash q$	If we have $p \to q$ and we know p, then we can deduce q.
MODUS TOLLENS	$\neg q, (p \to q) \vdash \neg p$	If we have $p \to q$ but also $\neg q$, then we can infer $\neg p$.
REDUCTIO AD ABSURDUM	$(\neg p \vdash q), (\neg p \vdash \neg q) \vdash p$	If $\neg p$ leads to a contradiction, then $\neg p$ is absurd; we conclude p.

Table 1.10: The rules of inference.

So far, we've developed a modestly-powerful formal language—capable of expressing some basic logical ideas—founded on *axioms*. This gives us a formal syntactic framework for expressing logical ideas, along with a basic semantics that relates these formal symbols to our natural language. The axioms in table 1.9 are all *equivalences*—substitution rules between propositions that preserve truth values—and we've now seen several examples of their use in proving some basic theorems.

Yet, you may have noticed that *some* of our reasoning in those proofs *was not based on equivalences*. This is most apparent in the proof of theorem 1.1, our very first theorem. We began that proof by introducing two arbitrary propositions and then *immediately assuming* that their conjunction was $\bot$ and their disjunction as $\top$. Making those assumptions was not justified on any of the equivalence axioms we'd introduced, so why were we allowed to say that in our proof? By a similar token, in the proof of corollary 1.1, we apply theorem 1.1 by saying that, since

we'd satisfied the premises of that theorem, we were allowed to write down the conclusion of that theorem. Why were we allowed to say that? In short: *because it makes sense!* The problem, of course, is that nothing yet in our system *formally* gives us the right or power to do these things, even though they make logical sense.

This then calls for the introduction of more axioms—ones that will allow us to construct these kinds of *one-way*, *inferential* arguments alongside our equivalence-based reasoning. We call these the *rules of inference*.

The rules in table 1.10 each take the form $\Gamma \vdash \varphi$,[1] where Γ represents a set of assumptions and φ is the conclusion that follows from them. The $\vdash$ symbol, sometimes called a turnstile, signifies that we can *prove* φ by assuming the statements in Γ and using the equivalence axioms, the rules of inference, and any theorems we've already proven. If there is nothing written to the left of the $\vdash$ symbol, this simply means that the conclusion φ can be derived *without* any additional assumptions.

prove

[1] "Γ *proves* φ" or "φ *follows from* Γ."

The most important of the rules of inference is *modus ponens*, enabling us to *follow through* on chains of conditional reasoning.[2] *Modus ponens* is, in a sense, the essence of classical rhetoric. Without it, the conclusion of a conditional statement's conclusion would not be meaningfully *conditioned* on its premise. There would be no point in establishing hypothetical arguments because the conditional chains of reasoning would never actually have any point to work towards. This rule has a sister—*modus tollens*—which conversely allows *breaking down* arguments counterfactually, denying antecedents with false consequents.[3]

[2] *Modus ponens* is short for the Latin phrase *modus ponendo ponens,* literally *"the method of putting by placing."*

[3] *Modus tollens* is short for the Latin phrase *modus tollendo tollens,* literally *"the method of removing by taking away."*

The next rule, named *reductio ad absurdum*,[4] gives us the ability to construct proofs by contradiction. Suppose we are interested in proving some proposition p. One way to reason about the validity of p is to think about what would happen if p were not the case. Hypothetically, assuming $\neg p$, if we were able to derive both q and also $\neg q$, then we would have derived a falsity ($q \wedge \neg q \equiv \bot$). If we were starting from *true* premises, this would be impossible since all of our axioms and rules of inference are *truth-preserving*. Clearly, this must mean that our assumption $\neg p$ was *not true,* leaving p as the only logical conclusion. This form of argumentation is like *"is like arguing with a hammer,"* according to a dear professor of mine from undergrad. It is incredibly powerful and has been in use since at least the year 400 BC.[5]

[4] *Reductio ad absurdum* is a Latin phrase meaning to *"reduce to absurdity."* This has also been called *argumentum ad absurdum*.

[5] In Plato's dialogues, Socrates frequently engages in this sort of reasoning by showing his opponents' seemingly-sensible statements can be systematically dismantled to absurdity.

Finally, the *deduction rule* is a technical rule of inference that ties together the meta-symbol $\vdash$ with the logical $\rightarrow$ symbol. It enshrines the parallel between a deductive *"q follows from p"* statement and a formal *"if p then q"* statement. If this distinction is confusing, just keep in mind that we are constructing a formal language to express mathematical ideas

with; the *propositions* we express are written in our language, but we write our *proofs* of these propositions in our natural language, and our natural language is what we use to write down the rules and axioms that our language must obey. The *deduction rule* tells us that the result of our *proofs* can be converted into statements *within the formal language.*

Although this is a rather small collection of rules, it is capable of representing any kind of expressible propositional rhetoric. Despite that, it's not a *minimal* set of rules for the zeroth-order logic. In fact, it's possible to have an even smaller set of rules without sacrificing the rhetorical strength of our language. *Modus tollens,* for instance, could actually be shown to follow from the other rules of inference as a *theorem,* reducing our total number of assumptions. Let's prove it now.

Theorem* 1.6 *(Modus Tollens).
We have $\neg q, (p \to q) \vdash \neg p$ for any propositions p and q.

定理

Proof. Let p and q be arbitrary propositions, and suppose $\neg q$ and also $p \to q$. We know that $p \to q \equiv \neg q \to \neg p$,[1] so we have $\neg q \to \neg p$. Then, by *modus ponens,* we can conclude $\neg p$.

Q.E.D.

[1] This result—that a conditional statement is equivalent to its *contrapositive*—is left as an exercise to the reader.

The interested reader might be excited to learn that *all* of propositional logic can be encoded using *just two connectives* ($\neg$ and $\to$) and *just three axioms* along with *modus ponens*. There are several classical *syllogisms* that have been studied since the time of the ancient Greeks. Before discussing these, we will first prove three important theorems.

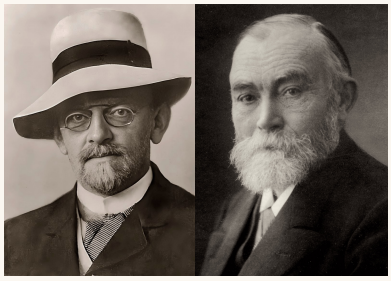

Figure 1.7: David Hilbert and Gottlob Frege were two of the most influential figures in the *logicist program* that was attempting to reduce mathematics to pure logic. Outside of logic, Hilbert was an extremely accomplished algebraist (maybe you've heard of Hilbert spaces in the context of linear algebra). Frege, while underappreciated during his life, is now recognized as one of the greatest and most profound mathematicians and philosophers of language of human history.

Hilbert's System

The logical system we've set up so far—the axioms that establish the propositional calculus as a Boolean algebra, and our comprehensive rules of inference—is very user-friendly, but for this reason it is not *minimal.* We could have made our logical system more *"elegant"*—in some eyes—by choosing a shorter list of axioms and relying on only one rule of inference, at the consequence of having *much uglier* theorems and substantially more tedious proofs. Nonetheless, there is still benefit to be had by studying one of these *more minimal* axiomatizations, as it will provide us invaluable insight into proving a very important theorem: *conjunction elimination.* This alternative axiomatization for the propositional calculus is attributed to Hilbert and Frege. A modern, more condensed version of their system can be written using only two axioms without losing any expressive power. We now prove their axioms as *theorems of our system* below.

Theorem 1.7 (Hilbert's First Axiom).
$\vdash \varphi \to (\psi \to \varphi)$ for any propositions φ and ψ.

定理

Proof. Let φ and ψ be arbitrary propositions and assume φ.

Suppose ψ. We have φ by assumption. Thus, we have $\psi \vdash \varphi$ since we derived φ from ψ. By the *deduction rule*, we then obtain $\psi \to \varphi$.

We now have $\varphi \vdash (\psi \to \varphi)$, since we derived $\psi \to \varphi$ from φ. Therefore, we conclude $\varphi \to (\psi \to \varphi)$ by the *deduction rule.*

Q.E.D.

Theorem 1.8 (Hilbert's Second Axiom).
$\vdash (\varphi \to (\psi \to \xi)) \to ((\varphi \to \psi) \to (\varphi \to \xi))$ for any φ, ψ, and ξ.

定理

Proof. Let φ, ψ, and ξ be propositions and assume $\varphi \to (\psi \to \xi)$. We want to show $(\varphi \to \psi) \to (\varphi \to \xi)$. Towards that goal, assume $\varphi \to \psi$. We now want to show $\varphi \to \xi$; so, towards this goal, assume φ. Now,

$$\varphi, (\varphi \to (\psi \to \xi)) \vdash \psi \to \xi$$

by *modus ponens* using our earlier assumption, so we obtain $\psi \to \xi$. Again, by applying *modus ponens* to our prior assumption, we see that

$$\varphi, (\varphi \to \psi) \vdash \psi$$

leaves us with ψ. We now take these two intermediate results to deduce

$$\psi, (\psi \to \xi) \vdash \xi$$

using *modus ponens.* Therefore, we have derived ξ from our initial assumption φ, letting us conclude $\varphi \vdash \xi$.

We now apply the *deduction rule* several times to arrive at the conclusion. From $(\varphi \vdash \xi)$, we deduce $(\varphi \to \xi)$. Next, from $((\varphi \to \psi) \vdash (\varphi \to \xi))$, we deduce $((\varphi \to \psi) \to (\varphi \to \xi))$. Lastly, we take our expression $((\varphi \to (\psi \to \xi)) \vdash ((\varphi \to \psi) \to (\varphi \to \xi)))$ and finally derive $((\varphi \to (\psi \to \xi)) \to ((\varphi \to \psi) \to (\varphi \to \xi)))$.

Q.E.D.

Classical Syllogisms

We now follow in the footsteps of classical students of rhetoric, who in antiquity would ponder over these (and other) *syllogisms*—a traditional term referring to an argument where a conclusion is drawn from some collection of premises—as a way to hone our skills in the sister arts of proof-writing and deductive reasoning.

The following theorem allows us to construct and follow extended chains of conditional reasoning. Combined with *modus ponens,* this fundamentally forms the basis for any nontrivial argument.

Theorem 1.9 (Hypothetical Syllogism).
We have $(p \to q), (q \to r) \vdash p \to r$ for any propositions p, q, and r.

定理

Proof. Let p, q, and r be arbitrary propositions, and suppose $p \to q$ and $q \to r$. We will first show that $p \vdash r$. Assume p. Since $p \to q$, we have q by *modus ponens.* Further, since we have $q \to r$, we get r by *modus ponens.* Thus, $p \vdash r$. Therefore, by applying the *deduction rule,* we can conclude $p \to r$.

Q.E.D.

The next theorem is the converse of the *deduction rule.* When these two are taken together, they establish the formal, syntactic equivalence between the $\to$ and $\vdash$ symbols, which are semantically distinct.

Theorem 1.10 (Conditional Elimination).
We have $(p \to q) \vdash (p \vdash q)$ for any propositions p and q.

定理

Proof. Let p and q be arbitrary propositions, and suppose $p \to q$. We will now show that $p \vdash q$. Assume p. Then, since we have $p \to q$, we can derive q by *modus ponens.* Thus, $p \vdash q$.

Q.E.D.

Theorem 1.11 (Conjunction Introduction).
We have $p, q \vdash p \wedge q$ for any propositions p and q.

This is known as *adjunction*.

定理

Proof. Let p and q be arbitrary propositions. Assume p, and also separately assume q. Towards a contradiction, suppose $\neg(p \wedge q)$. [1] We can plainly see

$$\begin{aligned} \neg(p \wedge q) &\equiv \neg p \vee \neg q && \text{by } \textit{De Morgan's laws} \\ &\equiv p \to \neg q && \text{by } \textit{conditional disintegration}. \end{aligned}$$

So, we have $p \to \neg q$, from which we can derive $\neg q$ by *modus ponens.* This shows us that $\neg(p \wedge q) \to \neg q$ by the *deduction rule.* However, since we also had q by assumption, we can derive $\neg(p \wedge q) \to q$ using the *deduction rule* again. ↯ [2]

Therefore, we can conclude $p \wedge q$ by *reductio ad absurdum*.

Q.E.D.

[1] When beginning a *proof by contradiction,* it is good form to explicitly alert the reader to this fact with a phrase like *"towards a contradiction."*

[2] The symbol ↯ is useful in *proofs by contradiction* to highlight to the reader where the *contradiction* is and when it is reached.

In table 1.11, we summarize these results and some other theorems. We leave the proofs of these as an important list of exercises to the reader.

MODUS TOLLENS	$\neg q, (p \to q) \vdash \neg p$	
HYPOTHETICAL SYLLOGISM	$(p \to q), (q \to r) \vdash p \to r$	
CONDITIONAL ELIM.	$(p \to q) \vdash (p \vdash q)$	*a.k.a. the consolidation rule*
CONJUNCTION INTRO.	$p, q \vdash p \wedge q$	*a.k.a. adjunction*
CONJUNCTION ELIM.	$p \wedge q \vdash p$	*a.k.a. simplification*
DISJUNCTION INTRO.	$p \vdash p \vee q$	*a.k.a. addition*
DISJUNCTION ELIM.	$(p \to r), (q \to r), (p \vee q) \vdash r$	*a.k.a. proof by cases*
EX FALSO QUODLIBET	$p, \neg p \vdash q$	*a.k.a. explosion*

Table 1.11: Some useful theorems.

2

First-Order Logic

"I am in a charming state of confusion."

– Ada Lovelace

The language we have described so far is often called the *classical logic*—since this is a modern development on Aristotelian logic—or the *propositional logic* because its basic syntactic unit is the proposition. Having the proposition as the most granular accessible referent helps keep this language manageable, but it will hold us back from being as expressive as we'd like to be. For example, suppose we are hungry, and in the course of our ruminations we discover that shepherd's pie is irresistibly delicious. We also happen to know the same thing about paella. Having recognized these facts, no simple substitute will do: we *must* have one of these two meals if we are to be satisfied at all. How might we express this logically? Let's introduce some definitions.

$$s := \text{"We eat } \textit{shepherd's pie.}\text{"}$$
$$p := \text{"We eat } \textit{paella.}\text{"}$$
$$n := \text{"We do not eat } \textit{anything.}\text{"}$$

The claim we are trying to express would formally look as follows.

$$(\neg s \wedge \neg p) \rightarrow n \tag{2.1}$$

From the syntax above, it doesn't seem like there is any relationship between the premise of that conditional statement and its conclusion. In fact, there doesn't even appear to be a relationship between s and p, even though they are both saying something really similar, because *syntactically* they just look like two distinct propositions! Suppose our friend felt the same way as we do about food, but he additionally knew about a *secret third food:* the tostada. Our friend might then resolve to have *that* meal as a fall-back if he can't get his hands on shepherd's pie or paella. He would let $t :=$ "We eat *a tostada.*" and say the following.

$$(\neg s \wedge \neg p) \rightarrow t \tag{2.2}$$

English *shepherd's pie,* as God intended.


The humble *paella,* national dish of Spain.


Tostada & café, a classic Cuban breakfast.

Now, despite our two claims having the *exact same syntactic form*, they express remarkably different ideas. To realize this, think about what it would take to prove (2.1): after verifying $\neg s$ and $\neg p$, we would then need to show we did not eat *any other food!* This is a *universal claim* we are making about *all* possible meals. However, our friend is not making this kind of claim: his conclusion is simply that *there exists* a particular meal he eats if $\neg s$ and $\neg p$ are satisfied. To prove himself right, he simply has to show that he ate that particular meal.

2.1 A More Expressive Language

It will quickly become frustrating for our language to limit our expressivity like this. The missing component in our language is the ability to distinguish the *object* of our speech from the *predicate* description we make about it when we declare a proposition.

Every man is mortal.
Socrates is a man.
$\therefore$ Socrates is mortal.

The argument above seems like a clear, sensible argument; it in fact looks like a simple application of *modus ponens*. Yet, we realize that a proof of this argument in the propositional logic could not actually invoke *modus ponens*. There is no way to symbolize the first sentence in such a way that we obtain a conditional $x \to y$ where the premise is *"Socrates is a man,"* and if we can't do that then we can't apply *modus ponens*. We fix this issue by augmenting our language with the ability to *syntactically* distinguish between *predicates* and the *terms* they describe.

Definition 2.1 (Term).

term

A *term* is a symbol denoting an object. Specific terms—*e.g.*, the natural number 5, Socrates, shepherd's pie—are called *constants*. Placeholder terms denoting objects that have not been specifically determined are called *variables*. Notice that *terms, on their own, do not form complete sentences! A term does not have a truth value!*

Definition 2.2 (Predicate).

predicate

universe of discourse

Let $x_1, \dots x_n$ be variable symbols. We say $\varphi(x_1, \dots x_n)$ is an *n-ary predicate* if replacing each of the n variables $x_1, \dots x_n$ by terms $t_1, \dots t_n$ from our results in a *proposition* $\varphi(t_1, t_2, \dots t_n)$, carrying a truth value. The collection of all terms that our language has referential access to is our *universe of discourse*.

We've now introduced a new problem into our language though. Suppose we have define the predicates $\mu(x) :=$ "x is a man" and

$\theta(x) :=$ "x is mortal" in an attempt to translate the previous argument. We can now translate the second premise and conclusion as $\mu(\text{Socrates})$ and $\theta(\text{Socrates})$ respectively. But we still can't translate the first line. For this, we need the ability to express *quantities*.

universal
$\forall$

Let $\varphi(x_1, \dots x_i, \dots x_n)$ be an n-ary predicate containing a variable x_i. The *universal quantification* of the variable x_i appearing in φ is denoted $\forall x\big(\varphi(x_1, \dots x, \dots x_n)\big)$ and says *any constant* replacing x will satisfy φ.

```
def forall(universe: Iterable, phi: callable) -> bool:
  for x in universe:
    if not phi(x):
      return False
  return True
```

Figure 2.1: A hypothetical implementation of $\forall x\big(\varphi(x)\big)$. If `False` is returned, then there is at least one `x` in `universe` such that `phi(x) == False`, which is equivalent to $\forall x\big(\varphi(x)\big) \equiv \bot$. Otherwise, every `x` in `universe` will satisfy `phi(x) == True`, which means exactly that $\forall x\big(\varphi(x)\big) \equiv \top$.

existential
$\exists$

The *existential quantification* of x_i is denoted $\exists x\big(\varphi(x_1, \dots x, \dots x_n)\big)$ and claims that *there is at least one* constant that, in place of x, satisfies φ. The *scope* of a quantifier is denoted by parentheses specifying its variable's lifetime; that variable is *bound* to that quantifier within that scope. A

free variable

variable that is not bound to any quantifier is called *free*. Statements with free variables *cannot have truth values*, they do not carry *meaning*. If a statement has free variables, those variables need to either be replaced by *terms*, or be bound to a *quantifier*.

```
def exists(universe: Iterable, phi: callable) -> bool:
  for x in universe:
    if phi(x):
      return True
  raise False
```

Figure 2.2: A hypothetical implementation of $\exists x\big(\varphi(x)\big)$. If `True` is returned, then there must be an `x` in `universe` such that `phi(x) == True`, which is equivalent to $\exists x\big(\varphi(x)\big) \equiv \top$. Otherwise, every `x` in `universe` will satisfy `phi(x) == False`, so that $\exists x\big(\varphi(x)\big) \equiv \bot$.

unique existential
$\exists!$

We also introduce the *unique existential quantification* of x_i as a way of saying that *there is exactly one* constant satisfying φ in place for x. We use the notation $\exists! x\big(\varphi(x_1, \dots x, \dots x_n)\big)$, to denote this, which is read as *"there exists a unique x such that $\varphi(x)$"* in English.[1]

[1] This will be useful in future chapters.

$$\exists! x\big(\varphi(x)\big) \;:\Leftrightarrow\; \exists x\Big(\varphi(x) \wedge \forall y\Big(\varphi(y) \;\Rightarrow\; (y = x)\Big)\Big).$$

This is a special case of existential quantification; using the unique existential quantifier means making an existential claim *and additionally* asserting that only one such example exists. So, we define the $\exists!$ quantifier *in terms of* the $\exists$ quantifier. Be careful to note that the ! symbol in $\exists!$ *does not correspond with negating anything!* Do not make the mistake of confusing ! with $\neg$ if you have experience with a programming language where the ! syntax corresponds to logical negation.

Forming Formulæ Well

Definition 2.3 (Atomic Formula).

atomic formula

We say a formula φ is *atomic* if it satisfies the following recurrence.

1. $\varphi = \top$ or $\varphi = \bot$.
2. $\varphi = \psi(t_1, \dots t_n)$, where ψ is an n-ary predicate, $t_1, \dots t_n$ are terms.

Definition 2.4 (Well-Formed Formula).

well-formed formula

We say λ is a *well-formed formula*—often abbreviated *wff*—if it satisfies the following recurrence.

1. λ is an atomic formula.
2. $\lambda = \neg(\varphi)$, where φ is a *wff*.
3. $\lambda = (\varphi) \wedge (\psi)$, where φ and ψ are *wff*.
4. $\lambda = (\varphi) \vee (\psi)$, where φ and ψ are *wff*.
5. $\lambda = (\varphi) \to (\psi)$, where φ and ψ are *wff*.
6. $\lambda = (\varphi) \leftrightarrow (\psi)$, where φ and ψ are *wff*.
7. $\lambda = \forall x(\varphi)$, where φ is a *wff*.
8. $\lambda = \exists x(\varphi)$, where φ is a *wff*.

sentence

A well-formed formula with no free variables is called a *sentence in the first-order logic*. Looking at the above definitions, a *wff* that has no free variables will boil down to a *proposition*, meaning it will have a definite, unambiguous truth value. Sentences will be our primary mode for expressing conjectures, theorems, and proofs.

2.2 *Rules of Inference*

UNIVERSAL INTRO.	$\varphi(t)$ for an arbitrary t	$\vdash \forall x(\varphi(x))$	If we know $\varphi(t)$ and t is an *arbitrary* term, then we can say $\forall x(\varphi(x))$.
UNIVERSAL ELIM.	$\forall x(\varphi(x))$	$\vdash \varphi(t)$ for *any* term t	If we have $\forall x(\varphi(x))$, then we can pick *any* t and say $\varphi(t)$.
EXISTENTIAL INTRO.	$\varphi(t)$ for a particular t	$\vdash \exists x(\varphi(x))$	If we know $\varphi(t)$ for a *specific* term t, then we can say $\exists x(\varphi(x))$.
EXISTENTIAL ELIM.	$\exists x(\varphi(x))$	$\vdash \varphi(t)$ for a *new term* t	If we have $\exists x(\varphi(x))$, then we have $\varphi(t)$ for some t *that has not yet appeared*.

Table 2.1: The rules of inference for quantified expressions involving predicates. Note that the *"new term"* referred to by *existential elimination* must be a symbol that has not yet appeared in your proof.

When we were building the propositional logic, we first defined a *syntax* for our logic by introducing the logical connectives and some other special symbols; we then gave it an *algebraic semantics* when we introduced the equivalence axioms and the rules of inference. Now that we are augmenting our language with *terms, predicates, and quantifiers,*

we have a similar need to establish semantics for interpreting our new symbols. We introduce these rules in table 2.1. In addition, we have three important theorems involving quantified expressions, each containing a *universal* fragment and an *existential* fragment. This first theorem establishes a form of *De Morgan duality* between the $\forall$ and $\exists$ quantifiers: *negating* a quantified sentence is equivalent to quantifying the *negated* sentence using the *other* quantifier.

Theorem 2.1 (Negation of Quantifiers).
If φ is a predicate of at most one free variable, these equivalences hold.

$$\neg\forall x\big(\varphi(x)\big) \equiv \exists x\big(\neg\varphi(x)\big) \qquad \neg\exists x\big(\varphi(x)\big) \equiv \forall x\big(\neg\varphi(x)\big)$$

定理

The next theorem illustrates a sort of *distributive law* for quantifiers. Be sure to *pay careful attention to the parentheses* in the following theorem.

Theorem 2.2 (Distribution of Quantifiers).
Let φ be a predicate of at most one free variable and p be a proposition. The four equivalences below are then satisfied; mind the parentheses.

$$\forall x\big(\varphi(x)\big) \wedge p \equiv \forall x\big(\varphi(x) \wedge p\big) \qquad \exists x\big(\varphi(x)\big) \wedge p \equiv \exists x\big(\varphi(x) \wedge p\big)$$
$$\forall x\big(\varphi(x)\big) \vee p \equiv \forall x\big(\varphi(x) \vee p\big) \qquad \exists x\big(\varphi(x)\big) \vee p \equiv \exists x\big(\varphi(x) \vee p\big)$$

Further, if ψ is also a predicate with at most one free variable and t is a term, then the following four one-way inferences hold.

$$\forall x\big(\varphi(x) \wedge \psi(x)\big) \vdash \forall x\big(\varphi(x)\big) \wedge \psi(t) \qquad \exists x\big(\varphi(x)\big) \wedge \psi(t) \vdash \exists x\big(\varphi(x) \wedge \psi(x)\big)$$
$$\forall x\big(\varphi(x) \vee \psi(x)\big) \vdash \forall x\big(\varphi(x)\big) \vee \psi(t) \qquad \exists x\big(\varphi(x)\big) \vee \psi(t) \vdash \exists x\big(\varphi(x) \vee \psi(x)\big)$$

However, those inferences above are *not* equivalences, as shown below.

$$\forall x\big(\varphi(x)\big) \wedge \psi(t) \nvdash \forall x\big(\varphi(x) \wedge \psi(x)\big) \qquad \exists x\big(\varphi(x) \wedge \psi(x)\big) \nvdash \exists x\big(\varphi(x)\big) \wedge \psi(t)$$
$$\forall x\big(\varphi(x)\big) \vee \psi(t) \nvdash \forall x\big(\varphi(x) \vee \psi(x)\big) \qquad \exists x\big(\varphi(x) \vee \psi(x)\big) \nvdash \exists x\big(\varphi(x)\big) \vee \psi(t)$$

Finally, the following four equivalences hold for conditional statements.

$$\forall x\big(\varphi(x) \to p\big) \equiv \exists x\big(\varphi(x)\big) \to p \qquad \forall x\big(p \to \varphi(x)\big) \equiv p \to \forall x\big(\varphi(x)\big)$$
$$\exists x\big(\varphi(x) \to p\big) \equiv \forall x\big(\varphi(x)\big) \to p \qquad \exists x\big(p \to \varphi(x)\big) \equiv p \to \exists x\big(\varphi(x)\big)$$

定理

The third and final theorem concerns the *order of quantifiers*, pointing out the important result that *quantifiers don't necessarily commute with each other*.

Theorem 2.3 (Quantifier Shift).
If φ is a predicate of at most two free variables, then the following hold.

$$\forall x\forall y\big(\varphi(x,y)\big) \equiv \forall y\forall x\big(\varphi(x,y)\big) \qquad \exists x\exists y\big(\varphi(x,y)\big) \equiv \exists y\exists x\big(\varphi(x,y)\big)$$
$$\forall x\exists y\big(\varphi(x,y)\big) \nvdash \exists x\forall y\big(\varphi(x,y)\big) \qquad \exists x\forall y\big(\varphi(x,y)\big) \vdash \forall x\exists y\big(\varphi(x,y)\big)$$

定理

2.3 *The Art of Writing Proofs*

The way approach a proof of a statement principally depends on the *form* of the what we're trying to prove. Depending on what the statement *looks* like, a valid proof may be allowed to take certain liberties or be required to satisfy certain constraints. We will end this chapter with some words of advice for writing proofs based on the rules of inference we have established and the semantic interpretation we have attached to our various logical symbols. Since propositions and sentences in the first-order logic are *recursive* constructions, the first thing we should do when presented a statement to prove is to *recursively* analyze its *form.*

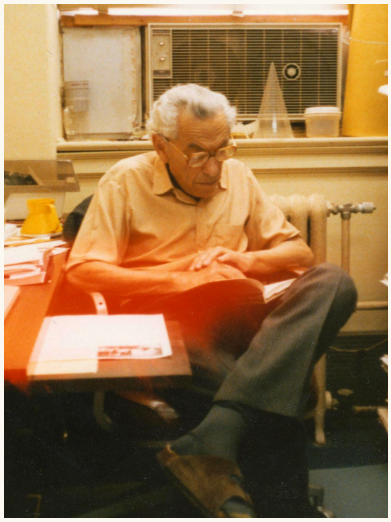

Figure 2.3: *"The purpose of life is to prove and to conjecture."* – Paul Erdős

Quantified Formulæ

If we are trying to prove a statement like $\forall x(\varphi(x))$, we can *check* $\varphi(t)$ *for all possible values of* t. This is usually not possible, as our universe of discourse often contains infinitely many objects. The natural alternative is to *introduce an arbitrary term* t and, without making any assumptions about t, to show that t satisfies φ. If we manage to do this without relying on any details pertaining to t specifically, then our argument *will generalize universally.* On the other hand, to prove a statement of the form $\exists x(\varphi(x))$, the task is to *find a specific object* t that we can prove satisfies φ. Existential claims are often the *most difficult* kind to prove because there is, generally, no clear strategy for *how* t should be found.

Figure 2.4: *"Another roof, another proof."* – Paul Erdős

Conditional Statements

Suppose we have a statement we want to prove that takes the form of a conditional $p \to q$. These are *by far the most common* kinds of statements we will be interested in proving. This involves showing we can derive q from p, so we first *assume* p in order to get to q. After assuming p is the case, we can think of how to derive q based on its *form* by again going through this analysis. Alternatively, instead of showing $p \to q$ directly, we can always think to prove $\neg q \to \neg p$ and apply our knowledge that a conditional statement is always equivalent to its contrapositive.

Junctions

Statements that look like $p \wedge q$ are relatively straight-forward: we have to show that *both* p *and* q are true. Similarly, showing $p \vee q$ requires deriving *one of either* p *or* q, but we are free to choose which one to pursue. Naturally, this will depend on what forms p and q take.

Nonconstructive Proofs

When, in the course of human events, it becomes necessary for one people to encounter a *contradiction*, a decent respect to the opinions of mankind requires that they should *reject the assumptions* that impelled them there. What we mean by this is: if you are ever stuck and feeling like a proposition p that you feel is insurmountable—for which you can see no way to make progress—try *assuming* $\neg p$ and seeing what happens. If this leads you to a *contradiction,* then you can invoke *reductio ad absurdum* and conclude p, washing your hands of the situation.