Notes on Proofs, Induction, Set Theory, and Foundations: A Companion to *The Logic Book*

Emerson P. Doyle*
Huron University College

September 7, 2017
v1.2

Abstract

The present set of notes have been developed over the years that I have taught Philosophy 2250 *Introduction to Logic* and Philosophy 2020 *Basic Logic* at the University of Western Ontario and the affiliated Huron University College. It is customary at Western to teach introductory logic using Bergmann, Moor, and Nelson's *The Logic Book*, published by McGraw-Hill and now in its sixth edition. The primary advantage of this text is its wealth of exercises with solutions, allowing students ample practice. However, the book has a habit of introducing certain conceptual, historical, and mathematical ideas without the explanation and background necessary to appreciate their importance or interconnections. Paradigmatic of this tendency is the mathematical-style proof of the relationship between an argument and its corresponding material conditional that appears without warning in §3.6. The text switches suddenly to a mathematical style in order to prove the result. Indeed, there is little to no discussion at all regarding the specifics or procedures of reading or producing mathematical proofs (even in the chapters on meta-theory, which include several exercises asking students to give proofs of important results!). Given that many students of introductory logic are encountering technical material for the first time, this situation is far from ideal. Thus these notes act to supplement and deepen discussions in the text technically, pedagogically, and historically.

Chapters are mostly self-contained, although they are designed to be read in order as students encounter the appropriate material in the text. All page numbers and references to "the text" or "our text" refer to the sixth edition of *The Logic Book*. Please contact the author with corrections or suggestions for improvement. Also feel free to distribute and use this material as you see fit, but please attribute the author where appropriate.

*These notes and some of the exercises owe substantially to a series of notes written by Nicolas Fillion for his own introductory logic course held at Western during the summer term of 2010.

Contents

Chapter 1

Set Theory

1.1 Sets and Objects

As we all know, there are many different types of English sentences. Some sentences—those dealt with in our text—have a Truth-Value. These are the sentences with which logic is concerned. In *SL*, such sentences are symbolized in one of two ways:

1. As compound sentences—if they have a Truth-Functional Connective expressible as $\sim, \&, \vee, \supset, \equiv$, or a combination of them;
2. As atomic sentences, denoted A, B, C,... if they are not compound.

Some sentences have a further character that will interest us in set theory, e.g.,

- Mario is a plumber;
- 3 is a prime number;
- Mammals are animals.

Each of these sentences suggests that one thing (or collection of things) *belongs to*, or is *a member of*, some other collection of things.

This suggests a fundamental distinction between two kinds of things, which is the conceptual heart of set theory:

1. *Objects*: denoted $a, b, c, \ldots$ (lowercase Roman letters);
2. *Sets*: denoted $\Gamma, \Delta, \Theta, \Omega, \ldots$ (uppercase Greek letters).[1]

Sets are collections of objects. The crew of the starship Enterprise, the planets of our solar system, the natural numbers, and the buildings on our campus are all examples

[1] In most books, including philosophical works, sets are denoted by $A, B, C, \ldots$, but in our textbook these are reserved as sentence letters. Thus we follow the text and use uppercase Greek letters to denote sets. Note: 'Γ' = *gamma*, 'Δ' = *delta*, 'Θ' = *theta*, 'Ω' = *omega*, 'Φ' = *phi*, and 'Ψ' = *psi*.

of sets. The individual objects that compose sets are said to be *members of*, or *elements of*, the set. For example, Spock is a member of the set of the crew of the starship Enterprise, Jupiter is an element of the set of planets in our solar system, 5 is a member of the set of natural numbers, and Talbot College is a member of the set of buildings on Western's campus. Sometimes we simply use the verb 'to be' in order to indicate set membership, as in 'Aristotle is a philosopher'. We indicate that an object a is a member of a set Γ by writing:

$$a \in \Gamma.$$

To indicate that a is not a member of Γ, we write:

$$a \notin \Gamma \quad \Leftrightarrow \quad \sim(a \in \Gamma).$$

Notice that, from the point of view of sentential logic, we are treating membership claims in a similar way to atomic sentences. Suppose a = Aristotle, and Γ is the set of all philosophers, then in *SL* we might symbolize the sentence that 'Aristotle is a philosopher' as 'A'. The language of set theory allows us to express this claim using a more subtle notation, but *when talking about set theory* we will treat such constructions as atomic sentences in *SL*, and they will behave in the same manner.[2]

As the text mentions, sets can be described in two ways:

1. *Extensionally*: an extensional description of a set consists in a list of the names of its members between curly braces;
2. *Intensionally*: an intensional[3] description of a set consists in giving a defining property for the elements of the set.

So, for example, we might define two sets thusly:

$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$
$\{\text{Mercury}, \text{Venus}, \text{Earth}, \text{Mars}, \text{Jupiter}, \text{Saturn}, \text{Uranus}, \text{Neptune}\}.$

The first is the set of one-digit decimal numbers, the second the set of planets in our solar system. Sometimes, when a set has a infinite number of elements, we end the list with '. . . ' as for the set of natural numbers:

$$\{0, 1, 2, 3, 4, 5 \ldots\}$$

[2] So technically, the use of parenthesis around membership claims is unnecessary. To extend the example above, we can symbolize 'Aristotle is not a philosopher' in *SL* as '∼A', 'Plato is a philosopher' as 'P', and their conjunction as '∼A & P'. Similarly, in set-theoretic notation we can write '$\sim a \in \Gamma \;\&\; p \in \Gamma$' without ambiguity. Or more concisely, '$a \notin \Gamma \;\&\; p \in \Gamma$'. Again, parentheses around membership claims are unnecessary because we treat them as atomic sentences. However, in these notes we will often enclose such expressions in parentheses to help with readability, e.g., '$(a \notin \Gamma) \;\&\; (p \in \Gamma)$'. Note that most texts and papers do not do this. We will also play fast-and-loose with the distinction between use and mention—not always enclosing expressions in single quotes. And similarly with the distinction between meta/object-language. This is to enhance readability, and such distinctions should be recoverable from the context.

[3] Notice the word *intensional*, with an 's', not *intentional*.

But this is a bit ambiguous, since there are many ways to continue the list. It is for this reason that we can also define sets *intensionally.* we write:

$$\{x : Px\}.$$

Here, we use 'x' as a variable, denoting any arbitrary object; 'P' denotes a property. The ':' can be read as 'such that'. So if $P =$ 'is a prime number', then $\{x : Px\}$ is the set of things that are prime numbers. When we have more logical and mathematical tools we usually write the property in symbols, as a formula. This formula is called a *characteristic function* or a *characteristic property.* But for the time being, we will write the defining property in words if we don't know how to write it in symbols. So for $\{x : Px\}$, we will write:

$$\{x : x \text{ is a prime number}\}.$$

When we want to talk about an arbitrary element of a set, we will use the end of the alphabet (with or without subscripts), $x, y, z \ldots$. As we have seen, individual known (i.e., named) elements use $a, b, c, \ldots$ (with or without subscripts).

There are a few other basic facts about set membership that we must now introduce. Let us consider the set $\Gamma = \{a, b, c\}$. This means that $a \in \Gamma, b \in \Gamma, c \in \Gamma$ and nothing else is an element of the set Γ. In logical terms, we can write:

$$x \in \Gamma \quad \equiv \quad ((x = a \vee x = b) \vee x = c).$$

It is also important to know that the order in which elements are listed in a set does not matter. So:

$$\{1, 3, 5, 7, 9\} = \{9, 7, 5, 3, 1\}.$$

And as the textbook notes, repetition is superfluous, so that:

$$\{\text{Kirk, Spock, Bones, Uhura, Spock}\} = \{\text{Kirk, Spock, Bones, Uhura}\}.$$

The reason for these two facts lies in the way we formally define certain relationships between sets, which we will do in the next section.

Finally, we allow for the possibility of sets being members of other sets. So we can write expressions such as $\Gamma = \{2, \{3\}, 4\}$. Here, Γ contains two objects, 2 and 4, plus one set, $\{3\}$. It is very important to recognize that even if $\{3\} \in \Gamma$, it does not follow that $3 \in \Gamma$, since an object is *not* the same thing as the set containing only that object. In the above case we actually have that $3 \notin \Gamma$. We allow sets to be members of sets in order to develop the concepts of ordered pairs and functions, to be examined below.

Exercise #1.

1. Which of the following are true?

   (a) $2 \in \{1, 2, 3\}$
   (b) $2 \in \{Paris, 4, Fido\}$
   (c) $\{1, 2\} \in \{\{1, 2, 3\}, \{1, 3\}, 1, 2\}$
   (d) {Elizabeth II} $\in$ $\{x : x$ is a king or queen of England$\}$

2. Find the characteristic property to describe these sets intensionally:

   (a) {Sweden, Finland, Norway, Denmark}
   (b) {Manitoba, Saskatchewan, Alberta}
   (c) $\{2, 4, 6, 8, 10, 12, 14, \ldots\}$

3. Describe these sets extensionally:

   (a) $\{x : x$ is a province of Canada$\}$
   (b) $\{x : x$ is a North-American country$\}$
   (c) $\{x : x$ is a possible truth-value for a sentence studied in *SL* and *PL*$\}$

1.2 Set Operations in Logical Terms

Now that we have a working understanding of sets, we can begin to manipulate them in more rigorous and complex ways. Notice in the following eight definitions, that on the left-hand side we will often introduce new notation, to be used between sets, and on the right-hand side we define the use of this new notation using familiar expressions (often *SL* symbols). This is typical of mathematical definitions. This leads us to our first quasi-formal definition of set theory:

Definition 1 (Set Equality)**.** *Let* Γ *and* Δ *be sets, then:*

$$\Gamma = \Delta \quad \text{iff, for all } x,\ (x \in \Gamma) \equiv (x \in \Delta)$$

That is to say, two sets are equal when they have the same membership conditions.

Here we have introduced the equality sign '=' as it is used *between two sets* in terms of our usual notion of truth-functional equivalence *between two atomic sentences*. What this definition tells us is that a set is fully determined by its elements. This is why the order in which the elements are listed doesn't matter, and why repetition in the extensional definition of a set is superfluous. This definition also leads to the interesting result that two sets can be equal even though they may be intensionally defined by radically different characteristic properties. For example, because it happens to be the case that animals have kidneys iff they have hearts, we have:

$$\{x : x \text{ is an animal with kidneys}\} = \{x : x \text{ is an animal with a heart}\}.$$ [4]

Therefore, to show that two sets are equal is the same as to show that they have equivalent characteristic properties.

Now consider the set of trees and the set of once living things. Clearly, every member of the set of trees is also a member of the set of once living things. Logically, we can say that an object's being a member of the set of trees *implies* that it is also a member of the set of once living things. In this case, we say that the set of trees *is a subset of* the set of once living things, or that the set of trees *is included in* the set of once living things. We write this as $\Gamma \subseteq \Delta$. So our second definition:

Definition 2 (Set Inclusion)**.** *Let* Γ *and* Δ *be sets, then:*

$$\Gamma \subseteq \Delta \quad \text{iff} \quad (x \in \Gamma) \supset (x \in \Delta)$$

That is to say, one set is included in another when membership in the first implies membership in the second.

Again, notice that '$\supset$' is a logical symbol and '$\subseteq$' is a set-theoretical symbol. This means that in sentential logic, if we have $(x \in \Gamma) \supset (x \in \Delta)$ and $(x \in \Delta) \supset (x \in \Gamma)$, then we have $(x \in \Gamma) \equiv (x \in \Delta)$ by the fact that a material bi-conditional between two sentences is just the conjunction of two material conditionals between them (remember the truth-tables!). Consequently, by our above definition of Set Equality, in such cases we have that $\Delta = \Gamma$. To make this relationship between set equality and inclusion even more explicit:

$$\Delta = \Gamma \quad \Leftrightarrow \quad (\Delta \subseteq \Gamma) \;\&\; (\Gamma \subseteq \Delta).$$

This fact will allow us to prove that two sets described differently are in fact equal, as observed above.

The concept of *Set Union* is introduced but briefly in the text on page 107. We now tackle it more thoroughly in our quasi-formal way, again using logical notation to define our set-theoretical symbols:

Definition 3 (Set Union)**.** *Let* Γ *and* Δ *be sets, then:*

$$\Gamma \cup \Delta = \{x : (x \in \Gamma) \vee (x \in \Delta)\}$$

That is to say, the union of two sets is itself a set with the characteristic property in which each object in the union is either a member of the first set, or a member of the second set, or both (recall '$\vee$' is the inclusive or).

[4] For those with a background in analytic philosophy, W.V.O. Quine often uses this example.

Notice that this definition is different than the first two, in the sense that it *defines a new set* by specifying a characteristic property. The first two definitions instead defined a *relationship* between sets. The characteristic property is determined using the logical notation '$\vee$'. So what the union operator does is to take take the members of two (or more)[5] sets and collects them into a *new set* that we call the *union* of the original sets. We can denote this set using the union symbol, e.g., '$\Gamma \cup \Delta$', or we can give it a brand new name, say another uppercase Greek letter. Our new set contains as members all and only the members of the original sets.

The text doesn't bother to define the notion of *Set Intersection*. It works in much the same way as set union, but as a conjunction rather than a disjunction:

Definition 4 (Set Intersection)**.** *Let* Γ *and* Δ *be sets, then:*

$$\Gamma \cap \Delta = \{x : (x \in \Gamma) \ \& \ (x \in \Delta)\}$$

That is to say, the intersection of two sets is itself a set with the characteristic property in which each object in the intersection is a member of both original sets.

Again, we allow for the creation of sets that are the intersection of more than two sets. Another important notion is of *Set Compliment*:

Definition 5 (Set Compliment)**.** *Let* Γ *and* Δ *be sets, then:*

$$\overline{\Gamma} = \{x : x \notin \Gamma\}$$

That is to say, the compliment of a set is itself a set that contains everything not contained in the original set.

We read '$\overline{\Gamma}$' as 'gamma-bar', or simply as 'the compliment of Γ'. The compliment of a set contains *absolutely everything* not contained in the original set. So if we take the set $\Phi = \{2, 3\}$, then $\overline{\Phi}$ contains the rest of the natural numbers, Captain Kirk, your textbook, all of us, every other set, etc. It is a large set indeed!

Our last useful set operation is:

Definition 6 (Set Difference)**.** *Let* Γ *and* Δ *be sets, then:*

$$\Gamma \setminus \Delta = \{x : (x \in \Gamma) \ \& \ (x \notin \Delta)\}$$

That is to say, the difference between two sets is itself a set that contains all of the members of the first set that are not also members of the second.

[5]Although our quasi-formal definition doesn't actually specify this because it would be more complicated to read, we will take it as given. For example we can have $\Psi = \Gamma \cup \Delta \cup \Omega$ as the union of those three sets, which is the set Ψ having as members the members of our original three sets.

You might be asking yourself why such an operation would be useful—don't we just end up with the set we started with?! The way to think about set difference is as *subtracting out* various sets from another set. Consider the following two sets:

$$\Gamma = \{\text{Plato, Aristotle, Descartes, Hume}\}$$
$$\Delta = \{\text{Descartes, Hume}\}$$

The difference $\Gamma \setminus \Delta = \{\text{Plato, Aristotle}\}$, the set containing all the members of Γ minus the members of Δ.

To complete our list of definitions, we outright define two special sets, one of which we are already very familiar:

Definition 7 (The Empty Set, $\varnothing$)**.** *The set containing no members. That is:*

$$\textit{For any } x,\ x \notin \varnothing$$

Definition 8 (The Universal Set, Λ)**.** *The set containing everything. That is:*

$$\textit{For any } x,\ x \in \Lambda$$

We reserve these symbols for these special sets (so the Greek letter lambda isn't used to indicate just any old set we want). There is only one of each, or in other words, there isn't more than one empty set, or universal set. You should be able to figure out why this is according to one of the definitions above.

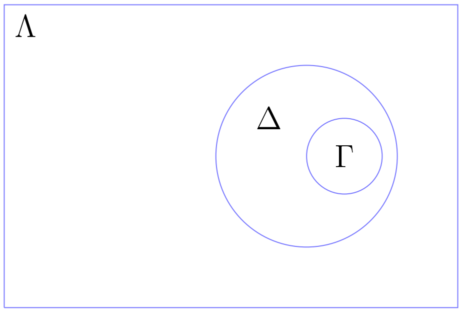
To summarize, we have provided six set-theoretic definitions in logical terms. Set equality in terms of truth-functional equivalence, set inclusion in terms of the material conditional, set union in terms of disjunction, set intersection in terms of conjunction, set compliment in terms of negation (because non-membership was specified in terms of negation), and set difference in terms of negation and conjunction. We also defined two additional sets, the universal and empty sets.

1.3 Venn Diagrams

We can represent the above set-theoretic relations and operations graphically, using so-called *Venn Diagrams*. They are so named because they were introduced in their modern form by John Venn in an 1880 article (although he wasn't so egotistical as to call them 'Venn Diagrams', he used the term 'Eulerian Circles' after 18th century mathemagician Leonhard Euler). Although Venn was certainly not the first mathematician to use Venn-like diagrams—in fact Euler introduced something similar—his article made their use much more common.

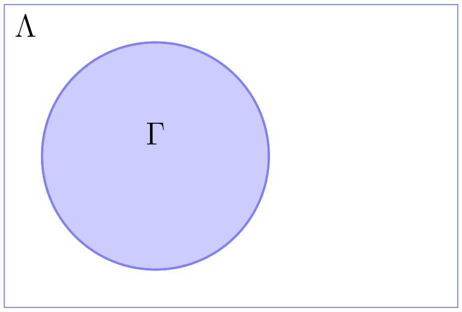
First we will see how to construct a Venn diagram. Begin with a rectangle representing everything (this is the universal set Λ). We then identify the region we want to characterize by colouring it:

The above diagram represents the fact that x could be anything, since it could be anywhere in the diagram. We can depict individual sets with circles. The following diagram represents two sets, one of which is a subset of the other (i.e., $\Gamma \subseteq \Delta$):

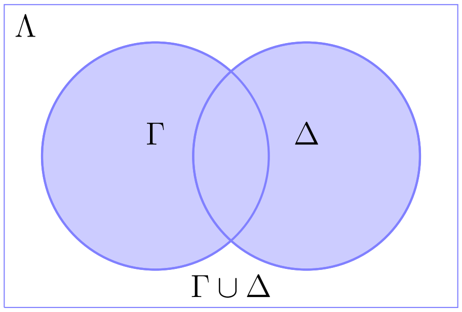


Notice that all sets are included in (are subsets of) the universal set, as it should be.

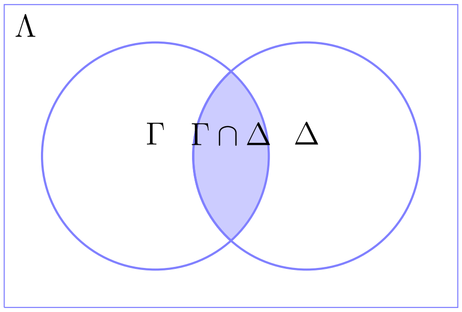
Now to membership claims. The following diagram represents that $x \in \Gamma$:



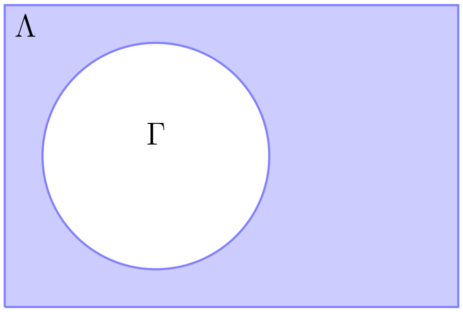
Given the condition $x \in \Gamma$, x could be anywhere in the circle, and so we colour that area to represent this assertion. The next diagram represents $x \in (\Gamma \cup \Delta)$:



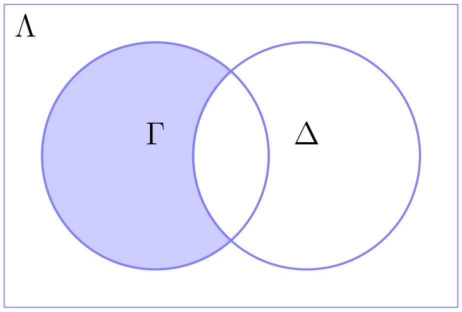
While the following represents $x \in (\Gamma \cap \Delta)$:



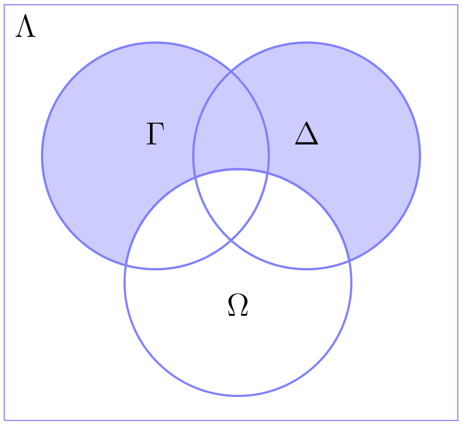
Next we have the representation of $x \in \overline{\Gamma}$:



And then $x \in (\Gamma \setminus \Delta)$:



If a certain object x is a member of $(\Delta \cup \Gamma) \setminus \Omega$, you can represent this as:



In this case, x satisfies $(x \in (\Delta \cup \Gamma))\ \&\ (x \notin \Omega)$, in other words:

$$((x \in \Delta) \vee (x \in \Gamma))\ \&\ (x \notin \Omega).$$

This is a useful intuitive way to reason about simple relationships between sets which contain objects only. It may be of use as you work your way through Chapters 7 and 8 of the text especially (the semantics of *PL(E)*, are cast in set theory explicitly). However, do keep in mind that this mode of thinking about sets breaks down for sets of sets, since Venn diagrams as presented here provide no way to represent sets themselves being members of other sets.[6] We also can't use the method presented here for membership claims between more than three sets at a time. Still, it's a useful heuristic for quickly getting a handle on set-theoretic relationships.

Exercise #2.

1. Determine which of the following are true, for any sets Γ, Δ, Ω. If a statement is false, give examples of Γ, Δ, Ω that make the statement false.

   (a) If $\Gamma = \Delta$ and $\Delta = \Omega$, then $\Gamma = \Omega$.
   (b) If $\Gamma \in \Delta$ and $\Delta \in \Omega$, then $\Gamma \in \Omega$.
   (c) If $\Gamma \subseteq \Delta$ and $\Delta \subseteq \Omega$, then $\Gamma \subseteq \Omega$.
   (d) If $\Gamma = \Delta$ and $\Delta \in \Omega$, then $\Gamma \in \Omega$.
   (e) If $\Gamma \in \Delta$ and $\Delta = \Omega$, then $\Gamma \in \Omega$.

2. In terms of set theory, what is wrong with the following argument:

   1. Socrates is a man.
   2. Men are numerous.

   $\therefore$ Socrates is numerous.

3. Which of the following are true for any set Γ, Δ, Ω?

   (a) $[(\Gamma \notin \Delta)\ \&\ (\Delta \notin \Omega)] \supset (\Gamma \notin \Omega)$
   (b) $[(\Gamma \neq \Delta)\ \&\ (\Delta \neq \Omega)] \supset (\Gamma \neq \Omega)$
   (c) $[(\Gamma \in \Delta)\ \&\ \sim (\Delta \subseteq \Omega)] \supset (\Gamma \notin \Omega)$

[6] This is something that has been left implicit so far, but it is important to keep in mind: The notions of *Membership* and *Inclusion* are different. Objects and sets can be members of other sets, and we take set membership to be a fundamental, undefined notion of set theory. Only sets can be included in other sets, and this is a defined notion—defined in terms of membership recall!

1.4 Some Constructions in Terms of Sets

Now that we have all of the concepts and notation we need, we can get down to the business of actually using set theory to construct some useful mathematical objects.

1.4.1 Properties and Relations Understood with Sets (n-tuples)

We saw above that the order in which the elements of a set are listed is not important. However, there are many cases where order does matter. For example, if you're arranging the natural numbers from least to greatest, 2 had better come before 100 or you've done something wrong. As another example, if a group of you are going to enter a haunted mansion, you may want to be the last to go in rather than the first. For this reason, we introduce yet another new notation, which the text introduces on page 331 for discussing the interpretation of predicates in *PL(E)*. When the order matters, instead of writing $\{a, b\}$ (with curly braces), we will write $\langle a, b\rangle$ (in angle brackets).[7] Thus, although $\{a, b\} = \{b, a\}$, it is usually the case that:

$$\langle a, b\rangle \neq \langle b, a\rangle$$

unless $a = b$. For this reason, we call such an thing an *Ordered Pair*.

Definition 9. *We define the Equality of two ordered pairs $\langle a, b\rangle$ and $\langle c, d\rangle$ as follows:*

$$\langle a, b\rangle = \langle c, d\rangle \quad \Leftrightarrow \quad (a = c) \ \& \ (b = d)$$

That is to say, two ordered pairs are equal just in case their members are the same and their members are in the same order.

The notion of an ordered pair is extremely important for logic and mathematics, because it allows us to represent *Relations* and *Functions* in a set-theoretic way. We'll come back to functions below. For now, we can think of a *Relation*, R as expressing some relationship between two or more sets of objects. For example:

- x is the mother of y;
- x loves y;
- x is the flavour of y;

are all examples of *Binary Relations*. The relations:

- x is is taller than y but shorter than z;
- x is between y and z

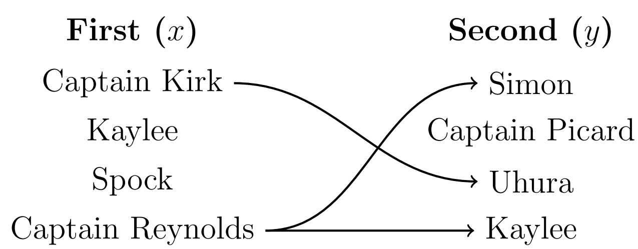
[7] Notice that we do not in what follows actually *define* ordered pairs in terms of sets, instead, we merely define a condition that any such definition must satisfy. The standard definition, for any objects a, b, is: $\langle a, b\rangle \Leftrightarrow \{\{a\}, \{a, b\}\}$.

are both *Ternary Relations*, because they express a relationship between three objects rather than two. We can of course generalize this to so-called n-ary relations that relate any n numbers of objects.

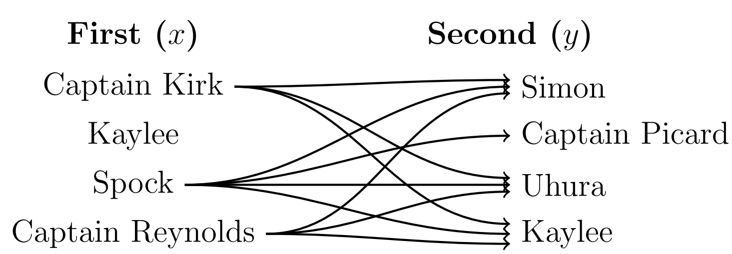
A relation itself can be thought of as really just big lists of objects (two lists for a binary relation, three for a ternary relation, etc.) between which are drawn arrows which correlate the objects according to the relation in question. So binary relations concern all possible arrows from each object x on our first list to each object y on a second list (the same objects might be on both lists of course). And a particular binary relation Cxy (say, 'x is the Captain of y') will pick out *some particular set* of arrows that represents the correlations between the objects as dictated by C. So if we take our lists to be as follows:

First (x)	**Second** (y)
Captain Kirk	Jane
Kaylee	Captain Picard
Spock	Uhura
Captain Reynolds	Kaylee

Then we can represent the relation Cxy as this particular set of arrows:



Some other relation, say Oxy = 'x is older than y', will be represented by another set of arrows:



You can see that this sort of representation quickly becomes unmanageable![8] So instead, we can represent our lists as *sets*, and the arrows as *a set of ordered pairs of*

[8]If you aren't familiar with the characters on the shows *Star Trek* or *Firefly*, this example probably won't make much sense, but hopefully the point is still clear. Note I'm guessing on some of the ages.

elements of the list sets. So take the set $\Delta = \{$Captain Kirk, Kaylee, Spock, Captain Reynolds$\}$, and the set $\Gamma = \{$Simon, Captain Picard, Uhura, Kaylee$\}$. Then the relation Cxy is represented by the set:

$$C = \{\langle \text{Cpt. Kirk, Uhura} \rangle, \langle \text{Cpt. Reynolds, Simon} \rangle, \langle \text{Cpt. Reynolds, Kaylee} \rangle\}$$

This is a set of ordered pairs (which we have named 'C' rather than using a Greek letter), which associates each captain in the first list with their crew from the second list. The relation Oxy will be a longer set of ordered pairs, but hopefully you see how we're picking out a *subset* of all possible arrows with each relation.

Now we will consider a more practical example. Consider the relation '$\leq$' (less-than-or-equal-to) over the natural numbers. It is true that $0 \leq 0, 0 \leq 1, 0 \leq 1000, 5 \leq 9$, etc., but false that $3 \leq 2$, for example. We can again represent this relation by *a set of ordered pairs*—namely all the pairs of numbers that satisfy the relation '$\leq$'. So given our example, we have:

$$\leq\ = \{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 0, 2\rangle, \dots\}$$

The set (which we have named '$\leq$' rather than using a Greek letter) will thus contain every ordered pair of natural numbers wherein the first element of the pair is less than or equal to the second element of the pair. Notice that in this case, our two 'lists' are exactly the same—the list of all the natural numbers (the counting numbers including zero), which mathematicians usually write as $\mathbb{N}$

What we have just done is taken a complex notion (the relation of being less-than-or-equal-to with regard to the natural numbers) and *reduced* it to a particular set of ordered pairs. Cool! This is one of the reasons mathematicians and philosophers like set theory. It allows us to think about a very wide range of complex notions using a few very basic concepts and some clever definitions.

In general, *we will consider binary relations to be sets of ordered pairs*. It is of course possible to extend this idea of an ordered set of elements beyond ordered pairs. For n elements, we call this a n-tuple (see page 331 in the text), and we write:

$$\langle a_1, a_2, a_3, \dots, a_{n-1}, a_n \rangle$$

So for example, we sometimes deal with ternary relations such as those mentioned above (like 'betweenness'). *We will treat ternary relations as sets of ordered triples.* And in general, *an n-ary relation will be treated as a set of n-tuples.* In our original haunted mansion example, the order in which the people enter the haunted mansion will be represented by an ordered n-tuple, where n is the number of people in the group, and their order is represented by their ordering in the n-tuple.

Using n-tuples, we are now able to define an important concept, the concept of the *Cartesian Product* of a collection of sets:

Definition 10 (Cartesian Product)**.** *Let* $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$ *be sets. The Cartesian Product of these* n *sets is the new set*

$$\Gamma_1 \times \Gamma_2 \times \ldots \times \Gamma_n = \{\langle x_1, x_2, \ldots, x_n\rangle : x_1 \in \Gamma_1 \ \& \ x_2 \in \Gamma_2 \ \& \ \ldots \ \& \ x_n \in \Gamma_n\}$$

That is to say, the Cartesian product of a collection of sets is the set of all possible n*-tuples of the members of the original sets.*

So suppose that we want to talk again about the relation '$\leq$' over the natural numbers. The set of ordered pairs that represents this relation is then a particular subset of $\mathbb{N} \times \mathbb{N}$.[9] In other words, the relation '$\leq$' is certain collection of ordered pairs (a subset) out of all possible ordered pairs of natural numbers. To go back to our "arrow talk", $\mathbb{N} \times \mathbb{N}$ represents every arrow from each natural number to each other one. But we want to focus only on those arrows that correlate the natural numbers in terms of the relation '$\leq$', and this will be some subset of all those arrows. We represent this subset as a set of ordered pairs of natural numbers—all pairs where the first number in each pair is less than or equal to the second number in each pair. The relation '$>$' would be another, different subset of $\mathbb{N} \times \mathbb{N}$. Usually, if $\Gamma = \Gamma_1 = \Gamma_2 = \ldots = \Gamma_n$, as in the example above (where Γ and $\Gamma_1 = \mathbb{N}$) we just write Γ^n. Thus, the relation '$\leq$' is a subset of $\mathbb{N}^2$.

This might be getting a little confusing again, so an example will probably help at this point. Let's define the set Φ as:

$$\Phi = \{\text{Socrates, Plato, Aristotle, Descartes}\}$$

Then the Cartesian product Φ^2 (or in other words, $\Phi \times \Phi$) is the set

$$\begin{aligned}&\{\langle Socrates, Socrates\rangle, \langle Socrates, Plato\rangle, \langle Socrates, Aristotle\rangle,\\ &\langle Socrates, Descartes\rangle, \langle Plato, Socrates\rangle, \langle Plato, Plato, \rangle,\\ &\langle Plato, Arisotle\rangle, \langle Plato, Descartes\rangle, \langle Aristotle, Socrates\rangle,\\ &\langle Aristotle, Plato\rangle, \langle Aristotle, Aristotle\rangle, \langle Aristotle, Descartes\rangle,\\ &\langle Descartes, Socrates\rangle, \langle Descartes, Plato\rangle, \langle Descartes, Aristotle\rangle,\\ &\langle Descartes, Descartes\rangle\}\end{aligned}$$

The length of this expression is a good example of why we use abstract, shorthand symbolizations! More importantly, the set just ends up being a list of every possible ordered pair of the elements of our original sets, which is how the definition of a Cartesian product works.[10] Now, consider the binary relation Txy = 'x was the teacher of y.' So far as the set Φ is concerned, we can represent this relation as a *subset* of our set $\Phi \times \Phi$, which we will call T:

$$T = \{\langle Socrates, Plato\rangle, \langle Plato, Aristotle\rangle\}.$$

[9] We say $\mathbb{N} \times \mathbb{N}$ as 'N cross N'

[10] In case you haven't already realized, the name "Cartesian Product" comes, of course, from philosopher and mathemagician René Des<u>cartes</u>.

This is all of the ordered pairs that satisfy the binary relation Txy over the set Φ.

It is important to remember that we can always find the Cartesian product of two or more *different* sets (e.g., $\Gamma \times \Phi$), not just of the same set twice like in this example and the example of '$\leq$'. Our example with the Captains uses two different sets, for example (although the sets happen to share one member, namely, Kaylee). Finally, here are some important properties that relations can have:

1. A relation R is *Transitive* iff, for all x, y, z, if $\langle x, y\rangle \in R$ & $\langle y, z\rangle \in R$ then $\langle x, z\rangle \in R$;
2. A relation R is *Symmetric* iff, for all x, y, if $\langle x, y\rangle \in R$ then $\langle y, x\rangle \in R$;
3. A relation R is *Reflexive* iff, for all x, $\langle x, x\rangle \in R$.

These properties are discussed in §7.5 of the text, pages 317–318, with regard to the special relation of *identity*. The language *PLE* is able to express relations of identity between objects.

Exercise #3.

1. Is $\langle 2, 3\rangle = \langle 2, 4\rangle$?
2. Is $\langle 2, 3\rangle = \langle 3, 2\rangle$?
3. Consider the following relations:

$$\begin{aligned}
\Gamma &= \{\langle 1,2\rangle, \langle 5,10\rangle, \langle 2,3\rangle, \langle 1,3\rangle, \langle 4,4\rangle\} \\
\Delta &= \{\langle 1,1\rangle, \langle 2,2\rangle, \langle 3,3\rangle\} \\
\Omega &= \{\langle 1,2\rangle, \langle 2,1\rangle\} \\
\Phi &= \{\langle 1,2\rangle, \langle 2,1\rangle, \langle 1,1\rangle, \langle 2,2\rangle\}
\end{aligned}$$

   (a) What properties of reflexivity, symmetry, and transitivity does each of the relations have?
   (b) Is Δ a subset of Γ?
   (c) Is Ω a subset of Φ?
   (d) List all the members of $\Omega \setminus \Delta$.
   (e) Does $(\Delta \setminus \Omega) = (\Omega \setminus \Delta)$?

1.4.2 Functions

Functions can be thought of as just a special kind of relation. You should already be very familiar with certain functions—specifically the *Truth-Functions*. Remember that what is essential about a function is that the "output" is determined, or computed, *entirely* by the value(s) of its "inputs". If $\Omega = \{T, F\}$ (the set of truth-values), then we can define a function with the set:

$$\sim \mathbf{P} = \{\langle T, F\rangle, \langle F, T\rangle\}$$

The first element in each ordered pair is called an *argument* (the "input"), and the second is called the *value* (the "output") of the function for its corresponding argument. This function associates T with F, and F with T. You should recognize this function, which is why we have given it a special name. It is exactly this set of ordered pairs that we represent when we draw this function's characteristic truth-table.

As noted, a function is a *special kind* of binary relation. In fact, a function f is a binary relation satisfying two conditions:

1. For all $x \in \Delta$, there is at least one $y \in \Gamma$ such that $\langle x, y\rangle \in f$;
2. For all $x \in \Delta$, there is at most one $y \in \Gamma$ such that $\langle x, y\rangle \in f$.

The set Δ is called the *domain*, it is where the arguments for the function come from (our first list). The set Γ is called the *range*, and it is where the of the function's values live (the second list). Condition 1 says that *each element* of the domain is related to some element of the range. Condition 2 says that an element in the domain *cannot* be related to *more than one* element of the range. This is because we don't want functions to be ambiguous—each set of inputs should correspond to one, and only one, output. So with the exception of $\sim \mathbf{P}$, *none* of our examples of relations so far count as functions. In the case of 'Cxy', Captain Reynolds is the captain of both Kaylee and Simon. In the case of '$\leq$', the number 2 is less than or equal to an infinite number of numbers. There's nothing wrong with this, it just means that these relations aren't functions. In a function, each argument (element of list one) is associated with one, and only one, value (element of list two).

You may have realized that many mathematical functions, and all of our truth-functions except for negation, take *multiple* arguments as "input". How is this possible, given what we've said about functions so far? Well, in the same way that sets can have other sets as members, it is possible that an ordered pair contains an ordered pair as one of its elements, in which case we have, e.g., $\langle\langle a, b\rangle, c\rangle$. In such cases, our domain is a set of ordered pairs, and our range is just a set of objects. The function itself is then a set of ordered pairs, with the first element of each pair being itself an ordered pair. We can still think of functions in terms of lists with arrows. What's important is that each element of list one (whether that be a single object, an ordered pair, an n-tuple, etc.) has only one arrow to an object of list two. We also require that *every* object on list one have an arrow to an object on list two. That's it.

Exercise #4.

1. Among the following functions,

$$R_1 = \{\langle\langle T, T\rangle, T\rangle, \langle\langle T, F\rangle, F\rangle, \langle\langle F, T\rangle, T\rangle, \langle\langle F, F\rangle, T\rangle\}$$
$$R_2 = \{\langle\langle T, T\rangle, T\rangle, \langle\langle T, F\rangle, F\rangle, \langle\langle F, T\rangle, F\rangle, \langle\langle F, F\rangle, F\rangle\}$$
$$R_3 = \{\langle\langle T, T\rangle, T\rangle, \langle\langle T, F\rangle, F\rangle, \langle\langle F, T\rangle, F\rangle, \langle\langle F, F\rangle, T\rangle\}$$
$$R_4 = \{\langle\langle T, T\rangle, T\rangle, \langle\langle T, F\rangle, T\rangle, \langle\langle F, T\rangle, T\rangle, \langle\langle F, F\rangle, F\rangle\}$$

which one is Truth-Functional Conjunction, Disjunction, Material Conditional, and Equivalence?

2. Write a function in terms of ordered pairs for the following connective:

P	**Q**	**P $\star$ Q**
T	T	T
T	F	F
F	T	T
F	F	F

1.5 Conclusion

Hopefully you now have enough of an understanding of set theory to engage with most papers in philosophy and other disciplines that discuss or use it. More to the present purposes, you should have a much deeper understanding of the *set-theoretic* part of the set-theoretical semantic theories used in our textbook. This will help you greatly in understanding the proofs and discussions in chapters 6 and 11 as well as the semantics of both *SL* and *PL(E)*, which rely upon set theoretic notions.

If you're interested to investigate set theory to an even greater degree, or if you're interested in many of the philosophical questions that set theory raises—and those it helps to address—I recommend the texts cited at the end of these notes.

1.6 Solutions to Exercises

Exercise #1. 1. a) T, b) F, c) F, d) F (because {ElizabethII} is a set, not a queen). 2. a) being a Scandinavian country, b) being a province in the Prairies, c) being an even natural number (probably). 3. a) {BC, Alberta, Saskatchewan, Manitoba, Ontario, Quebec, NB, NS, PEI, NF} b) {Canada, USA, Mexico} c) {T,F}

Exercise #2. 1. a) T, b) F, c) T, d) T, e) T. 2. The first premise is a set membership claim, while the second is a set inclusion claim. But then the conclusion, which is a set membership claim, does not hold. 3. a) F, b) T, c) F.

Exercise #3. 1. No. 2. No. 3. a) Γ is transitive, Δ is reflexive, Ω is symmetric, Φ is transitive, reflexive, and symmetric. b) No. c) Yes. d) It is just Ω. e) Yes in this case, but not in general.

Exercise #4. 1. In order: Material Conditional, Conjunction, Equivalence, Disjunction. 2. $\{\langle\langle T, T\rangle, T\rangle, \langle\langle T, F\rangle, F\rangle, \langle\langle F, T\rangle, T\rangle, \langle\langle F, F\rangle, F\rangle\}$.

Chapter 2

Mathematical Proofs

2.1 Analyzing and Understanding a Proof

The text throws its first mathematical-style proof at you without announcement on page 101. Here the text proves by what seems to be a rather informal discussion that:

Theorem 1. *An argument of SL with a finite number of premises is truth-functionally valid iff its corresponding material conditional is truth-functionally true.*

We'll get to the details of this proof in a moment. First, we should pause to note that the text does a terrible job of warning us that it's introducing an entirely new set of ideas and methods! While the discussion in the text following the phrase "We now prove..." may seem at first glance to be rather confusing and undirected, it is in fact a perfectly rigorous argument, simply an informal one. Rather than casting the argument in a formal language like *SL* and then deriving the conclusion from the premises using *SD*, the text instead utilizes the typical language and methods of mathematicians, which provide us great economy over the more formalized derivations learned in logic. The text might have told us it was doing this! In any case, mathematical-style proofs are used by the text throughout §3.6 and for all of Chapters 6 and 11 without really explaining what's going on. Many of the exercises in these sections also ask you to prove things in this mathematical style. Working through this chapter of the present notes will hopefully help to bridge the gap!

What's important to recognize right out of the gate—and throughout this chapter of our notes—is that the derivation rules of *SD* and *PD* are largely *modeled* upon the mathematical proof methods that mathematicians have been using for centuries, at least as far back as Euclid in 300 BCE, and probably much earlier. And that these methods are themselves rigorizations of our everyday modes of reasoning. Keeping this correspondence in mind—between *SD* and the structure of mathematical proofs—will help us in what follows.

So back to our Theorem 1. Here is the proof from the textbook, annotated in various ways that will be explained below:

Proof. Suppose that

$$\begin{array}{c}\mathbf{P_1}\\ \vdots \\ \mathbf{P_n}\\ \mathbf{Q}\end{array}$$

is a truth-functionally valid argument of *SL*. Then there is no truth-value assignment on which $\mathbf{P_1}, \ldots, \mathbf{P_n}$ are all true and $\mathbf{Q}$ is false. Because the iterated conjunction $(\ldots\ (\mathbf{P_1}\ \&\ \mathbf{P_2})\ \&\ \ldots\ \&\ \mathbf{P_n})$ has the truth-value $\mathbf{T}$ on a truth-value assignment if and only if all of $\mathbf{P_1}, \ldots, \mathbf{P_n}$ have the truth-value $\mathbf{T}$ on that assignment, it follows that there is no truth-value assignment on which the antecedent of the corresponding material conditional, $(\ldots\ (\mathbf{P_1}\ \&\ \mathbf{P_2})\ \&\ \ldots\ \&\ \mathbf{P_n}) \supset \mathbf{Q}$, is true while the consequent is false. Thus, the material conditional is true on every truth-value assignment and is therefore truth-functionally true.

Now assume that $(\ldots\ (\mathbf{P_1}\ \&\ \mathbf{P_2})\ \&\ \ldots\ \&\ \mathbf{P_n}) \supset \mathbf{Q}$ is truth-functionally true. Then there is no truth-value assignment on which the antecedent is true and the consequent false. But the iterated conjunction is true on a truth-value assignment if and only if the sentences $\mathbf{P_1}, \ldots, \mathbf{P_n}$ are all true. So there is no truth-value assignment on which $\mathbf{P_1}, \ldots, \mathbf{P_n}$ are all true and $\mathbf{Q}$ is false; hence the argument is truth-functionally valid. ■

Again, it can look very confusing if you've never encountered a mathematical proof before. However, I've taken the liberty of underlining several expressions that should be familiar. Recall from Chapter 1 of the text that expressions such as *it follows that*, *thus*, *therefore*, *so*, and *hence* are all *Conclusion Indicators*. While expressions like *because*, *assume that*, and *suppose that* are all *Premise Indicators*. The rest of the underlined expressions are just our usual truth-functional connectives. So we see that there's actually quite a bit of logical structure to this passage after all!

Now let's get to the business of analyzing and understanding the proof. Each time you approach a new proof, you can examine it in roughly the same way. The idea is to make the text of the proof clear enough that you can see how the theorem has been proven on the basis of the arguments in the body of the proof. There's often one or more *Eureka!* moments when the author's inferences become clear to you. To facilitate getting to such moments in a more methodical way, I've broken the analysis process down into four steps, which we might think of as guidelines for what to do—we won't necessarily need to go through each step explicitly with each proof. But for our first example, we will go through each one in great detail.

First Step: *Understand the Theorem Proven.* The first and most important thing to do when assessing a mathematical proof is to understand the theorem that is being proven. You might not understand the particular claims or some of the notation, but it is essential that you understand what *kind* of assertion the theorem is—i.e., its logical structure. So ignore the proof for a minute, and focus just on the statement of the theorem. If we were to translate this sentence into *SL*, it would be a bi-conditional. There really isn't any other truth-functional structure there, just two atomic sentences. So we have:

$$A \equiv M.$$

But a bi-conditional is really just the conjunction of two conditionals: an *only if* and an *if.* So we have:

$$A \supset M$$
$$M \supset A.$$

You'll notice the proof is two distinct paragraphs, and here we have two distinct material conditional claims:

1. IF an (A)rgument with a finite number of premises is TF-Valid, THEN its corresponding (M)aterial conditional is TF-True;
2. IF a finite-premised argument's corresponding (M)aterial conditional is TF-True, THEN the (A)rgument is TF-Valid.

So we get our first hints as to the likely structure of the proof on the basis of the logical form of the theorem. Happily in this case we are right, since each paragraph proves one of these conditionals, and in this order.

But seeing this takes a bit of work. The next thing we might do is think about *SD*. How could one derive a bi-conditional using the rules of *SD*? The simplest way is of course ≡I, which requires two sub-derivations, the first of which assumes the left-hand side and then derives the right under that assumption, and the second of which assumes the right-hand side and then derives the left under this other assumption. In other words, deriving the two material conditional sentences and then putting them together. Which, again, seems to be exactly what this proof is doing. However, we're still kind of guessing at this point, because we haven't actually examined at the proof in detail. So this brings us to...

Second Step: *Quasi-Paraphrase the Proof.* I've already got us started on examining the proof for logical structure. Read through it carefully, not in order to understand every claim made, but in order to identify the premise/conclusion indicators, and any familiar logical expressions. You can underline important expressions at this point, just as when performing the first step of symbolizing a typical English argument in *SL* (and you thought that skill would be useless!).

It's here that we get the first suggestions that our guesses as to the structure of the proof are correct—that the proof is divided into two paragraphs is telling, as noted above. More importantly, you'll note that each paragraph starts with a premise indicator, but very particular ones: Suppose that... and Assume that..., respectively. And each paragraph ends with a conclusion indicator: Therefore... and Hence..., respectively. Forgetting for a moment all the stuff in-between, what we've identified here is the basic logical structure of the proof, which matches up exactly with the form of the theorem. This brings us to the third step.

Third Step: *Identify the Major Structures.* It is useful to think of doing a derivation in *SD* again, where we know that *goal analysis* is our best friend. Doing this, we can set up the expected structure of the derivation long before filling in the details. When constructing or trying to understand a proof, we can do exactly the same thing:

> Let $\{\mathbf{P_1}, \ldots, \mathbf{P_n}\} \vdash \mathbf{Q}$ be an arbitrary argument with a finite number of premises, call this argument **Arg**.
>
> Suppose that A is true of **Arg**.
>
> [Proof of M goes here.]
>
> Therefore $A \supset M$ is true.
>
> Assume that M is true of **Arg**.
>
> [Proof of A goes here.]
>
> Hence $M \supset A$ is true.
>
> The above shows that for any argument in *SL* with a finite number of premises, $A \equiv M$ is true.

This captures the basic structure of the proof, and hopefully you can see it mirror how a derivation in *SD* would be set up to derive the same kind of sentence. In this case, however, we've used the premise and conclusion indicators, as well as an understanding of the theorem, to figure out the structure that the proof has. There are still two things missing, of course. In the first place, we need to fill in the details of each "sub-derivation". In the second place, we've added a line to the top of the structure at the same scope level as the conclusion. Why?

Most importantly, we need to recognize that the *argument* the proof is discussing throughout (and it's corresponding material conditional) is an *arbitrary* one. The proof isn't talking about some *specific* argument, because we want to prove the theorem *in general*, for any argument of *SL* with a finite number of premises. Hence the use of meta-variables, and the fact that our argument has n premises—any number, but at least one (in order to conform to the text's definition of an argument always having at least one premise). I've tried to capture this idea by adding the top line to our structure: If we can prove that $A \equiv M$ for an arbitrarily selected argument **Arg** (which isn't really an argument anyway, just a *schema*), then we will have shown that it holds for *every* argument of *SL*, just like we want. Many proofs will make this assumption clearer in the proof itself.

Fourth Step: *Work Through the Reasoning.* Now we just need to fill-in our understanding of the details of each "sub-derivation". In many mathematical proofs, these sorts of details will be some kind of algebraic manipulation or calculations. Just as often, as is the case here, we will have a clever application of *previously established results* and *definitions*, which are usually proven or defined earlier in the work. These will be strung together using simple inferences that again mirror the structure of our derivation rules. Of course, some proofs will be more complex than others, and so involve more "filling-in" or make inferences that take some work to grasp.

Notice I haven't said anything about the colours I've assigned to the various expressions in the proof. I've used the colours to indicate the various "definitions" invoked, including facts we know about how the truth-functional connectives work. We'll begin with the first paragraph, and I'll try to be as explicit as possible.

> The proof starts by *assuming* **Arg** is a TF-Valid argument of *SL* (this is *A*). The next sentence invokes the definition of TF-Validity, establishing that the conclusion of **Arg** must be true whenever the premises are.
>
> The proof then reminds us of the fact that a conjunction is true only when all of its conjuncts are; and of the fact that a material conditional cannot be true when its antecedent is true and its consequent is false.
>
> In this same sentence, the proof says that *it follows that* the corresponding material conditional of **Arg** cannot have a true antecedent and a false consequent (i.e., there is no truth-value assignment where this is the case for our argument **Arg**). Why does this follow?
>
> Remember that the proof *assumes* that **Arg** is valid So either **Q** is true, in which case the premises may be true or false, but the argument still valid; or **Q** is false, in which case one of the premises must also be false to keep the argument valid. The proof is trying to get from here to the claim that **Arg**'s corresponding material conditional is a TF-Truth. Recall that we *construct* a corresponding material conditional using the premises and conclusion of an argument.
>
> So these same options (**Q** true or **Q** false) and their consequences apply to the corresponding material conditional too. Either **Q** is true, in which case the premises could have been true or false in our original argument, and so the corresponding material conditional is true; or **Q** is false, in which case one of the premises must have been false in the original argument (in order to keep the argument valid), in which case *the corresponding material conditional is still true* (since F F $\rightarrow$ T for a material conditional).
>
> So whether **Q** is true or false, the corresponding material conditional of **Arg** is always true (on the *assumption* that **Arg** is valid). But this is just the definition of a truth-functional truth. Recall that this was *M*,

> what the proof was attempting to conclude given A! So the proof has established the first conditional.

That was long... but I'm trying to make it super-clear what the reasoning is, and the application of each definition. Hopefully you can see how the proof really just *transforms* the definition of a TF-Valid argument, along with the assumption that **Arg** is indeed valid, into the fact that its corresponding material conditional is TF-True. Here's a diagram:

Assume **Arg** Valid → Def Validity → Def Conj → Def MatCon →
Arg CMC → **Arg** CMC Always True → **Arg** CMC TF-True

The second half is the reverse of the first half. This shouldn't be surprising, since we have another conditional proof, but in the opposite direction. Faster this time:

> The proof assumes that the CMC of **Arg** is TF-True (This is M).
>
> If this is the case, by the truth-table of a material conditional, the antecedent of **Arg** cannot be true while the consequent (i.e., **Q**) is false.
>
> Recall a conjunction is true iff all its conjuncts are. So by the definition of a TF-Valid argument, whenever **Q** is false, one of the conjuncts must be (in order to make the antecedent false, and so the CMC true, which we assumed). But when **Q** is true, all of the $\mathbf{P_n}$ are. So putting these two together, the argument is TF-Valid. This is A, so the proof establishes the second conditional, and it's done!

What we've done is broken the reasoning down into its gruesome details. All of this content is there already in the original proof, but it leaves out any repetition, mention of what definition is what and how they apply, etc. As you get better at reading proofs, you will begin to fill this kind of thing in automatically. All that is in the proof are the direct indications of the structure, and the exact parts of the definitions and inferences needed to get from each step to the next. But this is all we need, and shows exactly *why* the theorem is true. *We* use the proof to rehearse the reasoning in our own minds and convince ourselves of the truth of the theorem.

At the other end of the spectrum, hopefully you see now how we could symbolize this argument in *SL* (well, really, we'd have to use *PLE*) and do a step-by-step derivation. And you can see how long the derivation would be! Note that each of the definitions we apply would have to be an initial premise to the argument.

2.2 Some Terminology

Before moving on to analyze a few more proofs and then make some generalizations about this process, let's slow down for a minute and get a grip on some of the terminology usually employed by the people who spend their time proving theorems (rather than just analyzing them!). Mathematicians use the word '*Theorem*' for any significant result that has been proven via the kind of arguments we're analyzing. However, almost every theorem relies upon some set of definitions, previous theorems, or axioms. Thus, our use of the word 'theorem' in *SD* is slightly different than the typical mathematical use, since a *Theorem in SD* is derived without any initial assumptions. You should note that the only reason this is possible in *SD* is because a whole lot of logical content is "built-in" to the derivation rules!

Mathematicians use the word '*Conjecture*' for an assertion that they think is a theorem, but that hasn't yet been proven. Two famous examples of conjectures that were recently proven are the *Four Colour Theorem* (about how few colours we need to colour in an arbitrary map without any two colours touching) and *Fermat's Last Theorem* (that no three positive integers a, b, c can satisfy the equation $a^n + b^n = c^n$ when $n > 2$). These theorems each have a very storied history. They are deceptively simple to state, but took the world's greatest mathematical minds hundreds of years to find a proof. Both have now moved from conjectures to theorems. Mathematicians use the word '*Corollary*' for an easier result that follows quickly from a more complex theorem. And they often use the word '*Lemma*' for an easy theorem that is used in the proof of the present theorem of interest.

Usually in mathematics, logic, economics, or physics textbooks, theorems, definitions, corollaries, etc. are numbered. Their proofs are often indicated by the word '*Proof*', as we have done, and the proof is usually given right after the statement of the theorem. Finally, you may have noticed the little block at the end of our proof of Theorem 1, the '■'. This is often called the *Tombstone*, and indicates the successful completion of the proof. Mathematicians sometimes append the initialism 'Q.E.D.' to the end of a proof instead. This is short for the Latin '*quod erat demonstrandum*', meaning "*that which had to be demonstrated*". Being able to finally—at the end of a long bout of work—draw the tombstone and say "Q.E.D." is a great feeling. It means the job is done and that you are, indeed, awesome.

2.3 Reading Mathematical Proofs

Let us get back to business. It will be instructive to go through a few more examples before we start constructing our own proofs. We will proceed more quickly in our analysis than above; however, we will afterward be able to make some generalizations that will help with the analysis and construction of future proofs.

Let's consider a proof from our favourite mathematical theory, Set Theory!

Theorem 2 (Simple Set Theory Theorem)**.** $(\Gamma \cap \Delta) \subseteq \Gamma$.

Proof. Let x be an arbitrary object and suppose that $x \in (\Gamma \cap \Delta)$. Then by definition $(x \in \Gamma)$ & $(x \in \Delta)$, from which it follows that $x \in \Gamma$. Hence if $x \in (\Gamma \cap \Delta)$, $x \in \Gamma$. Since x is arbitrary, it holds for all x. So by definition $(\Gamma \cap \Delta) \subseteq \Gamma$. ■

Again, that probably went by a little quickly. But now we can utilize the steps outlined above in order to perform an analysis and gain a better understanding of the theorem and how it has been proven.

The first step is understanding the claim being made. The theorem states that for any two sets Γ and Δ, the set we construct by taking their intersection is included in the original set Γ. Rather than giving this "new set" (the intersection) a name, the proof just calls it '$\Gamma \cap \Delta$', which is typical in set theory. So the key to understanding this theorem is going to be reminding ourselves what the definition of *Set Intersection* is, as well as *Set Inclusion*. I repeat these definitions from the previous chapter:

Definition 11 (Set Inclusion)**.** *Let Γ and Δ be sets, then:*

$$\Gamma \subseteq \Delta \quad \text{iff} \quad (x \in \Gamma) \supset (x \in \Delta)$$

That is to say, one set is included in another when membership in the first implies membership in the second.

Definition 12 (Set Intersection)**.** *Let Γ and Δ be sets, then:*

$$\Gamma \cap \Delta = \{x : (x \in \Gamma) \ \& \ (x \in \Delta)\}$$

That is to say, the intersection of two sets is itself a set with the characteristic property in which each object in the intersection is a member of both original sets.

I have not here performed Step Two, but I encourage you to do that now: What are the premise and conclusion indicators in the passage? What are the relevant truth-functional connectives?[1] The proof is also very short, so we will combine steps three and four below. It will be useful for you to try to sketch out the structure of the proof using the indicators first, before moving to my analysis below.

Let x be an arbitrary object.

  Suppose $x \in (\Gamma \cap \Delta)$ is true.

    By Definition 12, it follows that $(x \in \Gamma) \ \& \ (x \in \Delta)$.

    As a result, $x \in \Gamma$.

  Thus, IF $x \in (\Gamma \cap \Delta)$, THEN $x \in \Gamma$.

Thus, FOR ALL x, if $x \in (\Gamma \cap \Delta)$, then $x \in \Gamma$.

Therefore we have $(x \in (\Gamma \cap \Delta)) \supset (x \in \Gamma)$.

So, this time by Definition 11, $(\Gamma \cap \Delta) \subseteq \Gamma$.

[1] Partial Answer: An important premise indicator in this proof is *By definition...*.

The key steps of this proof are now much clearer, but perhaps still a bit obscure. Just as with the proof of Theorem 1, this proof begins by assuming that something is true for the purposes of an inference like $\supset$I. In this case, what's "filled-in" is the transformation of the original assumption using a definition. You'll note that according to our definition of Set Intersection, the conjunction *is just the same as* the original expression, but in a different form. So we can replace one with the other. By an inference very much like &E, the proof ends up with the fact that $x \in \Delta$. This allows the establishment of the conditional sentence, and so the proof can "leave the sub-derivation". In the next step, we make clear the fruit of the first line, that the object x was arbitrarily chosen to begin with. Just as with the proof of Theorem 1, this means that we can conclude the same result for any object x that happens to meet the requisite conditions (i.e., being a member of the intersection of Γ and Δ). The next line transforms this informal general statement into an expression using the horseshoe. The reason for this is because our definition of Set Inclusion uses this expression. So following that definition from right to left, the proof transforms the horseshoe expression with mention of the members of each set into a subset expression without need for mention of the members of the set. And that's it!

This is about the level of difficulty of proofs that the textbook will ask you to construct in Exercise 3.6E (it's not really much less difficult than Theorem 1, just written a bit more explicitly from the get-go). It's mostly just figuring out what the Theorem says, and then manipulating definitions in such a way to construct the statement. Exercises in chapters 6 and 11 are somewhat more difficult. And of course, the proofs that you'll be expected to understand will be quite a bit more involved—proving interesting facts about formal systems often takes *work!*

But sometimes really important proofs can be relatively simple, too. Just incredibly ingenious. Let's take a look at one such proof, one with real *substance*:

Theorem 3 (Infinity of Primes)**.** *There are infinitely many prime numbers.*[2]

Proof. <u>Suppose</u> there are finitely many primes. Let $p_1, p_2, \ldots, p_n$ be a list of all prime numbers. Let $m = p_1 p_2 \ldots p_n + 1$.

Now, m is either prime or a product of primes.[3] <u>Suppose first</u> that m is prime, then m is a prime not on the original list. This <u>contradicts our initial assumption</u>. <u>Suppose instead</u> that m is a product of primes, then by hypothesis m is divisible by some prime q. But m is not divisible by any of $p_1, p_2, \ldots, p_n$, because any such division results in a remainder of 1. So q must be a prime not on the list. This again <u>contradicts our initial assumption</u>, and thus there are infinitely many primes. ■

The aforementioned Greek mathemagician Euclid (ca. 300 BCE) offers a proof very similar to this in his book *Elements* (Book IX, Prop. 20)—a book taken to be

[2]Recall that a *Prime Number* is an integer divisible by only 1 and itself.

[3]This follows from another theorem that would either need to be assumed or proved prior this one, the *Fundamental Theorem of Arithmetic*: Every integer greater than 1 is either a prime number or a unique product of prime numbers.

the very model of rigorous deductive argumentation until the development of modern logical methods in the 20th century. It's certainly a substantial claim: that there are an *infinite amount* of prime numbers! But we can now easily come to understand the proof using the same tools of analysis as before.

So the statement is relatively simple to understand, although it may help to realize that something is infinite just in case it *is not finite*. Again, I will leave most of the quasi-paraphrasing as an exercise to the reader. Step three:

> Suppose that *there are only finitely many primes* is true.
>
> > So let $p_1, p_2, \ldots, p_n$ be a list of all the prime numbers.
> >
> > Let $m = p_1 p_2 \ldots p_n + 1$ (i.e., $p_1 \times p_2 \times \cdots \times p_n + 1$).
> >
> > It must be that EITHER m is prime OR m is a product of primes.
> >
> > > CASE 1: Suppose m is prime.
> > >
> > > > Take $q = m$.
> > > >
> > > > Contradicts the assumption that the list contained all primes.
> > >
> > > CASE 2: Suppose m is a product of primes.
> > >
> > > > [Proof q is a prime not on the list.]
> > > >
> > > > Contradicts the assumption that the list contained all primes.
> >
> > BOTH cases lead to a contradiction with the assumption.
>
> Thus it isn't the case that *there are only finitely many primes* is true.

There are a couple of interesting points here. First is that the overall structure of the proof is by *Contradiction* (think again of *SD*, particularly ∼I). The goal is to show that there are infinitely many prime numbers, so the proof supposes that there are only finitely many, which means that we can give a finite list of them all. What this amounts to is saying that there *does not exist* a prime number that is not on the list.[4] The goal now is just to find a prime number that cannot be on the list as given, since this would contradict the assumption that such a number does not exist.

This brings us to the second point, that the rest of the proof is by *Cases*. The proof first introduces a new number m, which it cleverly chooses because it will not be evenly divisible by any of the prime numbers on the original list. We can see this with an example. Suppose the list is $\{2, 3, 5, 7\}$, then $m = (2 \times 3 \times 5 \times 7) + 1 = 211$. While 210 is divisible by every number on the list (because it's just the product of those numbers), dividing 211 by the numbers will in each case leave a remainder of 1. Notice that the proof doesn't actually tell us which number m is, it's value is entirely dependent upon this supposed list of all the prime numbers.

[4] The proof calls this number 'q' in the second case. I have added a couple of steps involving q in the analysis of the first case to make the contradiction clearer.

Now it is a fact of arithmetic that every number is either a prime number, or what is called a "composite number", which means it is a unique product of prime numbers (e.g., $210 = 2 \times 3 \times 5 \times 7$, $42 = 2 \times 3 \times 7$, etc.). So the chosen number m, whatever it is, must either be a prime number, or a product of prime numbers. These are the only two options, and so the proof lays out the consequences of each, one after the other. Notice that I underlined the indicators in the proof that tell us it is proceeding by cases. It is a consequence of each case that there is a number q which must be a prime, but isn't on the original list. Since these are the only options, and in each case we derive a consequence that contradicts the assumption that such a number doesn't exist, the original supposition must be false.

Let's look at one more by contradiction. This time that doesn't involve cases:

Theorem 4 (Simple Algebra Theorem)**.** *If $x^2 + y = 13$ and $y \neq 4$ then $x \neq 3$.*

Proof. Suppose $x^2 + y = 13$ and $y \neq 4$. Suppose $x = 3$. Substituting this into the equation $x^2 + y = 13$, we get $9 + y = 13$, so $y = 4$. But this contradicts the fact that $y \neq 4$. Therefore $x \neq 3$. Thus, if $x^2 + y = 13$ and $y \neq 4$, then $x \neq 3$ ■

For anyone who doesn't like "mathy" things, at first glance this may actually be the worst one yet. But, again, we know that taking it one step at a time, it's easy to see what's going on in the proof. As always, we must first understand what the theorem is stating. It is a conditional sentence of the form $(\mathbf{P} \;\&\; \mathbf{Q}) \supset \mathbf{R}$. As expected, the proof proceeds by assuming both antecedent conditions, and then deriving the consequent. In this example however, the consequent is a negated sentence, i.e., $\sim (x = 3)$. So what it does is make the further assumption that x does indeed equal three, in the hopes of getting a contradiction and being able to conclude the opposite. It then simply does the algebra necessary, and ends up with $y = 4$. However, one of the antecedent conditions of the theorem was that $y \neq 4$, and so this can work as the contradiction required. Remembering how ∼I works, we see that since the assumption that $x = 3$ led to a contradiction, it must be the case that $x \neq 3$. Since this was the consequent of the original theorem, the proof obtains the whole conditional sentence. Here's the basic structure:

> Suppose $x^2 + y = 13$ and $y \neq 4$.
>
> > Suppose $x = 3$.
> >
> > > [Necessary algebra.]
> > >
> > > Given the assumption, $y = 4$, but from above, $y \neq 4$.
> >
> > Therefore $x \neq 3$.
>
> Thus, if $x^2 + y = 13$ and $y \neq 4$, then $x \neq 3$.

I think that's probably enough analysis. Let us move on to some generalizations that will help us construct our own proofs.

2.4 General Proof Structures

Hopefully you have started to see some general patterns emerging here. In the first place, much of the terminology in all of these proofs is similar, even though they are proving very different things about/with very different mathematical theories (we have meta-logic, set theory, number theory, and an algebraic problem). The terminology of 'suppose', 'let x be arbitrary', 'by definition', 'it follows that', etc. are conventions that mathematicians have settled upon because they clearly indicate the structure of the proof to the trained reader. Second, we should recognize a lot of these structures, and the individual inferences made within them, as very similar to the derivation rules familiar to us from *SD*.

With this in mind, we can list some common proof structures:

1. **Conditional Proof.** We saw this kind of structure in the proofs of Theorems 1 and 4. The reasoning mirrors our *SD* rule $\supset$I. We wish to prove a conditional statement, so we assume that the antecedent is true, and then try to obtain the consequent given that assumption. If we can do this, then we can conclude the entire conditional statement is true. The structure looks like this:

   Suppose **P** is true.

       [Proof of **Q** goes here.]

   Therefore, if **P** then **Q** is true.

   Obviously in the proof of Theorem 1 this strategy needed to be used twice because the theorem was a bi-conditional.

2. **Indirect Proof.** This is also called *proof by contradiction*. We saw it above in the proofs of Theorems 3 and 4. The reasoning mirrors our *SD* rules $\sim$I and $\sim$E. If we're trying to prove a statement, assume its negation is true and then derive a contradiction given that assumption. If we can do this, then we can conclude the original statement is true. The structure looks like this:

   Suppose **P** is false.

       [Proof of contradiction goes here.]

   Therefore, **P** is true.

   Notice that Theorem 4 was actually trying to prove a negation, that $x \neq 3$, and so it assumed that $x = 3$ was true and got a contradiction from that.

3. **Proof by Cases.** The proof of Theorem 3 also relied upon this structure. The reasoning actually mirrors our *SD* rule $\vee$E. If you have a situation in which there are several mutually exclusive and exhaustive possibilities, but you want to prove that some goal follows no matter what, you can break the situation down into its individual possibilities, and show that your goal follows separately

from the assumption of each possibility. If you can do this for all the possibilities, you can conclude that your goal is true. The structure looks like this:

> Case 1: Suppose **P** is true.
>
> [Proof of goal goes here.]
>
> Case 2: Suppose **Q** is true.
>
> [Proof of goal goes here.]
>
> Since we know $\mathbf{P} \vee \mathbf{Q}$, these cases cover all the possibilities. Therefore the goal must be true.

Often you will have many more than two cases. Notice here I said 'goal' rather than 'statement' or something like that. This is because it depends upon the particular proof we're doing. For example, in Theorem 3 the goal of the proof by cases was to establish a contradiction from the fact that the number m exists.

4. **For All x Proof.** Theorems 1 and 2 rely upon this structure. We do not have an *SD* equivalent, but we will have a rule just like this in *PD*. The idea is that we're trying to prove that some property or statement holds for all things of a certain type. If we choose one at random, then that gives us something to manipulate in the proof without sacrificing generality. Once we've proven what we want about that object, we conclude that it holds of all things of the same type. Importantly, *our proof cannot depend upon any special properties that the object selected might have*. The structure looks like this:

> Let x be arbitrary.
>
> [Proof that some property or statement holds of x.]
>
> Since x was arbitrary, we can conclude that the property or statement holds in general (i.e., for all x of the given type).

We've been talking here about an "object" x. This must be interpreted broadly. As we saw in Theorem 1, our "object" there was an argument of *SL*.

5. **Existence Proof.** Theorem 3 relied upon showing that some number existed, contradicting the assumption that it did not. The proof did this in each of the cases. Again, we don't have a rule in *SD* that corresponds to this kind of reasoning, but we will in *PD*. The idea is to introduce a new object with whatever value we need it to have, and then prove that it has the properties required by the theorem. The structure looks like this:

> Let $x =$ (the value you decide upon).
>
> [Proof that x has the properties required goes here.]
>
> Therefore, there exists an object with the properties required.

Theorem 3 was actually a particularly tricky case using this structure, because the object introduced at the beginning of the structure, m, was only the object shown to exist (i.e., q) in the first case. In the second case, the proof used the assumption of the existence of m to prove that another object, q, existed.

These five basic proof structures can be iterated, combined, nested, etc., in whatever way is required to prove the theorem. We have seen different combinations in each of the theorems analyzed above. It is also important to note that these are not the only possible proof structures, there are many more. One other very useful one is this:

6. **Prove the Contrapositive.** If you have to prove a conditional sentence, but Conditional Proof is not working, take the contrapositive of the original sentence (i.e., $\sim \mathbf{Q} \supset \sim \mathbf{P}$) and try to prove that by conditional proof. If you can, you can instead conclude the original sentence.[5]

This works because, as we know from our *SD* rule of replacement, the contrapositive of a conditional sentence is equivalent to the original sentence.

2.5 Constructing Your Own Proofs

The tools of proof analysis that we have thus-far developed work exactly the same way when you go to prove a theorem yourself. The first and most important step is to understand what the theorem is actually stating. Then you start sketching a proof structure by using goal analysis, in the same way you would construct a derivation. Rather than looking to the derivation rules however, you instead look to the proof structures outlined above. And just like doing a derivation, you work from the outsides (the bottom and top) toward the middle, so you're just left with easy details to fill-in in the middle.

As mentioned, there are more possible proof structures than have been outlined here, and not everything boils down to "applying" the right one. As we've seen, in proving a theorem we will often need to make creative use of definitions, transform sentences into a new form, make the right assumptions, know the right lemmas and previously established theorems, or perform some kind of calculation or manipulation. To appreciate this fact—that real mathematics requires *extreme creativity* and *intuition*—is to have a much deeper understanding of mathematics than most.

But what about more practical concerns... the theorems we will have to understand and prove in the text? As I've said, most will not be too difficult, and so the tools provided here will get you well on your way. There are also two general strategies that we should codify for when we attempt to prove a theorem ourselves:

[5]Proving the *Completeness of SD* and the *Completeness of PD* in chapters 6 and 11 of the text use the strategy of proving the contrapositive. Specifically, proving theorems 6.4.2 and 11.4.2 are done by proving the contrapositive. Once proven, the original formulations of these theorems are then used to prove the Completeness theorems (6.4.1 and 11.4.1 respectively).

1. **Understand What the Theorem States.** I've said this a bunch of times, and it's step one of analyzing a given proof. But I cannot repeat it enough. You won't be able to make an honest go at proving a theorem if you don't understand what it's actually saying. This involves recognizing not only the logical structure of the sentence, but also every single one of the terms and definitions that it relies upon. That leads us to strategy two...

2. **Understand and Use Definitions.** It is vital to understand the definitions of the concepts that make up the theorem you're trying to prove. Recall the proof of Theorem 1. The structure was extremely basic, and the 'work' was done by lining up the definitions in the right way to 'get' from one side of the bi-conditional to the other and back again. This was even more explicit in the proof of Theorem 2, which relied upon the fact that we defined the set-theoretic notions in terms of the truth-functional connectives.

Most often in the meta-logical theorems and proofs that you will encounter, definitional manipulation is the only 'work' that needs to be done once the basic structure is in place. This, again, often takes some ingenuity. But with practice you'll get the hang of it. With that in mind I recommend attempting Exercises 3.6E from the text before moving on to chapters 6 and 11.

2.6 Conclusion

You should now have a much easier time getting through chapters 6 and 11 of the text, and you'll have a much better understanding of the discussions therein. Perhaps most importantly, you now have a very good understanding of just what mathematicians (and logicians!) *do* much of the time. There is always another interesting conjecture about a mathematical theory or a logical system that somebody will think up, and the question then becomes one of trying to *find a proof*.

If you're interested in understanding mathematical and logical proofs even better, I highly recommend both the Hunter and the Velleman books listed in the bibliography at the end of these notes. The former deals specifically with the meta-logical concepts and proofs examined in the text, but in even greater detail. It is an excellent and very robust treatment of the science of our logical systems. The second book is like this chapter of the present notes but on steroids. In fact, this discussion has been modeled upon the Velleman, and I consulted that book throughout the preparation of this chapter and the next.

Chapter 3

Mathematical Induction

3.1 Induction—The First Steps

Recall that near the beginning of the text it makes a distinction between two different kinds of argument: (i) Deductive; and (ii) Inductive. What is characteristic about the first kind is that they are *truth-preserving*—a deductive argument will always result in a true conclusion as long as we begin with only true assumptions. An inductive argument does not have this character. A good inductive argument will *likely* take one from true premises to a true conclusion, but this is not a guarantee.

Mathematical Induction has *nothing to do* with this sort of inductive argument. It is an unfortunate accident of history that this structure of mathematical proof has been given the name 'induction'. Mathematical induction is a deductive inference, and so is truth-preserving just like every other logically valid form of argument. However, while we will here learn how to use mathematical induction, we will not attempt to *justify* its validity in the same sense that we can with the rules of *SD* and *SD*+ (using a truth-table). Although such a justification is entirely possible, it requires some very heavy-duty set theory that is beyond the scope of these notes. Instead, we will merely offer an intuitive justification to help you see for yourself that the form of argument is always valid when used correctly. From this point forward, any time these notes mention 'induction', it is mathematical induction that is meant.

We will begin our discussion of mathematical induction in 1780, with mathemagician extraordinaire Carl Friedrich Gauss, when he was a child of only seven or eight years old. As the story goes, he was misbehaving in school one day, so his teacher took him aside and as punishment made him sit down and add up all the natural numbers from 1 to 100. The teacher thought this would take some time, but within *a few seconds*, Gauss had the answer: *The sum is 5,050!* How did he do it?

The young Gauss realized that rather than doing each sum individually with a running total, we find that the addition of the two terms at opposite places on the list always yields 101. That is, $1 + 100 = 101, 2 + 99 = 101, 3 + 98 = 101$, etc. Since there are 50 such additions that need to be done for the whole list, we really only need to perform one multiplication: $50 \times 101 = 5050$! And so that's what Gauss did.

The hard part of course is recognizing that this pattern holds. Mathematicians are very good at recognizing the general structure of a problem rather than just trying to brute force a solution. And in fact, we can use Gauss' pattern to calculate the sum of all the first n numbers, up to any number we wish. The general form of this result is as follows:

$$1 + 2 + 3 + \cdots + n = \frac{1}{2}n(n+1)$$

So if we want to add up the first 6 natural numbers (i.e., $n = 6$), we simply multiply $0.5 \times 6 \times 7 = 21$. And similarly for any n. The question we will deal with in this chaper: How can we *prove* that this formula gives us the correct answer in each case? Certainly we cannot test each case individually, since there are an infinite number of them. Instead, we will offer a *Proof by Mathematical Induction*.

3.1.1 As a Formal Rule

The form of induction discussed in the text is *Strong Induction*. As the textbook notes, it is as capable as so-called *Weak Induction* (not surprising given the names!). For those with an analytic philosophy background, note that both of these are equivalent to the *Well-Ordering Principle*. If you don't know what that is, don't worry about it because it's not important for what follows. The problem with induction, from the point of view of expressing it as a sentence in the logical languages of the text (i.e., *SL*, *PL*, and *PLE*) is that we can't. Strictly speaking, none of our logical languages are expressive enough to completely capture the content of mathematical induction. However, this is okay because we can express induction as a rule of derivation in *PD*. The rule looks like this:

(MI)

	$\mathbf{P}(0)$
	$\mathbf{P}(\mathbf{k} \leq \mathbf{n}) \supset \mathbf{P}(\mathbf{n+1})$
$\triangleright$	$(\forall \mathbf{n})\mathbf{P}(\mathbf{n})$

This will be confusing if you haven't started learning *PL* yet. But the basic structure should be clear: We need two lines, one of which is a material conditional, and then we can conclude a third line. The symbols '**k**' and '**n**' are meta-variables that take as values the natural numbers. The sign '**P**' here does not stand just for sentences, instead, it stands for *a property of objects*. So '$A(2)$' might mean *2 is prime*, or *2 is even*, or *2 is my favourite number*, or whatever. Most importantly, the expression '$\forall$' just means "for all objects" of the specified type, so here it means "for all numbers n". We will not actually add this rule to *PD*, since we only use it when doing informal mathematical proofs. The point is that this is how we can express it in one of our logical languages.

3.1.2 As an Argument, with Example

Here is the same thing, but expressed in words as an informal argument form:

1. The statement is true when n is the minimal case.
(usually 0, sometimes 1 or more)

2. IF the statement is true for all $k \leq n$, THEN the statement is also true for $n + 1$. (i.e., when $n + 1$ is substituted for n.)

$\therefore$ The statement is always true (i.e., the statement holds for all n).

Again, there are two premises which lead to a general conclusion. The first premise is called the *Basis Clause*—we show that the property **P** holds when n is the smallest possible, usually zero. The second premise is called the *Inductive Step*—notice that this sentence is a material conditional. The antecedent of this conditional is called the *Inductive Hypothesis*. We show that IF the property **P** holds for all natural numbers k up to and including some arbitrarily selected number n (i.e., if the inductive hypothesis is true), THEN it also holds for the next number, which we call $n + 1$ (because n was chosen arbitrarily in the antecedent). If we can show both of these premises, then we can conclude that the property **P** holds generally, for all numbers n. I think it would be helpful to dive into an example, so let's go back to the young Gauss and his formula for adding the first n natural numbers:

Theorem 5 (Gauss' Equation)**.** *For all n, $1 + 2 + 3 + \cdots + n = \frac{1}{2}n(n + 1)$.*

Proof. What we are trying to prove is that this equation holds generally, for all sequential series of numbers from 1 to any number n. In other words, that rather than adding up each number in the series and keeping a running total, we can just plug the last number in the series into the right-hand side of the equation (to the right of the equals sign), and we'll get the right answer. This is what Gauss did for the numbers from 1 to 100. He made $n = 100$, and so got $0.5 \times 100 \times 101 = 5050$. Since we can't check every number, we need a general method, that's induction.

The first step is to prove the *Basis Clause*. As will often be the case we take $n = 0$. In other words, we want to prove that the basis clause is true with the minimal possible number. This is easy, we just substitute 0 into the equation:

$$0 = \frac{1}{2} \cdot 0 \cdot (0 + 1) = 0$$

Since $0 \times \frac{1}{2} = 0$, and $0 \times 1 = 0$, the above substitution is true, and we have thus proven the basis clause.

Next we must prove the *Inductive Step*. It's important to recall how to approach a proof, and realize what we need to prove here. We just need to prove a conditional: that *if* the equation holds for all numbers up to n, *then* it holds for the next number too, that is, for $n+1$. So what we're *not* going to do is to show it holds for every number up to n—this would be to prove the inductive hypothesis, which we don't have to do. Instead, we're going to rely upon our old friend *Conditional Proof*, or $\supset$I if you'd like. We'll *assume* the inductive hypothesis (i.e., the antecedent of the conditional), and then show that under that assumption, the equation holds for $n+1$. So we suppose the equation is true for n:

$$1+2+3+\ldots+n = \frac{1}{2}n(n+1) = \frac{n(n+1)}{2}.$$

Notice I've re-written the assumption in a different form on the far right. This is just for pedagogical purposes, so it'll be easier to identify in the algebra below. Now, we want to show that given this assumption, the equation is true for $n+1$. I.e., that:

$$1+2+3+\ldots+n+(n+1) = \frac{n(n+1)}{2} + (n+1)$$

Here we've added the next number to the series, i.e., $n+1$ is added to the left (in blue). We want to see if it can be added also to the right-hand side, and that we can then obtain an expression of the same form as the original. We can indeed obtain it by some (perhaps forgotten) high school algebra:

$$\begin{aligned}1+2+3+\ldots+n+(n+1) &= \frac{n(n+1)}{2} + (n+1)\\ &= \frac{n(n+1)+2(n+1)}{2}\\ &= \frac{n^2+3n+2}{2}\\ &= \frac{1}{2}(n+1)((n+1)+1)\end{aligned}$$

This equation that we've ended up with is the same as the right-hand term our original equation (in green), but we've replaced n with $n+1$. Since we've shown that the equation holds when $n=0$ (the basis clause), and we've proven that when the equation holds for n, it holds for $n+1$, we've shown that it holds for all numbers. ■

And that's the end of the proof. We've shown that Gauss' method works in general, for all series from 1 to any number. What we've done in these last few steps might still be a little confusing. Let's go back to an actual instance in order to see what happened. When $n=6$ we have:

$$1+2+3+4+5+6 = \frac{6(6+1)}{2}$$

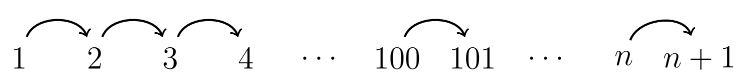
Now remember, this is our inductive hypothesis. Of course, with an actual number plugged in, it's obviously true (since each side of the equation = 21). When n is arbitrary, this isn't obvious. But it is assumed to be true for the purposes of conditional proof. What we want to show in the inductive step is that *given* this is true, the next case will also be true, i.e., that the case for $n+1$ will be true. So what we have to do is add $n+1$ to each side of the equation, and see if it still holds:

$$1+2+3+4+5+6+(6+1)=\frac{6(6+1)}{2}+(6+1)$$

Again, this is easy to see with actual numbers instead of variables. What we did with the algebra is to show that we can obtain an expression of the same form as the original, with $n+1$ instead of n, simply by adding $n+1$ onto the case for n. This shows that if the equation holds for n, then it will hold for $n+1$ too. And that's it.

3.1.3 Informal Justification, with Another Example

The point of arguments by induction is usually to show that some property holds of all the elements of an ordered series. To do this we show that the property holds in the smallest case, and then we show that if it holds up to some arbitrary case, this implies that it holds for the next case too, and thus the property holds of all cases. The proof structure takes advantage of the fact that the natural numbers form an ordered list, such that if we start at the beginning and keep going from one to the next, we can eventually reach any given number, along with the fact that it's often very easy to create a *correspondence* (more on correspondences in the next chapter!) between the natural numbers and any given ordered series of interest.



The analogy that people often use when introducing induction is a row of dominoes standing upon their ends. If we know:

1. The first domino will fall.
2. For any given domino, if it falls, then its neighbour will also fall.

We can conclude that all the dominoes will fall, and that this fact is inevitable. The first step is the *Basis Clause*, the second is the *Inductive Step*.

Another useful analogy is perhaps climbing a ladder. The basis clause is climbing onto the first rung of the ladder. The induction step states that if we can climb up to a given rung, then we can take then next step and climb onto the next rung. If we can show both of these, it's clear that we can climb the whole ladder.[1] So proofs

[1] I will not suggest we throw it away after we have climbed up it.

by mathematical induction work by going through each “step”, from one to the next. But of course, with the natural numbers there are an infinite number of steps. What we do is use the power of logic to reason about an arbitrary step and its successor, and then generalize this to every step. Here’s another example.

Consider the following list of calculations, which suggests a surprising pattern:

$$2^0 = 1 = 2^1 - 1$$
$$2^0 + 2^1 = 1 + 2 = 3 = 2^2 - 1$$
$$2^0 + 2^1 + 2^2 = 1 + 2 + 4 = 7 = 2^3 - 1$$
$$2^0 + 2^1 + 2^2 + 2^3 = 1 + 2 + 4 + 8 = 15 = 2^4 - 1$$

Can you see the pattern? It seems to be that:

$$2^0 + 2^1 + \cdots + 2^n = 2^{n+1} - 1.$$

In other words, the series of all exponents of two up to some exponent n, added together, equals the next exponent of two, minus one. We’ve shown this above for the first four possible cases. But does the pattern continue to hold for every exponent of two—i.e., for all values of n? Let’s prove it with induction!

Theorem 6 (Exponent Pattern Theorem)**.** *For all* n*,* $2^0 + 2^1 + \cdots + 2^n = 2^{n+1} - 1$.

Proof. Again, the basis clause is easy. We already did it above when we started listing calculations which suggested the general pattern in the first place. We take $n = 0$ and plug it into our theorem, then we have:

$$2^0 = 1 = 2^1 - 1.$$

It is trivial to verify that each part of the equation is equal to 1, so the basis clause is true. Often it will be easy to prove the basis clause. The tricky part is usually figuring out what to do for the inductive step. In the inductive step, we have to let n be any arbitrary natural number. We then want to show (likely by conditional proof) that if we assume $2^0 + 2^1 + \cdots + 2^n = 2^{n+1} - 1$ is true, then we can prove that the next case is true, which is, $2^0 + 2^1 + \cdots + 2^{n+1} = 2^{n+2} - 1$

So we can think of $2^0 + 2^1 + \cdots + 2^{n+1} = 2^{n+2} - 1$ as our new goal. We need to show that this is true, on the assumption that the inductive hypothesis is true. The key to this proof (and the previous one, actually) is to recognize that the left-side of our goal here is exactly the same as the left-side of the inductive hypothesis, but with an extra term added on. In this case 2^{n+1} (in the last proof it was $n + 1$). So the strategy is to take the inductive hypothesis and simply add the extra term to each side of the equation, giving us:

$$(2^0 + 2^1 + \cdots + 2^n) + 2^{n+1} = (2^{n+1} - 1) + 2^{n+1}$$

But of course (remembering back to high school algebra again), $2^{n+1} + 2^{n+1} = 2^{n+2}$. This means that we have:

$$2^0 + 2^1 + \cdots + 2^{n+1} = 2^{n+2} - 1$$

And this just was our new goal, thanks to the inductive hypothesis. So we're done! We've proved that the pattern holds indefinitely. ■

Does this proof really show the general result? For all numbers n? Think about it like this. In the basis clause, we proved that $\mathbf{P(0)}$ is true. This should be uncontroversial. We also calculated (above the proof) that $\mathbf{P(1)}, \mathbf{P(2)}$, and $\mathbf{P(3)}$ holds, but this doesn't matter for the proof. In the inductive step, we proved $\mathbf{P(n)} \supset \mathbf{P(n+1)}$ is true. Now, since $\mathbf{P(0)}$ is true, we can substitute that into the conditional of the inductive step, and we get $\mathbf{P(0)} \supset \mathbf{P(1)}$. But this, along with the fact that $\mathbf{P(0)}$, allows us to apply $\supset$E and conclude $\mathbf{P(1)}$. Now that we have this, we can substitute this into the conditional of the inductive step again, and get $\mathbf{P(1)} \supset \mathbf{P(2)}$. Again using $\supset$E, we get that $\mathbf{P(2)}$ is true. But of course, we can substitute this into the conditional of the induction step, and get $\mathbf{P(2)} \supset \mathbf{P(3)}$, and then by $\supset$E, get $\mathbf{P(3)}$... So on and on like this, with repeated application of the inductive step, the dominoes fall, the ladder is climbed, and we see the result holds generally, for all n.

3.1.4 As a Mathematical Proof Structure

The way we have applied mathematical induction in the two examples above is not as a formal rule of inference. Instead, we have treated it like the informal mathematical proof structures of the previous set of notes. And indeed, this is usually how mathematicians work with it. Here's the structure:

> Basis Clause: [Proof of $\mathbf{P(0)}$ goes here.]
>
> Inductive Step: [Proof of $\mathbf{P(n)} \supset \mathbf{P(n+1)}$ goes here. Often it's:]
>
> > Let n be an arbitrary number.
> >
> > > Suppose $\mathbf{P(n)}$ is true.
> > >
> > > > [Proof of $\mathbf{P(n+1)}$ goes here.]
> > >
> > > Therefore, if $\mathbf{P(n)}$ then $\mathbf{P(n+1)}$ is true.
>
> Therefore, for all n, $\mathbf{P(n)}$ is true.

So speaking in terms of the previous chapter, often the proof of the inductive step contains a *Conditional Proof*, itself inside a *For All x Proof*. Usually the basis clause will be pretty straightforward, although it might involve a *Proof by Cases*, which means the inductive step will include that as well. Although sometimes finding the basis clause can be a bit tricky, as noted above the trickiest part of proofs by induction is usually figuring out what the inductive step should be.

3.2 Finding the Basis Clause & Inductive Step

The greatest part about mathematical induction is that we can also use it for things other than proving universal properties of natural numbers. The first example in chapter 6.1 of the textbook does this:

Example 1 (Parentheses)**.** *All sentences of SL have the same number of left and right parenthesis.*

Basis Clause: Sentences of *SL* with 0 connectives (i.e., atomic sentences) have the same number of left and right parentheses.

Inductive Step: If all sentences of *SL* with n connectives or less have the same number of left and right parentheses, then all sentences of *SL* with $n+1$ connectives have the same number of left and right parentheses.

Without worrying about the actual proof (the textbook gives a good run-through of the proof), what's interesting here is that we're using induction to prove a fact about *SL*. We're not proving anything *about* the natural numbers, but we're setting up a correspondence between the ordering of the natural numbers and some ordered aspect of our formal language for which we want to prove something. This is how induction is used to prove things that aren't directly about the natural numbers.

Approaching this theorem for the first time, you might be tempted to try and set up a proof by induction by inducing on the number of parentheses in a sentence directly. This would be a mistake. The basis clause would be sentences with zero parentheses, and then the inductive step would be... I'm not really sure. You might try saying that if all sentences with n parentheses have the same number of left and right parentheses, then all sentences with $n + 1$ parentheses have the same number of left and right parentheses. But that's not going to work, because the theorem is interested in showing that sentences always have the *same number of left and right parentheses*, and adding one would make them unequal. So you might try setting up the inductive step to say something about all sentences with an even number of parentheses, or something like that. In any case, once you got down to trying to prove the inductive step, you'd see that something went wrong.

The trick with this proof is to recognize that parentheses are governed in *SL* as a result of the *Definition of 'Sentence of SL'* way back in chapter 2 of the book. And that definition is organized by *Connectives*. Each connective gets its own step in the definition, and that tells you how to add parentheses—and it tells you to always add *two*—one left and one right. So the induction is upon the *number of connectives*, with reference to the rules of *SL* for sentence formation. Again, understanding the definitions that go in to your theorem is essential for constructing a proof.

Exercise #1

1. Find the *Basis Clause* and *Inductive Step* for each of the following theorems; then prove them by mathematical induction.

   (a) No sentence of *SL* that contains only binary connectives, if any, is TF-False (that is, every TF-False sentence of *SL* contains at least one '∼').

   (b) For all $n \geq 4$, $2^n < n!$.[2]

   (c) Every sentence of *SL* that contains no binary connectives is Truth Functionally Indeterminate.

   (d) Every truth-function can be expressed by a sentence of *SL* that contains no sentential connectives other than '∼', '∨', and '&'.

   (e) $(ab)^n = a^n b^n$

3.3 Solutions to Exercises—with (some!) Proofs

Theorem 7. *No sentence of SL that contains only binary connectives, if any, is TF-False (that is, every TF-False sentence of SL contains at least one '∼').*

Basis Clause: Sentences of *SL* with 0 binary connectives and no '∼' (i.e., atomic sentences) are not TF-False.

Inductive Step: If all sentence of *SL* with n binary connectives or less and no '∼' are not TF-False, then all sentences of *SL* with $n+1$ binary connectives and no '∼' are not TF-False.

The proof of this question is on page 111 of the Solutions Manual. The key is to recognize how sentences of *SL* are "built"; and that all binary truth-functions in *SL* are true when both immediate components are true.

Theorem 8. *For all* $n \geq 4$, $2^n < n!$.

Basis Clause: $n = 4$.

Inductive Step: If $2^n < n!$, then $2^{n+1} < (n+1)!$.

Proof. The basis clause is easy. Where $n = 4$, we have

$$2^4 = 16 < 24 = 4 \times 3 \times 2 \times 1.$$

[2]Recall that '$n!$' (said 'n factorial') is $1 \times 2 \times 3 \cdots \times n$. For example, $5! = 1 \times 2 \times 3 \times 4 \times 5 = 120$.

For the inductive step, we suppose $2^n < n!$. Then we have:

$$\begin{aligned} 2^n &< n! \\ (n+1)2^n &< (n+1)n! \\ \frac{(n+1)}{2}2^{n+1} &< (n+1)! \end{aligned}$$

and since $2^{n+1} < \frac{(n+1)}{2}2^{n+1}$, we get that

$$2^{n+1} < (n+1)!.$$

Thus for all $n \geq 4$, $2^n < n!$. ■

Theorem 9. *Every sentence of SL that contains no binary connectives is truth-functionally indeterminate.*

Basis Clause: Atomic sentences of *SL* are truth-functionally indeterminate.

Inductive step: If all sentences of *SL* containing n or fewer occurrences of '∼' are TF-Indeterminate, then all sentences of *SL* with $n+1$ occurrences of '∼' are TF-Indeterminate.

The proof of this question is again on page 111 of the Solutions Manual. The key is to understand how the truth-function '∼' works. When we negate a sentence, we "flip" it's truth-values. But if the original sentence was TF-Indeterminate, then the resultant sentence will be too (but with opposite truth-values).

Theorem 10. *Every truth-function can be expressed by a sentence of SL that contains no sentential connectives other than '∼', '∨', and '&'.*

Basis Clause: All truth-functions of 1 argument can be expressed by a sentence of *SL* that contains no sentential connective other than '∼', '∨', and '&'.

Inductive Step: If all truth-functions of n arguments can be expressed by a sentence of *SL* that contains no sentential connectives other than '∼', '∨', and '&', then all truth-functions of $n+1$ arguments can be expressed by a sentence of *SL* that contains no connective other than '∼', '∨', and '&'.

The textbook devotes the entirety of chapter 6.2 to proving this theorem, which it lists as **Metatheorem 6.2.1**. It develops an "algorithm" for the proof, giving us a set of instructions for constructing a sentence which represents each possible truth-function. The proof can instead be done by mathematical induction. Try it!

Theorem 11. $(ab)^n = a^n b^n$

Basis Clause: $n = 0$.

Inductive Step: If $(ab)^n = a^n b^n$, then $(ab)^{n+1} = a^{n+1} b^{n+1}$

Proof. Taking our basis clause to be $n = 0$, we have:

$$(ab)^0 = 1 \text{ and } a^0 b^0 = 1 \times 1 = 1.$$

For the inductive step, we suppose that $(ab)^n = a^n b^n$. Then we find:

$$\begin{aligned}(ab)^{n+1} &= (ab)^n (ab) \\ &= a^n b^n ab \\ &= a^n a b^n b \\ &= a^{n+1} b^{n+1}.\end{aligned}$$

Therefore, for all n, $(ab)^n = a^n b^n$. ■

Chapter 4

Conceptual Review Plus: Infinity & Foundations!

4.1 A Conceptual Review of *The Logic Book*

You'll recall that the text began by identifying eight *Logical Concepts* that apply to sentences, sets of sentences, or inferences in a language:

1. L-Truth
2. L-Falsity
3. L-Indeterminacy
4. L-Equivalence
5. L-Consistency
6. L-Entailment
7. L-Validity
8. L-Soundness

These concepts are both *informal* and *general*. They are informal in the sense that we define them in virtue of our intuitive notions of truth, falsity, and the possibility or impossibility of a sentence's being true or false. For example, we said that a sentence is L-True iff it's *not possible* for the sentence to be false. Notice that we gave no further analysis of a sentence's being 'false' or the notion of something being 'not possible', we simply took our intuitive understanding of these notions for granted. When I say that these concepts are also general, what I mean is that they are applicable to our deductive reasoning and indicative statements regardless of complexity or language. This generality is something that our initial informality bought us.

Much of the text amounts to learning two *formal languages*, in order to explicate these properties for certain portions of our reasoning and language:

- Sentential Logic (*SL* and *SD*)—Truth-Functional analysis
- Predicate Logic (*PL* and *PD*)—Quantificational analysis

The resultant TF-Concepts and Q-Concepts can be thought of as rigorizations of the L-Concepts, but applicable to only some subset of our reasoning and language. The purpose of introducing these formal languages is that while we lose the generality of the L-Concepts, formalization and a restriction to these definite levels of analysis gain us (i) Significant precision and understanding; and (ii) Rigorous methods to test for (formal language correlates of) the logical concepts.

For example, restricting ourselves to TF-Concepts means that we can offer only a truth-functional analysis of our sentences and inferences: We break complex sentences down into atomic sentences and the truth-functional connectives. A consequence of this way of analyzing our natural-language arguments is that any L-Valid argument that relies upon logical connections other than those that are truth-functional has no guarantee of being TF-Valid. Consider:

1. All kittens are cute.
2. Merlot is a kitten.

∴ Merlot is cute.

This argument is clearly valid (i.e., L-Valid), but it is not TF-Valid.

But what we lose in generality we gain in understanding. Above I pointed out that we originally gave no analysis of the notion of 'possibility' at work in the definitions of the L-Concepts. With a truth-functional analysis however, we do. Think about it for a moment… Recall that a sentence is a TF-Truth iff it is true on every truth-value assignment (i.e., true on every row of its truth-table). So using *SL* we replace a nebulous and intuitive notion of 'possibility' with the very clear and well-understood notion of 'true on every TV-A'. This is a significant gain in clarity.

Our truth-functional analysis in terms of *SL* also benefits us with the existence of methods—in the case of *SL*, a complete *decision procedure*—for testing finite sets of sentences for the TF-Concepts. For *SL*, these methods were truth-tables and truth-trees, which are no doubt extremely useful.[1] We also gain in the understanding of the scope of our methods. Recall that the text proves in §6.2 that *SL* can represent *every possible* truth-function—or in other words, *SL* is *Truth-Functionally Complete*. What this means is that *SL* offers a complete characterization of language and inference at the level of those logical relationships relying only upon truth-functions.

Finally, the precision of *SL* allows us to develop a corresponding derivation system, *SD*, which models correct (intuitive) human reasoning at this level of analysis. More important, perhaps, is that this system is entirely *syntactical*—it relies upon no assumptions about the meaning or reference of any aspect of *SL*, but is just a system of rules for manipulating signs. Yet the text proves in chapter 6 that *SD* will never lead us from truths to falsehoods (*SD* is *Sound*); and that we can, in a finite number of

[1] Recall from p. 356 of the text that a method is a *Decision Procedure* if it is mechanical, certain, and yields the desired result in a finite number of steps. Such procedures are sometimes called *Effective Procedures* in older works. Languages that have such a procedure to determine—*in every case*—whether the semantic concepts obtain is said to be *Decidable* or *Effective*. *SL* is decidable.

steps, derive every correct TF-Entailment (and so every TF-Truth, TF-Equivalence, and TF-Inconsistency) using just this syntactical system (*SD* is *Complete*).

Extending the language *SL* to *PL* greatly increases the set of sentences and inferences we can rigorously consider. This is because *PL* offers a more complex analysis: We break down *SL*'s atomic sentences as well, into terms and predicates. The resultant Q-Concepts offer a characterization of those logical relationships which rely not just upon the truth-functional connections between sentences, but also upon those connections which depend upon relationships between objects and their properties. So for example, the argument above about Merlot the kitten, while not TF-Valid, is indeed Q-Valid. With the inclusion of identity, *PLE* allows for the analysis of sentences and inferences involving definite descriptions, and about specific numbers of objects.[2]

The text's analysis of the L-Concepts in terms of the Q-Concepts again provides an increase in our understanding. Whereas in the semantics of *SL* we explicate the intuitive notion of 'possibility' in terms of TV-As, in the semantics of *PL* we explicate this in terms of *Interpretations*. More significantly, we add in *PL* an explication of our intuitive notions of 'truth' and 'falsity' in terms of the formal definition of the *Satisfaction of a formula* of *PL(E)*. In essence we provide a set of explicit conditions under which a given interpretation *makes a sentence true*, all of which can be made completely rigorous in mathematical terms.

The trade-off for these significant increases in expressive power and understanding is the loss of a full decision procedure for testing all finite sets of sentences for the Q-Concepts—*PL* is *Undecidable*. Although somewhat surprising, recognize that quantification introduces the need to consider interpretations of our sentences that have infinite universes of discourse. Recall that for any finite set of sentences of *SL*, there will only ever be a finite number of *models* which must be considered (just the number of rows of the relevant truth-table) before we obtain a positive or negative answer as to whether a set of sentences of *SL* has some TF-Property. For every sentence of *PL* however, we must consider an infinite number of models (there are always an infinite number of distinct interpretations for the relevant predicates and constants), some of which have an infinite UD. Thus we may be forced to continue the search for a model *indefinitely* if the answer to our question is negative—say that a set of sentences of *PL* is not Q-Consistent—since a mechanical procedure literally can't decide when to stop the search and say "no models exist".[3]

[2]Notice that this is why the *Löwenheim Theorem* and the *Löwenheim-Skolem Theorem* fail to hold of *PLE*. A sentence like $(\forall x)(\forall y)x = y$ is satisfied only under interpretations with a universe of discourse that has exactly one object. (see pp. 356 and 366 in the text, and Exercises 11.4E.)

[3]It's still useful to extend the truth-tree method to *PL(E)*, especially in the form of *Systematic Trees*. The procedure for systematic trees provides a completely mechanical method for finding a *finite* model of a set of sentences if one exists. But, of course, the method is not definite in all cases (say if a set of sentences is only satisfied in models with an infinite universe of discourse), and so not *Mechanical* in the text's sense.

Given undecidability, it comes as an even bigger surprise that there exists a derivation system for *PL*, *PD*, which is also entirely syntactical and likewise models correct (intuitive) human reasoning at this much deeper level of analysis. And again, while *PD* is just a system of rules for manipulating signs, the text proves in chapter 11 that *PD* is both *Sound* and *Complete*. This is even though *PL* is undecidable and its semantics requires considering infinite numbers of models, some with infinite universes of discourse. All of this holds for *PLE* and its corresponding derivation system too.

These meta-theoretical results, that:

$$\Gamma \models \mathbf{P} \quad \text{iff} \quad \Gamma \vdash \mathbf{P} \text{ in } \textit{SD}$$

and

$$\Gamma \models \mathbf{P} \quad \text{iff} \quad \Gamma \vdash \mathbf{P} \text{ in } \textit{PD(E)}$$

show that our formal, syntactical derivation systems are adequate as models of our correct (intuitive) reasoning, and of our understanding of the logical status of a wide range of sentences and sets of sentences. Backed up by the fact that *SL* is truth-functionally complete and our formal semantics for *PLE*, the analyses the text explores and the methods it teaches are powerful indeed!

4.2 Second- and Higher-Order Logics

A question that remains is whether or not *PLE*, along with its derivation system, provides an *exhaustive* explication of the L-Concepts. In other words, can we adequately represent *all* of our correct intuitive deductive reasoning, or *all* correct applications of the L-Concepts, using *PLE*? Or again: Is every L-Valid inference Q-Valid? The answer is almost certainly 'no'. Consider for example the following sentence:

> Some critics admire only one another.[4]

Where we take as UD: all people, Cx: x is a critic, and Axy: x admires y, we might try to symbolize the above in *PLE* as:

$$(\exists x)Cx \ \& \ (\forall y)(\forall z)((Cy \ \& \ Ayz) \supset (y \neq z \ \& \ Cz))$$

This says: *There exists something that is a critic, and if anything is a critic and admires anything else, then those two things can't be the same thing and the thing admired is also a critic.*

The problem with this as a symbolization is that it misses the idea that the original natural-language sentence is speaking about a *collection* of critics, of which

[4] This example is due to W.V.O. Quine, but was made philosophy-famous by George Boolos' 1984 paper. In that paper Boolos actually *proves* that the second-order formalization below has no first-order equivalent. The idea for the proof is due to philosopher David Kaplan.

none admire themselves or anyone outside of the collection. The sentence isn't just asserting that there is *at least one* critic, or that *there are two* critics, or that there are any definite number of critics at all. Instead, it's using a plural form to talk about *the collection* of critics itself, regardless as to number of members it may have. We can capture this idea by allowing ourselves to quantify not just over objects in the universe of discourse, but also over *collections of objects* in the universe:

$$(\exists X)[((\exists x)Xx \;\&\; (\forall w)(Xw \supset Cw)) \;\&\; (\forall y)(\forall z)((Xy \;\&\; Ayz) \supset (y \neq z \;\&\; Xz))]$$

This says: *There exists a collection of objects, and there exists at least one thing in that collection, and everything in that collection is a critic, and if anything is in that collection and admires anything else, then those two things can't be the same thing and the thing admired is also in that collection (and so a critic).*

If that example doesn't have you convinced, then recall our discussion of the proof-structure *Mathematical Induction*. We observed in chapter 3 above that we can use *PL* to represent induction as a rule of inference in *PD*:

(MI)

$$\mathbf{P}(0)$$

$$\mathbf{P}(\mathbf{k} \leq \mathbf{n}) \supset \mathbf{P}(\mathbf{n}+1)$$

$$\triangleright \quad (\forall \mathbf{n})\mathbf{P}(\mathbf{n})$$

Here we have used the meta-variable '**P**' to stand for any predicate whatever, and I'm using parentheses in an atypical way to indicate that the non-standard or complex expressions '0', '$\mathbf{P}(\mathbf{k} \leq \mathbf{n})$', '$\mathbf{n}+1$', and '**n**' are (meta-variables for) terms.[5] So in that sense the rule as given above is a *schema*—it asserts that this *form of inference* will always be valid. But we should be able to *express* this thought itself as a sentence in a logical language. What this rule really says is:

> **(MI):** *For all properties of the natural numbers*, if they obtain of zero, and their obtaining of every natural number k up to n implies their obtaining of the number $n+1$, then *those properties* obtain of all natural numbers.

Again, we cannot represent this as a sentence in *PLE* because it is not only quantifying over all objects in the universe of discourse (i.e., all natural numbers), but also over *all properties* of those objects. So what we will call the *Second-Order Formulation* of Mathematical Induction is:

$$(\forall P)([Pa \;\&\; (\forall x)(\forall y)((Px \;\&\; Syx) \supset Py)] \supset (\forall z)Pz)^{6}$$

[5] With the exception of the sign '0', which is not a meta-variable but a constant denoting zero.

[6] This is actually a formulation of the so-called *weak* principle of mathematical induction, whereas the version above is the so-called *strong* principle of mathematical induction. Hence the leaving out of the $k \leq n$ part, which would make our symbolization much more complex. Not to worry however, since you will recall that the two versions are in fact equivalent!

Where UD: $\mathbb{N}$, a: 0, and Sxy: x is the successor of y. The symbol 'P' is a *predicate variable*, needed so that we can quantify over *all properties*.

So it seems there are lots of sentences (and arguments) that cannot be adequately expressed in *PLE*. Specifically, it seems that we sometimes have to quantify not only over objects, but also over *predicates* (and relations and functions too), or equivalently, over *collections of objects*.[7] We call this *Second-Order Logic* just because it allows for quantification at the second level—over properties of objects—rather than just at the first level—over objects alone.

The easiest way to extend our formal language *PLE* to a second-order language is to add infinitely many *predicate variables*: $W, X, Y, Z \ldots$, and then to amend our definitions of *Expression*, *Formula*, *Sentence*, *Quantifier*, etc., to allow for quantification over predicates. This gives us sentences such as:

$$(\exists X)Xa$$

which says: *The object a is in the extension of at least one predicate.* We might alternately phrase this: *There exists at least one property which obtains of a.*

Besides being able to express mathematical induction as a sentence in our new formal language (which we will call *SOL* for short), we can take advantage of its expressive power to define various notions whose meanings remained implicit in our less expressive languages *SL* and *PLE*. For example, in *SOL* we can *define* the *PLE* notion of identity as:

Definition 13 (Identity)**.** *Let* $\mathbf{t_1}$ *and* $\mathbf{t_2}$ *be terms, then:*

$$\mathbf{t_1} = \mathbf{t_2} \quad \Leftrightarrow \quad (\forall X)(X\mathbf{t_1} \equiv X\mathbf{t_2})$$

That is to say, the denotation of two terms are identical when they have all the same properties.

This is known as *Leibniz's Law*, after 17th century philosopher, scientist, and mathemagician Gottfried Leibniz. While he proposed this definition of identity, he of course did not have the formal tools necessary to symbolize it. But now *we* do!

What about the semantics and meta-theory of *SOL*? A proper discussion of the semantics will have to wait until after we've introduced some new ideas below. With regard to meta-theory, recognize first that *SOL* is *Undecidable*. This should be unsurprising, since *SOL* is more expressive than *PL*. The typical derivation systems for *SOL* are *Sound*, but there can be no *Complete* derivation system for standard *SOL* so long as we constrain derivations to be *finite* sequences of sentences of the language.

[7] Quantification over predicates and quantification over collections of objects are actually not always equivalent—it depends upon the semantics you're comfortable with. The previously mentioned paper Boolos (1984) discusses this. We will ignore the philosophical and technical subtleties here.

The reasons for this will be mentioned below. Similarly, the *Compactness Theorem* and *Löwenheim-Skolem Theorem* also fail to hold in *SOL*.

The possibility exists of extending our logical languages to be even more expressive. Consider the following sentence:

> Batman's cape is black, and black is a colour.

This sentence not only says something about an object (namely, Batman's cape), but it also says something about a *property*—that it itself possesses a property, what we sometimes call a higher-order property. We might symbolize this sentence as follows:

$$Bb \mathbin{\&} \mathcal{C}B$$

In this case we use the curly-script 'C' to stand for a property of a property, namely '... is a colour'. We might also want to quantify over such positions, in which case we again amend our definitions and introduce the new variables $\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}, \ldots$.

This hierarchy of properties of properties of properties of..., and our ability to quantify over each order, can be extended indefinitely. Taking the union of all these extensions is one way to generate a *Higher-Order Logic* (*HOL* for short). The originators of modern logic at the end of the 19th century—Gottlob Frege, Bertrand Russell, C.S. Peirce, and others—developed *HOL* systems first, without initially recognizing the practicality of isolating the more moderate fragments like *SL*, *PLE*, or *SOL*.

4.3 Infinity with Set Theory

When I introduced talk about *collections of objects* earlier in this chapter, your mind might have immediately gone to *Set Theory*. After all, sets are just collections of objects constructed according to certain rules. Recall that sets can also contain other sets, and so it stands to reason that sets are objects themselves, such that we can quantify over them at the first-order. This is typically the case, but for reasons that shall become clear we'll sometimes continue to use the language *SOL* in what follows.

Set theory was originally developed by mathemagician Georg Cantor (1845–1918), starting around 1870, while researching the properties of infinite series. Cantor needed a way to compare the size of two collections with different kinds of elements in them—say, the number of possible triangles of a certain type as compared to the number of prime numbers of a certain type. Given two sets, perhaps the most obvious way we might determine which set is larger is to count the elements in each set. But this quickly becomes unwieldy for the kinds of sets we deal with in logic and mathematics (just consider the size of many of the universes of discourse we have assumed!), and thinking about infinities in this intuitive way can also lead to problems that have been giving philosophers and mathematicians trouble for thousands of years.

As an example of such conceptual difficulties consider *Galileo's Paradox*, which the famous scientist presented in his last book *Two New Sciences*, published in 1638. This is not the actual text from the dialogue (which is too long to copy here), but hopefully it still gets to the heart of the matter:

> First observe that some numbers are squares, while others are not. Therefore, all numbers (including both squares and non-squares) should be more numerous than just the squares themselves. But on the other hand, for each square there is exactly one natural number that is it's positive square root (e.g., $\sqrt{16} = 4$), and each natural number has only one square (e.g., $3^2 = 9$). Therefore, the natural numbers cannot be more numerous than just the squares themselves. This seems like a contradiction.

The problem that Galileo has encountered here might be characterized as a conflict between two fairly strong intuitions:

1. Euclid's famous dictum: *The part is less than the whole*; and
2. The idea that: *infinity is infinity.*

We can make the first intuition a little clearer by introducing the set-theoretic notion of one set's being a *Proper Subset* of another. Recall **Definition 2** from above, which tells us that a set Γ is a subset of Δ, written $\Gamma \subseteq \Delta$, iff membership in the first implies membership in the second. Now we add:

Definition 14 (Proper Subset)**.** *Let* Γ *and* Δ *be sets, then:*

$$\Gamma \subset \Delta \quad \text{iff} \quad ((x \in \Gamma) \supset (x \in \Delta)) \;\&\; \Gamma \neq \Delta$$

That is to say, a set is a proper subset of another when membership in the first implies membership in the second, but they are not equal (i.e., there is at least one member in the second not in the first).

So for example, the set $\Gamma = \{0, 1, 4\}$ is a proper subset of $\Delta = \{0, 1, 2, 3, 4\}$ because Γ is a subset of Δ and is *missing some elements* that are in Δ. Of course here our sets Γ and Δ are just the first few square numbers, and the first few natural numbers, respectively. So we can say that the set of all square numbers is a proper subset of the set of all natural numbers (squares *and* non-squares), and so, in accord with our first intuition, it seems like there should be less of one than the other.

With regard to the second intuition, we have the idea that the members of an infinite set simply "go on forever", and since both sets have an infinite number of members, the two sets should be the same size—infinite! Galileo observes that every natural number *corresponds* to exactly one square number. In other words, we can make two lists, one with all the natural numbers, one with all the square numbers, and in some sense "pair up" each natural number with its square:

Naturals		**Squares**
0	⟶	0
1	⟶	1
2	⟶	4
3	⟶	9
4	⟶	16
⋮		⋮

So in accord with our second intuition, it seems like the sets are the same size.

So which intuition is correct? Does the set of squares have (i) fewer members than the set of all natural numbers; or (ii) the same number of members? Or is this really a paradox? A fundamental contradiction in our intuitive notion of number, square number, infinity, or some such? Philosopher John Locke, in his 1690 book *An Essay Concerning Human Understanding*, supposes that the very idea of infinity is absurd, just because it leads to such problems:

> [I]f a Man had a positive Idea of infinite… he could add two Infinites together; nay, make one Infinite infinitely bigger than another, Absurdities too gross to be confuted. (p. 222)

By 1891, Cantor had basically told Locke to check himself before he wrecked himself.

The issue is that, because we have no direct experience with infinite sets in our everyday lives, our intuitions easily lead us astray when we try to extrapolate from the finite cases with which we are very familiar. Our discussion of infinity so far has relied upon just our intuitive way of answering the question *how many?* That we can *count* the number of members of a set, even if just in principle, grounds both of our intuitions above. We suppose that a given set is larger than any of its proper subsets because it has *more members* than its subsets—we can count the members of a set *left out* by any of its proper subsets. At the same time, we suppose that trying to count the members of any infinite set just goes on forever, so every infinite set has the *same number* of members. Certainly the idea that two sets are the same size just in case they have the same number of members is true of finite sets, but need it be true of infinite sets? Can we really count an infinite set in the first place? What we need is some way to compare the size of sets *independent of counting.*

In this respect we can take a cue from the game musical chairs. When the music stops, everyone rushes to grab a chair and sit down. We then immediately know whether the set of chairs is larger, smaller, or equal in size to the set of players. If the two sets are indeed the same size, we say that the two sets are *Equipollent.*[8] Whether

[8] Historically people said that the two sets are *Equinumerous*, but this terminology has fallen out of favour for whatever reason.

this relation obtains is obvious by inspection, because each member in the set of players P is paired up with exactly one member of the set of chairs C. If there are no unpaired members of either set at the end, we say there is a *One-One Correspondence* between them, written $P \approx C$. So we have no need to actually *count* the members of either set—we just try to pair them up, and then see whether there are any left over.

This idea, and its formulation in the mathematical realm, was Cantor's big idea. To do this one usually employs *functions* as represented in set theory. Recall that we characterized relations as just sets of n-tuples, e.g., the interpretation of x *is the Captain of* y is the set of ordered pairs $\langle x, y \rangle$,where x is a Captain and y is a crew-member under that Captain. Intuitively, we can think of this as just two lists of people, with the particular relation Cxy represented by all the arrows from the first list to the second for which the relation holds. Of course each Captain may have many crew-members, and so there will be many arrows from any given element of the first list to elements of the second. Recall that a function is just a special kind of relation such that each *argument* has one, and only one, *value*. So for example, the function $y = x^2$ can be thought of as a relation between two sets of numbers, x and y. What's different about $y = x^2$ from Cxy is that there is exactly one square for each number—exactly one *value* for each *argument*.

So back to one-one correspondences. One of the problems with counting in the first place was that the process becomes unmanageable when sets get too large. Actually pairing up the members of each set by hand can be similarly problematic. Instead, we say that two sets are equipollent iff *there exists a function* between the two sets that puts their members in one-one correspondence. More precisely:

Definition 15 (Set Equipollence)**.** *Let* Δ *and* Γ *be sets, then:*

$$\Delta \approx \Gamma \quad \text{iff} \quad (\exists f)(\forall y)(\exists! x)((x \in \Delta \ \&\ y \in \Gamma) \supset y = f(x))$$

That is to say, two sets are equipollent if there exists a function between the first and the second such that for each member of the second set there exists exactly one member of the first set which is the argument of the function for that member.

While this is a lot of new notation, it's all very straightforward. The symbol '!', when combined with the existential quantifier, just means *there exists exactly one....* This replaces clauses of the form: ... & $(\forall z)((\text{conditions}) \supset x = z)$ in any *PLE* sentence that says there's exactly one of something. So this notational shortcut is pretty handy! (No, you can't use it on the exam) In the definition we also quantified over *functions*, stating that at least one function exists. Now that we've introduced *SOL*, we have at least two ways to think about the expression '$(\exists f)$':

(i) We can think of functions as sets (reviewed above), and since sets are objects, we can quantify over them in *PLE* no problem;

(ii) Or we can think of functions as relations, but rather than analyzing relations as sets, we instead quantify over predicate variables directly. In which case the sentence above is in *SOL*.

Either way of thinking about things is fine. Finally, the expression '$y = f(x)$' is just stating that y is the value of the function f, when applied to argument x.[9] So to go back to our example of $y = x^2$, a particular instance would be $x = 3$ and so $9 = f(3)$ (pronounced "nine is *f-of-three*"). But of course we can have functions of things other than numbers, e.g., between players and chairs. So the definition tells us that two sets are the same size just in case there is a function between them that pairs up all the members of each set without remainder. And this is just what we wanted.

Notice that what Galileo did in the second part of his discussion was to *construct a one-one correspondence* between the set of all natural numbers $\mathbb{N}$, and the set of square numbers (call it $\mathbb{G}$). That correspondence is just the one we displayed in the diagram above, with our example $y = x^2$ as the function that demonstrates the correspondence! So it seems that we have:

$$\mathbb{G} \subset \mathbb{N} \quad \text{and} \quad \mathbb{G} \approx \mathbb{N}$$

In other words, the set of square numbers is a proper subset of the set of natural numbers *and* the two sets are equipollent.

But does this mean that the two sets are the same size? To answer this question it will be useful to reflect again upon our conflicting intuitions, and recognize that our intuitive notion of a set's *size* is actually ambiguous. When we say that one set is "smaller" than another set, there are at least two things we might mean: (i) that the "smaller" set is a proper subset of the "larger" set; or (ii) that the two sets are not equipollent. In the case of *finite* sets, these two notions always coincide, but clearly in cases of infinite sets they need-not. This is perhaps why we run into problems when trying to think about infinities in an intuitive way.

What Cantor did to answer this question was to consider the "size" of a set in terms of the notion of equipollence. We take the *Cardinality* of a set Γ, written $|\Gamma|$, to be a measure of the size of the set, and then define:

Definition 16 (Cardinal Equality)**.** *Let Γ and Δ be sets, then:*

$$|\Gamma| = |\Delta| \quad \text{iff} \quad \Gamma \approx \Delta$$

That is to say, two sets have the same cardinality when they can be put into one-one correspondence.[10]

[9] So notice that the expression '$\mathbf{f}(\mathbf{x})$' is a *term*, just like variables and constants. E.g., where we substitute 4 for '$\mathbf{x}$', $f(x)$ will just denote the the object 16.

[10] Notice that we haven't actually *defined* the notion of Cardinality. Instead, like our notion of an *Ordered Pair* from above, we've merely defined a condition (equality) that any such definition must satisfy. Although we can define Cardinality directly, it involves much more technical development. What we have here will do for our purposes.

It follows from our definitions so far that $|\Gamma| < |\Delta|$ iff both:

1. A proper subset of Δ and the whole of Γ can be put into one-one correspondence;
2. The whole of Δ cannot be put into one-one correspondence with any proper subset of Γ.

As an example, consider the sets $\Gamma = \{1, 2, 3\}$ and $\Delta = \{A, B, C, D\}$. In this case, a proper subset of Δ can be put into one-one correspondence with all of Γ—just take any three elements from Δ and pair each one with one of the elements from Γ. At the same time, the whole of Δ cannot be put into one-one correspondence with any proper subset of Γ, because Δ has four letters as members and Γ only has three numbers as members. Therefore $|\Gamma| < |\Delta|$. Note that we can also prove that for any two sets with cardinality $|\Gamma|$ and $|\Delta|$, that either $|\Gamma| < |\Delta|$, $|\Gamma| > |\Delta|$, or $|\Gamma| = |\Delta|$. This result actually requires a lot of work, so we will simply accept it here.

We can think about the cardinality of a set as a *generalization* of the notion of the *number of members of a set*, but without having to *count* anything. Notice that in every finite case, the cardinality of a given set will just equal the cardinality of a set containing the first n natural numbers. So, for example, the cardinality of the number of fingers on my hand equals the cardinality of the set $\mathbb{N}_5 = \{0, 1, 2, 3, 4\}$. In this way, we see that cardinality matches up with our intuitive notion of counting in every finite case. Cool! In the infinite cases however, the definition is more precise than our intuitive notion of *number of members of a set*, because it respects the difference between the notions of *proper subset* and *equipollence*. In fact, we can use this distinction as one way to *define* an infinite set:

Definition 17 (Dedekind Infinite Set)**.** *Let* Γ *be a set, then:*

$$\Gamma \textit{ is Dedekind Infinite} \quad \text{iff} \quad (\exists\Delta)((\Delta \subset \Gamma) \;\&\; (\Delta \approx \Gamma))$$

That is to say, a set is Dedekind Infinite when at least one of its proper subsets is equinumerous to the original set.

A set is *Dedekind Finite* just in case it is not Dedekind infinite. This definition of infinity was proposed by mathemagician Richard Dedekind in 1888. Notice that this definition is *completely independent* from any notion of the natural numbers. Still, all of our intuitively finite sets are indeed Dedekind finite; and as we've seen, $\mathbb{N}$ is Dedekind infinite (so is $\mathbb{G}$). So this ends up being a pretty great analysis of our intuitive understanding of infinity!

So, finally, we can say that:

$$|\mathbb{G}| = |\mathbb{N}|$$

and so the two sets are, according to our definitions, *the same size*. So in this case, our second intuition seems to have been correct. But must this always be the case for two Dedekind infinite sets? Does the maxim *infinity is infinity* always hold? Well, let's examine some number systems, and see whether or not we can find functions which act as a one-one correspondences between them.

Recall that the set of *Natural Numbers*, $\mathbb{N}$, are the counting numbers, $\{0, 1, 2, \ldots\}$. The textbook prefers to leave out zero and talk about the set of *Positive Integers*, $\mathbb{Z}^+ = \{1, 2, 3, \ldots\}$. There is also the set of *Negative Integers* of course, $\mathbb{Z}^-$, and when appended to the positive integers with zero added back to the mix, we get the full set of *Integers*, $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, +1, +2, +3, \ldots\}$. Between each integer are an infinite number of fractions, and we can also represent each integer as a fraction. These constitute the *Rational Numbers*, $\mathbb{Q}$. Finally, we can add all of the irrational numbers, like π and $\sqrt{2}$, to get the *Real Numbers*, $\mathbb{R}$. We already know that the set of square numbers, $\mathbb{G}$ has the same cardinality as $\mathbb{N}$, since Galileo inadvertently proved this by constructing a one-one correspondence between them. But the question remains as to whether or not all these other infinite sets have the same cardinality as well. That is, is this true:

$$|\mathbb{G}| = |\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = |\mathbb{R}|,$$

or is this true:

$$|\mathbb{G}| = |\mathbb{N}| < |\mathbb{Z}| < |\mathbb{Q}| < |\mathbb{R}|,$$

or something in between? Think about it for a moment, then we'll see what Cantor has to say about it.

In what follows we offer *proof sketches* of the relevant results, rather than full proofs. This is because to write everything out rigorously just adds in some technical details and obscures the conceptual elegance of the results. In each case we will try to construct a one-one correspondence between two sets. In fact, this method is so common when dealing with cardinals we can think of it almost like a new proof structure. In any case, on to the proofs!

Theorem 12. $|\mathbb{N}| = |\mathbb{Z}|$

Proof. Notice that if we try to list the set of integers, they trail off in two directions. This makes it hard to put them in one-one correspondence with the natural numbers. However, if we *interlace* the integers (i.e., alternate between positive and negative elements), as Cantor does, then we can construct a correspondence very easily:

$$\begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 0 & +1 & -1 & +2 & -2 & +3 & -3 & \cdots \end{array}$$

Thus $\mathbb{N} \approx \mathbb{Z}$, and so $|\mathbb{N}| = |\mathbb{Z}|$. ■

That result is perhaps not so surprising, given what we already know about the set of square numbers as compared to the set of natural numbers. What about the rational numbers? Note that the rationals are what mathematicians call *Dense*, that is, between every two rationals there is a third. That means that between any two integers (indeed, between any two rationals), there are *infinitely many* rational numbers. We will consider only the positive rationals, for simplicity's sake.

Theorem 13. $|\mathbb{N}| = |\mathbb{Q}|$

Proof. Again, because the rationals are dense, at first glance it seems difficult to find a one-one correspondence with the natural numbers. Here is a table demonstrating Cantor's ingenious method:

		Numerators					
		1	2	3	4	5	$\cdots$
	1	1/1	2/1	3/1	4/1	5/1	$\cdots$
	2	1/2	2/2	3/2	4/2	5/2	$\cdots$
Demoninators	3	1/3	2/3	3/3	4/3	5/3	$\cdots$
	4	1/4	2/4	3/4	4/4	5/4	$\cdots$
	5	1/5	2/5	3/5	4/5	6/5	$\cdots$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Here we have listed all possible natural number numerators and denominators. The interior part of the table thus uses the axes to compose all the rational numbers. To construct a one-one correspondence, f, we pair zero with zero, and pair the rest of the natural numbers up with a fraction, starting with 1/1 in the top-left corner, then proceeding to the right to 2/1, then diagonally down and to the left to 1/2, then down, then diagonally up and to the right, and so on along a path covering every rational. The first part of the correspondence will be:

$$\begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 0 & 1/1 & 2/1 & 1/2 & 1/3 & 2/2 & 3/1 & \cdots \end{array}$$

Thus $\mathbb{N} \approx \mathbb{Q}$, and so $|\mathbb{N}| = |\mathbb{Q}|$. ∎

In fact, this method of enumerating the rationals includes each rational more than once (consider that $1/1 = 2/2 = 3/3 = \ldots$), and so this isn't strictly speaking a one-one correspondence. What we've shown then is that the rationals *plus* the duplicates are still no greater in cardinality than the natural numbers. But since the naturals are a proper subset of the rationals, it follows that they are equinumerous.

So far we have:

$$|\mathbb{G}| = |\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|,$$

and this last one may have been surprising. Before considering the real numbers, it will be helpful to add a few more definitions.

Definition 18 (Countable Set). *Let* Γ *be a set, then:*

$$\Gamma \textit{ is countable} \quad \text{iff} \quad (\exists\Delta)((\Delta \subseteq \mathbb{N}) \ \& \ (\Delta \approx \Gamma))$$

That is to say, a set is called Countable when it is equinumerous with some subset of the natural numbers.

This definition brings us back to the intuitive idea of *counting members* of sets, but with more rigor. We call sets that have the same cardinality as $\mathbb{N}$ *Countably Infinite* sets, otherwise *Countably Finite* sets—more often just *finite* sets. We've seen that $\mathbb{G}, \mathbb{Z}$, and $\mathbb{Q}$ are all countably infinite sets, and are all the same size. The number of fingers on my hand, the number of books on my bookshelf, and the number of atoms in the universe are all examples of countably finite sets, and their members can each be paired up with a set consisting of some finite first part of the natural numbers. So again, this captures a part of our intuitive notion of counting—the finite part. As we have seen, things work a bit differently in the infinite realm. For example, the union of two countably infinite sets is of the same cardinality as the original sets. In fact, the union of countably infinite many countable sets still yields a countably infinite set! And we have seen that a proper subset of a countably infinite set may still be countably infinite. So all of these sets are the same size. We also define:

Definition 19 (Uncountable Set). *Let* Γ *be a set, then:*

$$\Gamma \textit{ is uncountable} \quad \text{iff} \quad \sim (\exists\Delta)((\Delta \subseteq \mathbb{N}) \ \& \ (\Delta \approx \Gamma))$$

That is to say, a set is called Uncountable when it is not equinumerous with any subset of the natural numbers.

Are there any uncountable sets? These would have a strictly *greater cardinality* than $\mathbb{N}$. This brings us to:

Theorem 14 (Uncountability of the Real Numbers). $|\mathbb{N}| < |\mathbb{R}|$.

Proof. We will prove this by contradiction. Recall first that some real numbers can be expressed as non-repeating (or infinitely repeating), non-terminating decimal series. For example, $\pi = 3.1426\ldots$, while $1/3 = 0.3333\ldots$. Assume there exists a function f which is a one-one correspondence between the set of real numbers and the set of natural numbers. For simplicity, we will consider only the set of real numbers between 0 and 1, and we will express terminating decimal series as non-terminating decimal series. So, for example, $0.5 = 0.49999\ldots$. By our assumption, there should be some one-one correspondence as follows:

$$\begin{array}{ccl} 0 & \leftrightarrow & 0.\mathbf{d_1}d_2d_3d_4\ldots \\ 1 & \leftrightarrow & 0.d_1\mathbf{d_2}d_3d_4\ldots \\ 2 & \leftrightarrow & 0.d_1d_2\mathbf{d_3}d_4\ldots \\ 3 & \leftrightarrow & 0.d_1d_2d_3\mathbf{d_4}\ldots \\ \vdots & \leftrightarrow & \vdots \qquad\qquad \ddots \end{array}$$

Notice the bold decimal digits in each of the real numbers on the table. These represent the n-th decimal digit of the n-th real number on the list. We will use these numbers to construct a *new* real number, r, that is demonstrably excluded from any such table, and so, from our assumed function f. We construct r as follows:

> If the n-th decimal digit of the n-th real number on the list $= 1$, then the n-th decimal digit of $r = 2$. Whereas, if the n-th decimal digit of the n-th real number on the list $\neq 1$, then the n-th decimal digit of $r = 1$.

So for example, if the first four lines of our correspondence are:

$$\begin{array}{ccc} 0 & \leftrightarrow & 0.\mathbf{1}329\ldots \\ 1 & \leftrightarrow & 0.6\mathbf{6}66\ldots \\ 2 & \leftrightarrow & 0.82\mathbf{1}3\ldots \\ 3 & \leftrightarrow & 0.451\mathbf{4}\ldots \end{array}$$

then the first part of $r = 0.2121\ldots$. So in this way we have constructed a real number that cannot be in the correspondence, because it will be *different* from every real number in the correspondence in at least one decimal place. This contradicts the assumption that f was a one-one correspondence between $\mathbb{N}$ and $\mathbb{R}$. Thus the reals *cannot* be put in one-one correspondence with the naturals. But since the naturals are a proper subset of the reals, it follows that the cardinality of the real numbers is greater than the cardinality of the natural numbers. In other words, $|\mathbb{N}| < |\mathbb{R}|$! ■

Cantor first proved this extremely celebrated result in 1874. The method of proof used here, called *Cantor's Diagonal Method*, was developed by Cantor in another proof of this result in 1891. It is an extremely powerful method, and can be generalized to prove all kinds of important results in logic and set theory.[11] Some of these general results were also proved by Cantor, involving another important set-theoretic definition:

Definition 20 (Power Set)**.** *Let* Γ *be a set, then:*

$$\mathcal{P}\Gamma = \{x : x \subseteq \Gamma\}$$

That is to say, the power set of a set is itself a set with the characteristic property of having as members all of the subsets of the original set.

Consider for example some finite set $\Gamma = \{A, B\}$. The power set of this set will have four members, that is:

$$\mathcal{P}\Gamma = \{\varnothing, \{A\}, \{B\}, \{A, B\}\}$$

[11] A question you might now have is why we cannot run the same sort of diagonal method on the *Rational* numbers to show they're actually uncountable? In fact you can of course run the same sort of construction, but the new number you generate won't be a rational! It'll be a real number.

In general for any set Γ, its power set will have a higher cardinality than the original set. While this is obvious for finite sets, it is one of the few properties that also carries over to infinite sets.[12] This fact is actually a very important theorem:

Theorem 15 (Cantor's Theorem)**.** *For any set* Γ*, there exists a set which is the power set of* Γ *and is of strictly larger cardinality. That is,* $|\Gamma| < |\mathcal{P}\Gamma|$.

Cantor proved this result with a diagonal argument similar to the one we used to prove Theorem 14. Notice that this entails that the power set of the natural numbers is *Uncountably Infinite*. We can also prove that the cardinality of the natural numbers is the *smallest* infinite cardinality. Finally, notice that this theorem entails that there is actually a never-ending hierarchy of sets with ever-increasing cardinality!

Now, a question you might have is whether the power set of the natural numbers, that is $\mathcal{P}\mathbb{N}$, has the same cardinality as the set of real numbers. This stands to reason, since both $|\mathcal{P}\mathbb{N}|$ and $|\mathbb{R}|$ are uncountably infinite. In fact, Cantor proved it is true that $\mathcal{P}\mathbb{N} \approx \mathbb{R}$, and so $|\mathcal{P}\mathbb{N}| = |\mathbb{R}|$. Of course, Cantor's Theorem still entails that $|\mathbb{R}| < |\mathcal{P}\mathbb{R}|$, so don't go thinking that there's only "one size" of uncountable infinity!

But what about in between $\mathbb{N}$ and $\mathbb{R}$? Are there any sets with a size that falls between these two? In 1878, Cantor himself conjectured:

Conjecture 1 (Continuum Hypothesis)**.** *There is no set* Γ *with a cardinality between the natural numbers and the real numbers, that is:* $\sim (\exists\Gamma)(|\mathbb{N}| < |\Gamma| < |\mathbb{R}|)$.

Or in other words, that the cardinality of the power set of the natural numbers (and so the set of real numbers) is the "next biggest" infinity.[13]

Cantor searched for his entire life trying to find a proof of this conjecture. Most mathematicians since Cantor took it to be a pretty obvious truth. If it doesn't seem obvious to you, recall from above that taking the union of however many infinite sets does not give us a set with a higher cardinality. The only operation we seem to have that always gives us a set with a higher cardinality is the power set operation. It thus came as little surprise when, in 1940, logician extraordinare Kurt Gödel proved that the Continuum Hypothesis cannot be *disproven* in standard formulations of set theory. However, it came as an incredible shock to the mathematical community when, in 1963, mathemagician Paul Cohen proved that neither can the Continuum Hypothesis be *proven* in standard formulations of set theory! The combined result is that the Continuum Hypothesis is *Independent* of set theory, meaning that standard formulations of set theory cannot settle the question. Cohen was awarded the Fields Medal (basically the Nobel Prize for mathematics) in 1966 for his proof.[14]

---

[12] This also holds of the empty set. Recall that the empty set has no members. The power set of the empty set, $\mathcal{P}\varnothing = \{\varnothing\}$, which is a set with one member.

[13] The name of this conjecture derives from the fact that $\mathbb{R}$ is often called the *Continuum* by mathematicians, because the set of real numbers is equipollent to the set of points on a line.

[14] You'll sometimes hear mathematicians, logicians, and scientists (especially Roger Penrose) talk

4.4 Number Theory, Paradox, and Logicism

Cantor, Dedekind, and others developed their theories of sets and infinities using *informal mathematics*, as we have done throughout these notes, and as the text does for the meta-theories of *SL* and *PL(E)*. As we have observed, this is how mathematicians usually work. In Cantor's case, he really had no choice, since the notion of a *formal language* didn't really exist at the time that Cantor started working. Gottlob Frege's publication of his *Begriffsschrift* in 1879 is usually considered to be the first work that makes the notion of a formal language completely explicit. During the next fifty or so years, work along the same lines culminated with Bertrand Russell and A.N. Whitehead's three-volume *Principia Mathematica* (1914–1918), Leopold Löwenheim and Thoraf Skolem's work in the 1910s and 1920s, Kurt Gödel's Completeness (1930) and *Incompleteness* theorems (1931), David Hilbert and Paul Bernay's *Foundations of Mathematics* (1934/1939), Alan Turing and Alonzo Church's work on undecidability, and finally, Alfred Tarski's development of formal semantics from 1933 to 1956.

During this work, the question arose as to whether or not various mathematical theories, notions, and proofs can be *formalized*, or in other words cast in a completely precise formal language like *PLE* or *SOL*. Now, there are many grades of formality. For example, our derivation systems are completely formal, and can be used to represent much of our informal mathematical reasoning. On the other hand, our discussion of set theory has been much less formal. Both Cantor and ourselves have developed set theory by simply *defining* a bunch of notation and providing intuitive explanations of the defined concepts, but we glossed over or left out a bunch of stuff. This is fine, but we must be careful when we do this, because it can lead to problems.

For example, in our discussion of set theory above we introduced the idea of the *Universal Set*. I repeat the definition here:

Definition 8 (The Universal Set, Λ)**.** *The set containing everything. That is:*

$$\text{For any } x,\ x \in \Lambda$$

This idea was really useful for drawing Venn diagrams, but it's actually quite problematic. Just consider that this set will have some definite cardinality. So what is its

about the "incompleteness of mathematics", or "Gödel's Incompleteness Theorems". The following are what they are referring to. In 1931, Gödel proved that *every* formal language expressive enough to represent basic arithmetic will always include sentences which are independent from the language, meaning they can neither be proven nor disproven using a derivation. His proof, while beautiful and definitely correct, doesn't offer a really robust example of a mathematical conjecture that's independent. Instead, it just constructs a formal analogue of the liar paradox. Mathematicians thus hoped that perhaps no important conjectures would ever be independent, and so Gödel's results could be ignored for all practical purposes. The *Continuum Hypothesis* is an example of a significant conjecture that is in fact independent, which is why it made such waves when this was proven. This notion of 'completeness' is different from the soundness/completeness of a derivation system itself. *PLE* is *Sound* and *Complete*, but the theory of formal arithmetic (more on that below!) as formulated in the language *PLE* is incomplete in Gödel's sense.

cardinality? It really should have the *greatest possible* cardinality, since Λ is the set that contains the members of *every other set* as its own members. So suppose:

$$(\forall\Gamma)(|\Gamma| \leq |\Lambda|)$$

But now recall Cantor's Theorem, which presented more formally tells us that:

$$(\forall\Gamma)(\exists\Delta)(\Delta = \mathcal{P}\Gamma \ \&\ |\Gamma| < |\Delta|)$$

In other words, every set has a smaller cardinality than its own power set. Taking $\Lambda = \Gamma$ in the above, this entails that $|\Lambda| < |\mathcal{P}\Lambda|$. But by definition we asserted that Λ has the greatest cardinality, and so it should be the case that $|\mathcal{P}\Lambda| \leq |\Lambda|$. The result is that the power set of the universal set has both greater cardinality than, *and* cardinality less than or equal to, the universal set. This is a contradiction!

This result is… no good. And it really is an honest contradiction, unlike Galileo's only apparent paradox. The problem is that Cantor's Theorem as stated is incompatible with the idea of a universal set. The usual solution is to deny the existence of a universal set, and we will follow that path here.

Cantor himself discovered this situation, and it has come to be known as *Cantor's Paradox*. During the early days of set theory, most mathematicians didn't really worry much about it, since we can just deny the existence of a universal set. There was, however, a much more pressing problem that was soon discovered—one that could not be ignored. Recall that Frege and Russell originally developed modern logic as a *HOL*, from which we have since isolated less expressive but better-behaved languages like *SL* and *PL(E)*. However, Frege's goal was more than just to offer an analysis and formalization of our informal notions of deductive reasoning and indicative sentences—in other words, he wanted to do more than simply explicate the L-Concepts, although this was part of his task. Frege developed the logico-philosophical program of *Logicism*, which means that he conjectured that mathematics can be completely *reduced* to logic. This thesis of *Logicism* really has two parts:

1. All the *concepts* of mathematics can be defined in a purely logical language;
2. All the *theorems* of mathematics can be derived as theorems in a suitable purely formal, logical derivation system.

This is quite a task! But if this could be done, it would provide us great understanding in many areas of mathematics. As with our analyses of the L-Concepts in *SL* and *PL(E)*, by giving completely logical definitions of all mathematical concepts, we could sort out all of the problems surrounding things like our intuitive concept of infinity; and by deriving all theorems in a derivation system, we would be able to see the logical dependencies of each theorem, definition, and concept. More important still, the success of this project would give us an uncontroversial epistemic grounding for our mathematical knowledge—we could be certain that mathematics is *analytic* and can be known *a priori*. Very cool.

As just an example of how this might be done, we can take Frege's language to be the language *SOL* that we discussed above. This isn't exactly historically accurate, but it is close enough for our purposes. Now recall that we can define the notion of *Identity* using Leibniz's Law:

$$\mathbf{t_1} = \mathbf{t_2} \quad \Leftrightarrow \quad (\forall X)(X\mathbf{t_1} \equiv X\mathbf{t_2})$$

Employing a bit of set-theory talk for readability (recall from above that we can take quantification over *collections* to be set-theory talk or *SOL* talk), think about how we might define the concept *Zero*...

$$0 \quad = \quad \{x : x \neq x\}$$

That is, zero is defined as being the set of all things that are not self-identical, which is nothing! So perhaps Frege's idea doesn't seem so crazy.

Even more encouraging was the fact that by the time of Frege's work at the end of the 19th century, mathematicians had succeeded in defining many of the more complicated structures in mathematics in terms of the theory of the arithmetic of the natural numbers. For example, we can construct the positive rational numbers as the quotient of two naturals (i.e., by dividing). Similarly for the integers, the real numbers and much else.[15] So with regard to the first part of Frege's conjecture, the real work was just to define the concepts of arithmetic in terms of logic—and we're already well on the way with zero and identity!

With regard to the second part of Frege's conjecture, we must introduce another new idea—that of an *Axiom System*. Around 300 BCE ancient Greek mathemagician Euclid presented the first clear axiomatic system in the form of an axiomatization of geometry. Of course he didn't have a formal language to work with, so he presented everything in carefully-phrased words and his proofs were like modern mathematical proofs rather than derivations. Still, it was all extremely clear and the proofs valid.

Translating this all into a formal logical language would buy us even more precision, and convince us even further of the validity of the proofs. But the thing about a logical language is that it is very *general*—we can use it to talk about anything. For example, we can take the sentence:

$$(\forall x)(\forall y)(\exists z)Bzxy$$

to mean "*Between any two points on a line there is a third*" so long as we give the sentence the right interpretation. On the interpretation $Bxyz$: x is between y and z on

[15] I cannot express, nor even fathom, the ingenuity and amount of work it took mathematicians to figure all of this out!

a line, with UD: points and lines, this sentence is true. On some other interpretation, the sentence is false—this sentence is *not* Q-True, and so it's not a theorem of *PD*.

However, the sentence *is* a theorem *of geometry*, in the mathematical sense. So another way of thinking about it is to say that it is a G-Truth—it is always true in the theory of geometry. What we want to do then, is to come up with a set of sentences in our logical language that *characterizes* the basic truths of just this topic—just geometry. We call these basic truths the *Axioms of Geometry*, and if we've found the right ones, then all and only the G-Truths should follow from their assumption in a suitable logical derivation system, say *PD*.

Let's make these ideas a little bit more precise. We can give a *Mathematical Axiomatization* of a certain body of mathematical work—geometry, set theory, number theory, etc. This is what Euclid did for geometry, and none of this requires a formal language. Instead, we just write down as axioms what we take to be the most fundamental truths of the theory, and try to prove the rest of the truths from these basic assumptions. These axioms will include sentences like:

> There exist at least two points on a line.
> There exist at least three points that do not lie on a line.

All of this happens at the level of informal mathematical reasoning, but it is still a huge advance over disparate collections of mathematical concepts and theorems.

As mentioned, in the 19th century people started taking this idea to the next level: Formalization in a logical language. We say a *Logical Theory*, T, in a logical language is the set of all sentences of the language that can be derived from the set Γ_T of axioms for T. So in the case of geometry, we will take *PLE* as our logical language. We supplement this language with a few predicates or constants (Px: x is a point; Lx: x is a line; etc.) that get a fixed interpretation, just like our identity predicate in *PLE*. The set of axioms are then symbolized as sentences in this extension of *PLE*. The theory of geometry, T, is then all the sentences of *PLE* such that:

$$\Gamma_T \vdash \mathbf{P}$$

Our theory is considered successful if (i) The theory allows us to derive all and only the truths of geometry; (ii) The axioms aren't inconsistent (we can't end up deriving $\mathbf{Q}$ and $\sim \mathbf{Q}$, unless what we've figured out is that the original topic was itself inconsistent!); and (iii) The set of axioms are mathematically elegant in the sense that the logical theory makes clear why theorems depend upon them, which concepts are related or extraneous, etc.

Today we have very well-understood mathematical axiomatizations, and logical theories, for arithmetic, analysis, geometries of all kinds, modal logics, set theories of all kinds, and most of the rest of mathematics. Indeed, a mathematical topic isn't really considered very clear until there is a suitable axiomatization for it! (both mathematical and logical) We can also give axiomatizations of other things: Newtonian Mechanics, parts of economics, decision theory, etc.[16]

Back to Frege. The second part of the thesis of *Logicism* is partly a call to find a suitable logical theory for the arithmetic of the natural numbers. The first part of the thesis then takes care of any undefined constants or new predicates with a fixed interpretation that must be added to the language, by requiring that they also get a purely logical definition using just the usual logical tools in the language. So this really would be a *reduction* of arithmetic to logic. Again, very lucky for Frege, a couple of mathematicians were at the same time in the processes of discovering a mathematical axiomatization for arithmetic. This set of axioms—which are now the standard axioms for arithmetic—have come to be known as the Dedekind-Peano Postulates, after their discoverers Richard Dedekind (who we've already mentioned) and Giuseppe Peano, in 1888 and 1889.

The main set of axioms are beautifully simple to state, and depend upon only three undefined notions: the constant '0', and the two predicates '$\#x$: x is a number', and '$x\prime$: the successor of x', which are given fixed interpretations. Then we have:

1. 0 is a natural number.
2. For every natural number n, the successor of n (written $n\prime$) is a number.
3. 0 is not the successor of any natural number.
4. For all natural numbers m and n, if $m\prime = n\prime$, then $m = n$.
5. For all properties of the natural numbers, if they obtain of zero, and their obtaining of every natural number k up to n implies their obtaining of the successor of n, then those properties obtain of all natural numbers.

With just these five axioms it's possible to reconstruct *all* of the arithmetic of the natural numbers. Notice that the first four axioms are only true in *Countably Infinite* UDs. This is exactly what we would expect. More significantly, every mathematical theorem about the natural numbers, from the Infinity of the Primes and $2 + 2 = 4$ to $1 \neq 2$, follows from them (once we've defined the necessary concepts within the theory). So we can say that these five axioms really *define* the structure of the natural numbers in the sense that every model of these axioms will be structurally equivalent to the natural numbers. *Question:* What logical language seems most suitable to cast these axioms, if we were to formalize them? *SOL* of course! Since you'll notice that the fifth axiom is just the principle of mathematical induction! And who discovered *SOL*? Frege of course! So Logicism is shaping up nicely.

[16]Logician and philosopher of science Rudolf Carnap took the idea of providing axiomatizations to the extreme—he thought the process was very philosophically useful. In his book *Introduction to Symbolic Logic* (1958) Carnap provides logical theories for all the usual logico-mathematical topics, but also a logical theory for talking about *things*, for parts of physics, and even for parts of *biology*!

The problem that's left is that it's not enough for Frege to simply symbolize each of these sentences in *SOL*, which is easy enough to do. Recall that Frege doesn't just want a logical theory in that sense, where we take these five axioms as automatic premises in every one of our derivations. In other words, he doesn't think that the theorems of number theory are just N-Truths, he thinks they're honest-to-goodness L-Truths! So we have to ask ourselves what allows us to derive certain sentences as theorems in a derivation system like *PD*. The answer of course is that we have a complete and sound set of derivation rules. Keep in mind that the particular set of rules in our textbook are not the only complete and sound set of rules, other textbooks use a different set (like *PD+*) that are equivalent. Instead of having a whole mess of rules, like we do, some textbooks only have two or three rules (usually just ⊃E, ∀I, and a substitution rule), and then a bunch of *logical axioms* that allow us to derive everything else. All of these axioms must be L-Truths, (depending on the logic, TF-Truths, or Q-Truths, or whatever) and they will all be theorems in a corresponding rules-based derivation system like ours.[17]

So Frege's goal is to find a set of axioms that are *L-Truths*, symbolize them in *SOL*, and then derive the Dedekind-Peano Postulates in this system. This is *really* hard to do—as it turns out, it's impossible. But Frege got *really close*, whatever that means. It turns out that everything Frege wants follows from just *one* sentence:

$$(\forall X)(\forall Y)(\#X = \#Y \equiv X \approx Y) \tag{HP}$$

This is now called *Hume's Principle* (HP), because Frege introduces the principle just after talking about philosopher David Hume. What it says is: *The number of Xs equals the number of Ys iff the Xs and the Ys are equipollent.*

It's a little hard to read because on each side of the biconditional is an atypical sentence. As we know from above, on the right-hand side is a sentence that states that two sets can be put in one-one correspondence. On the left is an identity claim. The symbol '#' is like a function symbol from above, in the sense that it takes an argument and spits out a value. But here the arguments it takes are *predicates*, or in other words, properties or collections. Intuitively, we give it a collection, and it tells us the number of objects in that collection. So, for example, if we take Dx: x is an object on my desk, then $\#D$ equals something like 25 or 30 (my desk is very messy). So the expressions '$\#X$' and '$\#Y$' must be *terms*, i.e., they denote *objects*.

Which objects do they denote? The natural numbers of course! So what Hume's Principle is really doing is stating an equivalence between claims about numbers and claims about equipollence (one-one correspondences). This principle is almost exactly what Frege wants in the sense that the following is provable:

[17] An example of an axiom for derivation systems of sentential logic using axioms is $\mathbf{P} \supset (\mathbf{Q} \supset \mathbf{P})$, which is a TF-Truth, and is a theorem of our derivation system *SD*.

Theorem 16 (Frege's Theorem)**.** *The Dedekind-Peano Postulates follow from Hume's Principle in* SOL*.*

This is what Frege proved in his work from 1893–1903. And it is an amazing fact. Everything about the natural numbers can be completely characterized using *just the sentence HP* and second-order logic. In a sense then, the content of our correct judgements about number are entirely explicated by this principle.

But I said above that Frege got "really close", not that he succeeded, so what gives? The problem is that he still needs a definition for the symbols like '$\#X$'. This is why he can't take HP as an axiom of his logical system, he needs to discover some more fundamental axioms that will define this notion. Before giving the axiom he eventually settled upon, it will be worthwhile to spend a minute discussing Frege's extremely deep and celebrated analysis of number.

Frege assumed that numbers are *objects*, just as the text has assumed throughout. His big idea was that while this is true, statements *making numerical claims* are actually about *properties*. According to Frege, when I say "There are 2 objects on my desk" (to make things easier), I'm actually making a claim about the property Dx, that it satisfies the higher-order property "…is a property with two elements in its extension". Why bother with this higher-order business? Because we can *count anything*, even properties. So we need this extra complexity so it works for sentences like: *Your two best qualities are your curiosity and your love of mathematics.*

What does this have to do with defining the numbers, and HP? Recall that we defined *zero* above as the set of all objects that are not self-identical. Frege didn't have set-theory, because Cantor was working on it at the time. Instead, Frege developed his own theory of *extensions*. Our text assumes that extensions just are sets—recall that the interpretation of a predicate in *PL(E)* is just the extension containing all the objects that have that property, in other words, the set. But not all theories of extensions/sets are the same. So what Frege's definition of zero really says is:

> *Zero is the extension containing all objects that satisfy the property of not being self-identical*

Frege needs to say this *within his logical system*. This is why it takes *SOL* (really, *HOL*). More generally, Frege defines the expression *the number of Xs* as:

Definition 21 (#X)**.**

$$\#X \quad = \quad \{x : (\forall Y)(Y \approx X)\}$$

That is, the number of Xs of a property is the extension (set) 'x' containing all properties equipollent with X.

Equipollence is *essential* here so we don't talk about counting, thereby presupposing the natural numbers in our analysis of them!

So for example, if I say "There are two objects on my desk", Frege's analysis takes that to mean: "The property D is in the extension of the higher-order property of *being equipollent with every extension with two members*". The use of the word 'two' here is misleading—the definition is not circular. The last step is to recognize that the extension of a property will have two members just in case that property satisfies this completely logical sentence:

$$(\exists x)(\exists y)([(\mathbf{P}x \,\&\, \mathbf{P}y) \,\&\, x \neq y] \,\&\, (\forall z)(\mathbf{P}z \supset (z = x \vee z = y))$$

which is our familiar way of symbolizing a sentence that says exactly two things have the property $\mathbf{P}x$. Analogously with zero, *number 2 itself* is defined as the extension (set) of all extensions that satisfy the above sentence.

In order to make this all work then, Frege needs to define what an extension is, and he needs to state that every property has an extension. The axiom he settled upon he called *Basic Law V* (Frege called axioms "basic laws"):

$$(\forall X)(\forall Y)(\varepsilon X = \varepsilon Y \equiv (\forall x)(Xx \equiv Yx)) \tag{BLV}$$

Okay... so it looks like I just replaced the symbol '#' from HP with *another* weird new symbol 'ε'. Which I did. But I'm not just messing with you, seriously, we're almost at the end here. What this says is:

> *For all properties X and Y, the extension of the property X equals the extension of the property Y, iff every object that has the property X or the property Y also has the other.*

The symbol 'ε' is like '#' above, but it outputs *extensions* rather than *numbers*. This gets Frege just what he needs—he uses BLV to introduce extensions and derive HP, and then uses HP for everything else. Again, what Frege's BLV states is that the extensions of two properties are identical just when for every object in the UD it has the first property iff it has the second. *This seems really true.*

Unfortunately for Frege, it isn't. Unfortunately for our own informal set theory as well it turns out! The mistake that both Frege and we made in developing our set/extension theory is that we assumed that *every property* determines an extension. This is an extremely plausible assumption—on the face of it. We can state this assumption in our set-theory language as such:

$$(\forall X)(\exists x)(\forall y)(y \in x \equiv Xy) \tag{NC}$$

What this says is: *For every property X, there exists a set that contains just the members that have that property.* This principle is called *Naïve Comprehension* (NC), because at first people just take it for granted—people like Frege and Cantor. We've also assumed it in all our discussions about sets, extensions, and when creating interpretations for our sentences in *PL(E)*. Sure, some properties apply to no objects, so the extensions of those properties will be the empty set, but we never for a minute thought that there are properties that *don't have* extensions.

But in 1902, Bertrand Russell found one. Consider the property of sets/extensions "x is not a member of itself", which we can symbolize as: $y \notin y$. You can see where this is going... Do a second-order $\forall$E on NC above, removing the outer quantifier and instantiating the variable X with our property:

$$(\exists x)(\forall y)(y \in x \equiv y \notin y)$$

Now do $\exists$E using the new constant b:

$$(\forall y)(y \in b \equiv y \notin y)$$

Finally, apply a first-order $\forall$E with the same constant:

$$b \in b \equiv b \notin b$$

This is a whopper of a contradiction! Usually it's called *Russell's Paradox*.

So what does this mean? It means that Frege's program fails. His BLV is inconsistent, but he needs BLV to derive Hume's Principle, and in order to ensure that every property will have an extension so that his definitions work out. Removing BLV makes the rest of Frege's system consistent, but then it is unclear if Hume's Principle is really an L-Truth, most people don't think that it is.

For our own informal set theory as developed in these notes and in the textbook, this is less of a problem. As we've been discussing the theory so far, it is indeed inconsistent. However, this was the case for Cantor's set theory too—Russell's Paradox is a problem for any theory of collections that assumes NC. Once this was discovered, mathematicians developed an *axiomatic* treatment of set-theory that avoids the paradox. Mathematician Ernst Zermelo developed a set of axioms for set theory in 1908 that isn't inconsistent. We were clever in that each of the definitions in our first chapter above (except for Definition 8 about the universal set!) is based upon one of his axioms (we ignored a few axioms that we didn't need), but we didn't state a comprehension axiom at all. Here is Zermelo's version, the *Axiom of Separation*:

Axiom 1 (Axiom of Separation)**.** *Let* Γ *be a set, then:*

$$(\forall X)(\exists x)(\forall y)(y \in x \equiv (y \in \Gamma \ \&\ Xy))$$

That is, there exists a set whose members are all the objects with a given property provided that those objects are already members of another set.

This is very similar to NC. The key difference is that the axiom does not state that the set x exists outright. Instead, it says that x exists *provided that* its members are *already* members of some other set Γ. Notice that Γ *must be given*—in other words, it must already be a known set. So for example, if you're given the set $\Gamma = \{1, 2, 3, 4, 5\}$, then you can use the axiom of separation to construct a new set x whose members consist of all the objects from the original set that have the property *x is an even number*. Again, the key is that the set Γ is already given. Similarly, in giving an interpretation for sentences of *PL(E)*, we must *always* specify a UD, to ensure that we are implicitly relying upon the axiom of separation rather than NC in order to assign extensions to our predicates.

So what about *Logicism*, *Set Theory*, and the theory of *Infinity* that we've been discussing? In each case, we must include an axiom which *explicitly states* that certain sets exist. Recall that we included definitions of the empty set and the universal set. The idea of the universal set is inconsistent, so we throw that one away. That leaves us with the empty set axiom. From the empty set axiom we can iterate the operations of union, subset, intersection, power set, separation, etc. in order to get sets of any finite size. If we want to talk explicitly about infinite sets, then we must add another axiom, the *Axiom of Infinity*, which explicitly states that there exists at least one non-finite set. This is enough to get back everything we've talked about in these notes, and much else besides. However, if such an axiom is always required (and it seems so), it means that *Logicism* (at least in terms of Frege's conjecture) is strictly speaking false. The axiom of infinity is not an L-Truth, it is a substantial assumption about the mathematical world: That an infinite number of mathematical objects exist. This doesn't mean that logic isn't still incredibly useful, it just means that we can't use logic to generate substantial claims by itself. But we learned that lesson in our first class—logic is about *form*, not *content*.

4.5 Conclusion

Congratulations!! This was a long and dense set of notes, but I hope you found them interesting and not too tough. The take-away from these notes is that while the tools of logical analysis and formalization are indeed limited, they are still incredibly *useful*. The amount we have learned about mathematics, the world, and our own reasoning and conceptual abilities by using the tools of logic is astounding and profound. While we can't symbolize *everything all at once*, there are a huge range of arguments, sentences, and concepts that logical analysis clarifies.

4.5.1 P.S.... Derivations in *SOL*

So I said near the middle of this chapter that I'd explain why derivation systems for *SOL* must be incomplete. We've covered enough material now to figure it out, so let's do it! Think about it like this:

The first thing we need is to identify some correlates of the L-Concepts for *SOL*. Let's call them 2-Valid, 2-True, 2-False, etc., and we'll call the best derivation system possible corresponding to these concepts *2D*. The basic fact of the matter is that there are more 2-Truths than possible Theorems in *2D*.

Here's why. In *PL(E)*, we specify a UD and then offer sets of n-tuples of the UD as the interpretation for a predicate. So when we evaluate a sentence of *PL(E)* with a quantifier like $(\forall x)Ax$, it will be true just in case every object x in the UD has the property A. This means that our definition of satisfaction must "check" every object in the UD against every object in the extension of A to make sure that each of those objects is actually in the extension. If the sentence is true, then all of the following will be derivations in *PD*:

$$(\forall x)Ax \vdash Aa \quad (\forall x)Ax \vdash Ab \quad (\forall x)Ax \vdash Ac \quad \ldots$$

Because the UD may be an infinite set, there are a *Countably Infinite* cardinality of such derivations. This isn't a problem.

But now think about the semantics of *SOL*. The semantics of objects and their quantifiers works exactly the same as in *PL(E)*, and so does the interpretation of predicates. However, a predicate like Px in *SOL* acts like a *constant* from *PL(E)*. We can "open" this position by replacing it with a variable—this time a predicate variable! And just like sentences in *PL(E)*, sentences in *SOL* must not have any free variables. So we also need to offer satisfaction clauses for second-order quantifiers (we just add these two clauses to the definition of satisfaction from *PL(E)*). But now, if we evaluate a sentence of *SOL* with a quantifier like $(\forall X)Xa$, it will be true just in case *every property* obtains of the object a. But what do we mean by "every property"? Well, the extension of a predicate is a set of n-tuples of the UD, so we mean *every set of n-tuples* of the UD. This is just the *power set of the UD*. So if this sentence is true, all the following will be 2-Valid:

$$(\forall X)Xa \models Aa \quad (\forall X)Xa \models Ba \quad (\forall X)Xa \models Ca \quad \ldots$$

But if the UD was a *Countably Infinite* set (say $\mathbb{N}$, as usual), then there will be an *Uncountably Infinite* cardinality of such valid arguments!

Finally, consider that we usually take it as a condition of sentences in a formal language that they be *finitely* long. Similarly for *derivations* in a derivation system, since any such system would be pretty useless if the derivation had no possibility of ending! This means that there are only a *Countably Infinite* cardinality of possible derivations in any finite derivation system. So it's impossible to capture all of the above 2-Valid sentences in any finite derivation system—so *2D* is not complete.

Bibliography

Bergmann, M., Moor, J., & Nelson, J. (2014). *The Logic Book*. New York: McGraw-Hill, sixth edn.

Boolos, G. (1984). To Be is to Be a Value of a Variable (or to Be Some Values of Some Variables). *The Journal of Philosophy*, 81: 430–449.

Carnap, R. (1958). *Introduction to Symbolic Logic and its Applications*. New York: Dover Publications.

Hunter, G. (1971). *Metalogic: An Introduction to the Metatheory of Standard First Order Logic*. Berkeley: University of California Press.

Machover, M. (1996). *Set Theory, Logic and Their Limitations*. CUP.

Shapiro, S. (2000). *Thinking about Mathematics*. OUP.

Suppes, P. (1972). *Axiomatic Set Theory*. New York: Dover Publications.

Velleman, D.J. (2006). *How to Prove It: A Structured Approach*. CUP, 2nd edn.