8 A Theory of Sets

Having studied a good amount of arithmetic, we should now move on to studying something a bit more substantial. The theory of arithmetic we studied—despite being the basis for all *"computation"*—was actually very restrictive. It was fundamentally a theory about natural numbers, whose axioms all tried to describe what a *"natural number"* was, and whose quantifiers only range over objects representing natural numbers; these place strong limitations on what we can talk about. For example, it is impossible in the theory of arithmetic to express ideas like the following:

- Every nonempty list $x_1, x_2, x_3, \ldots$ of natural numbers has a least element.[1]
- Any collection of natural numbers with *"positive upper density"* contains arbitrarily long arithmetic progressions.[2]

[1] The *Least-Element Principle.*

[2] *Szemerédi's Theorem.*

In fact, even sentences that *can* be expressed in the theory of arithmetic, like the ones below, might require so much work to write down that we will feel compelled to move to a *stronger*, more *expressive* theory.

- The n^{th} prime number is asymptotically $n \cdot \log(n)$.[3]
- Every natural number greater than 1 has a unique prime factorization.[4]

[3] The *Prime Number Theorem.*

[4] The *Fundamental Theorem of Arithmetic.*

But… setting up a whole new theory is a *lot* of work. Every time we do this, we have to specify a universe of discourse, a list of constant symbols, a list of predicate symbols, a list of functional symbols, and a list of axioms before we can even get started… and *then* we have to start proving all of our basic theorems from scratch, which usually involves a lot of tedious and repetitive work. Ideally, we would have *one* theory in which we could do *all* of mathematics![5] This theory should attempt to capture something *"fundamental"* enough to the mathematical practice that it can be used as a *foundational "model"* for any particular mathematical inquiry we might want to make. But if the natural numbers are not *"fundamental"* enough, then what is?

[5] As it turns out, despite our strongest desires for life to be simple, this is *not* possible. We will give it our best attempt anyways, and this will serve as useful context for those interested in learning more about this in the future.

8.1 What Is a Set Supposed to Be?

It's not unreasonable to say that a good amount of mathematical discussion involves the study of *"quantities."* What is a quantity? We can imagine that a quantity is the result of a measurement performed on a *"collection of things."* Measure different things about the collection, and you might get different quantities. Take a sack of marbles, for a concrete example, and measure its mass, its volume, the number of marbles in the sack, the inertia of the sack, *etc.*—these are all (physical) quantities. More abstractly: the sum of interior angles of a triangle, the magnitude of the greatest common divisor of two numbers, the density of connections in a network of nodes, the greatest prime divisor of a pair of integers—these are all *"measurements"* made about collections of things. It would seem that, in order to have a meaningful notion of *"quantity,"* we need *collections of things* to measure in the first place!

Our idea, then, is to describe more precisely what a *"collection of things"* is. We will call such collections *"**sets**,"* and the theory we will study is the ***theory of sets*** [6] *according to Zermelo and Fraenkel*.[7] We will specify a universe of discourse similar to the one we set up in the theory of arithmetic: the objects of the theory of sets are *those things we can prove exist*, and we call those things ***sets*** by definition. For now,

[6] This is also called *"Zermelo-Fraenkel set theory with Choice,"* commonly abbreviated ***ZFC***. Although we won't study the *Axiom of Choice* itself in detail, we will introduce an axiom that is *equivalent* to choice later this semester. Until then, these axioms will establish what's known as ***ZF*** *set theory.*

[7] The theory is named after Zermelo and Fraenkel, but there were significant contributions to ZFC by John von Neumann, Kurt Gödel, and Thoralf Skolem that should not be discounted.

a set is nothing more than *"something that exists in our theory."* Instead of trying to describe what a set is *informally* through the universe of discourse, we will specify what count as sets and how sets behave through our *axioms,* just like how the axioms of arithmetic (attempt to) describe who the natural numbers are.

∈

Before we get to our axioms, we introduce the symbol $\in$ to denote the binary predicate for ***elementhood***, the *only* predicate symbol in set theory.[1] When say write a sentence like $x \in y$—which we read out-loud in English as *"x is an element of y"*—we can think of this intuitively as meaning something like *"the object x is physically inside of the container y."*

[1] This is not to say that we can't *define* other predicates later, but any new predicates or statements we express must all be written in terms of the two predicate symbols $\in$ and $=$ that we are basing our theory on. Remember that $=$ is a binary predicate that we inherit from the language of the first-order logic (along with its various axioms).

8.2 Do Any Sets Exist?

Well… *something* should exist; otherwise, we'd have an empty universe of discourse. Having just spent a lot of time and energy studying the theory of arithmetic, which was meant to describe the *"natural numbers,"* it would be ridiculous if that time and effort all went to waste—we certainly need to retain the ability to add and multiply numbers[2] when we study more complicated topics. A natural thing to do then would be to assert something along the lines of *"the collection of all natural numbers exists."* This would be our first axiom—***Axiom 0***—an existential axiom guaranteeing not only that there is a set in our universe of discourse, but also that this set contains the natural numbers as we have come to know them.

[2] And perform ***induction***.

Idea: A Proposal for Axiom 0.
The set containing all—and only—the natural numbers exists. We will eventually denote this set by the symbol $\mathbb{N}$, and it will contain the natural numbers 0, 1, 2, … and so on as we have come to know them through our study of arithmetic. ⌟

N.B. An important thing to keep in mind about the set of $\mathbb{N}$ natural numbers and the individual natural numbers 0, 1, 2, … is that they are *all* ***sets*** *because* ***everything is a set*** *in set theory!* Right now, we don't know how to interpret something like 2 as a set; when we revisit this axiom, we will know.

We will come back to this axiom later; before understanding it in detail, we should study the other axioms of set theory, all of which are *universally quantified.*

8.3 What Are Sets?

A set is meant to represent our idea of a *"collection of things,"* and it's difficult to think of something more important to know about a collection of things than *what is* ***in*** *the collection.* This motivates the elementhood $\in$ predicate, which tells us when something is inside of something else. It's important to note through that this is our *only*[3] predicate symbol in the theory of sets. We have no constant symbols, and we have no functional symbols, so any other predicate we want to define for helping us describe sets *must* be expressed in terms of the elementhood $\in$ predicate.

[3] Apart from the equality $=$ predicate, which comes packaged with the language of the first-order logic.

One of the fundamental questions to ask about the theory of sets is: *"What is a set?"* A complete answer to that question requires us to consider *all* of our axioms as a whole, which is not a straight-forward ask of even a seasoned mathematician. A simpler version of this question—whose answer we absolutely *need* to understand if we hope to do *anything* meaningful in set theory—is: *"How do we know when sets are* ***equal****?"*

Axiom 1: Extensionality.
We consider sets to be equal *iff* they have exactly the same elements as each other.

$$\forall x \forall y \big(x = y \Leftrightarrow \forall z (z \in x \Leftrightarrow z \in y) \big)$$

⌟

Our first axiom provides the answer to this question: we consider sets to be equal when they have exactly the same elements as each other. This is what *Axiom 1* is telling us formally: when we say $x = y$ about any two sets x and y, that's equivalent to saying $\forall z(z \in x \Leftrightarrow z \in y)$, and *that* says precisely that elementhood in x is equivalent to elementhood in y. This axiom actually follows immediately from the fact that we only have one predicate symbol in set theory. If we agree with the principle that *equal objects should share all of the same properties*,[1] then any property we can express about x must also hold about y if we want $x = y$; since elementhood is the only thing we can express at the most fundamental level, this means it should *always* be the case, for any object z, that $z \in x$ whenever $z \in y$, and $z \notin x$ whenever $z \notin y$. That is exactly what we mean when we say $\forall z(z \in x \Leftrightarrow z \in y)$.[2]

[1] This is sometimes called the *"indiscernibility of equals"* or *"Leibniz's law."*

[2] Recall, for any propositions p and q, that when we say $p \leftrightarrow q$, we are saying p and q have *the same truth value.* That means they are *both true* or *both false;* they can't have different truth values.

The most important thing we can know about a set is what its elements are. Therefore, we introduce a convenient notation for

Definition 8.1: Set Roster Notation.
Given a natural number n and a finite list of n objects $x_1, ..., x_n$, we define the notation $\{x_1, ..., x_n\}$ to refer to the set whose elements are precisely the objects in that finite list. For every z, we define what it means for z to be an element of $\{x_1, ..., x_n\}$ below.

$\{x_1, ..., x_n\}$

$$z \in \{x_1, ..., x_n\} :\Leftrightarrow (z = x_1 \vee \cdots \vee z = x_n)$$

We should be careful to note that this syntax only makes sense when the number n (and, if possible, the identities $x_1, ..., x_n$) of the objects are known *a priori*.[3] In particular, the use of *"..."* within brackets, as in *"*$\{2, 4, 6, ..., 14\}$*,"* is not defined. ⌟

[3] This notation is ***not defined*** when you, the human, do not know what the natural number n is specifically. Although we often write expressions like $\{1, 2, ..., 10\}$ or $\{0, 1, 2, ...\}$ as a quick way of describing sets by listing their elements,

This provides us with a quick and simple way to describe finite sets by simply listing out their elements. For example, $\{1, 4, 7\}$ refers to the set whose elements are precisely 1, 4, and 7. This implies $\forall x(x \in \{1, 4, 7\} \Leftrightarrow (x = 1 \vee x = 4 \vee x = 7))$, specifying exactly who $\{1, 4, 7\}$ is by *Axiom 1*. We can use this definition to infer that $4 \in \{1, 4, 7\}$ because $4 = 4$, and that $0 \notin \{1, 4, 7\}$ because $0 \neq 1 \wedge 0 \neq 4 \wedge 0 \neq 7$.

8.4 How Can We Construct Sets?

Now that we have a convenient notation for writing down sets by simply listing their elements, we should pause and reflect: when we use *"set roster notation"* and say that $\{0, 1, 2, 3\}$ is a *"set,"* do we *actually* now it is a *"set?"* In other words—since something has to *exist* before we can officially call it a *"set,"* do we know that using set roster notation to take some objects 0, 1, 2, and 3 and put them all in a collection $\{0, 1, 2, 3\}$ together actually corresponds to something that we can prove exists?

Well, so far, we only have one axiom—*Axiom 1*—and that axiom doesn't involve any existential quantifiers, so we can't prove that *anything* exists right now. But we certainly *want* to be able to take finitely many objects and put them together in a set. That's the bare minimum we should expect from the concept of a *"set."* However, trying to write this down in general brings up the problem with our notation, and a fundamental problem with the language of the first-order logic: if we don't know how many elements there are beforehand, then we don't even know how long the sentence we want to write down is, so we have no hope of actually writing it down as an axiom.

Instead of trying to do this in general—meaning, all-at-once for every natural number n in one axiom (or axiom schema)—here's a reasonable alternative. We can give ourselves the ability to construct *unordered pairs* of objects, like $\{0,1\}$, and $\{2,3\}$, and $\{0,3\}$, and $\{1,2\}$. Then, we can give ourselves the ability to *merge sets together* to form larger sets. These two combined will allow us to make any finite set we want—like $\{0,1,2,3,4,5\}$—by first pairing the elements up into separate sets $\{0,1\},\{2,3\},\{4,5\}$, and then merging those sets together into a larger set $\{0,1,2,3,4,5\}$. We'll have to be careful in how we describe this *"merging"* process, but the *"pairing up"* should be pretty uncontroversial, and we can implement it with the following simple axiom.

Axiom 2: Pairing.
Given two sets x and y, we can construct the set z whose only elements are x and y.

$$\forall x \forall y \exists z \forall w \big(w \in z \Leftrightarrow (w = x \vee w = y) \big)$$

An equivalent way to say the same this using *"set roster notation"* is given below, emphasizing that, given two sets x and y, we can construct the *unordered pair* $\{x,y\}$.

$$\forall x \forall y \exists z \Big(z = \{x,y\} \Big)$$ ⌟

While the simplicity of set roster notation makes it convenient to use and reason about, its simplicity is also obviously a serious limitation on our expressive power. In situations where there are too many objects to list one-at-a-time, or when we don't know the identity—or even *the quantity*—of objects we want to talk about, set roster notation fails us. For example, it would be *infeasible* to use set roster notation to explicitly list out the elements of the set of even natural numbers below 100,[1] and it would be *impossible* to talk about the set of *all* even natural numbers. It would be better if we could describe a set by a *property* that is *true* precisely about its elements.

[1] Informally, the set being described is $\{0,2,4,6,8,...,90,92,94,96,98\}$.

Definition 8.2: Set Comprehension Notation.
Given a *wff* φ with at most one free variable, we use the notation $\{z \mid \varphi(z)\}$ to represent the collection of all objects z that make $\varphi(z)$ *true*. Formally, this means that the following sentence is *true* by definition for any object z.

$\{z \mid \varphi(z)\}$

$$z \in \Big\{ x \;\Big|\; \varphi(x) \Big\} \;\; :\Leftrightarrow \;\; \varphi(z)$$

A special form of this notation is easily capable of representing the set of all elements of some particular set X that have the property specified by the *wff* φ.

$\{x \in X \mid \varphi(x)\}$

$$\Big\{ x \in X \;\Big|\; \varphi(x) \Big\} \coloneqq \Big\{ x \;\Big|\; x \in X \wedge \varphi(x) \Big\}$$ ⌟

This notation is easily capable of representing the set of all even natural numbers below 100 as $\{x \mid x \in \mathbb{N} \wedge \exists y(y \in \mathbb{N} \wedge x = 2y) \wedge 0 \leqslant x < 100\}$ and the set of *all* even natural numbers as $\{x \mid x \in \mathbb{N} \wedge \exists y(y \in \mathbb{N} \wedge x = 2y)\}$.

Using this notation, we can describe this idea of *"merging"* together the contents of two sets by defining the pairwise union between sets. But… why would we stop at merging just two sets? Why not give ourselves the ability to merge arbitrarily many sets together at once? Merging very many sets together may produce a very large result, but since we're the ones writing down the axioms, there's no need to hold ourselves back without a good reason.

Definition 8.3: Unions.
Given two arbitrary sets x and y, we define the *binary union between x and y*, which we denote $x \cup y$, to be the set of all things that are elements of either x or y.

$x \cup y$

$$x \cup y \coloneqq \Big\{ z \;\Big|\; z \in x \vee z \in y \Big\}$$

We also define what it means to *iterate* the binary union operation *over* all of the elements of a set. Given an arbitrary set x, we define the *unary union over x*, which we denote by $\cup x$, to be the set of all elements of elements of x.[1]

$\cup x$

$$\cup x \coloneqq \Big\{ w \;\Big|\; \exists z (z \in x \wedge w \in z) \Big\}$$ ⌟

We defined above a binary union operation $\bigcirc \cup \bigcirc$ between sets as well as a unary union operation $\cup\bigcirc$ performed on one set, which conceptually corresponds to an iterated union over *all* of the elements of a set. Let's see why this happens with an example. Consider the set $A = \{\{0,2,3\},\{2,3,5\},\{5,10\}\}$, and let's look at what happens when we take the union over the elements of A and keep track of 0 and 5.

$$\begin{aligned}
\big(\{0,2,3\} \cup \{2,3,5\}\big) \cup \{5,10\} &= \{0,2,2,3,3,5\} \cup \{5,10\} \\
&= \{0,2,3,5\} \cup \{5,10\} \\
&= \{0,2,3,5,5,10\} \\
&= \{0,2,3,5,10\}
\end{aligned}$$

Why does 0 wind up in the final result? Because $0 \in \{0,2,3\}$ and $\{0,2,3\} \in A$. Similarly, why does 5 end up in the union? Because $5 \in \{2,3,5\}$ and $\{2,3,5\} \in A$, but we could just as well have said it's because $5 \in \{5,10\}$ and $\{5,10\} \in A$. This is precisely what the definition of $\cup A$ is trying to tell us because being an element of A means being one of the sets $\{0,2,3\}$, $\{2,3,5\}$, or $\{5,10\}$, and each of their elements belong in the final result.

$$\begin{aligned}
\cup A &= \Big\{ x \;\Big|\; \exists a (a \in A \wedge x \in a) \Big\} \\
&= \Big\{ x \;\Big|\; x \in \{0,2,3\} \vee x \in \{2,3,5\} \vee x \in \{5,10\} \Big\} \\
&= \Big\{ x \;\Big|\; x \in \{0,2,3\} \cup \{2,3,5\} \cup \{5,10\} \Big\} \\
&= \{0,2,3,5,10\}
\end{aligned}$$

Axiom 3: Union.
Given an arbitrary set x, we can collect all of the elements of *elements of x* into a set.

$$\forall x \exists y \forall z \big(z \in y \Leftrightarrow \exists w (w \in x \wedge z \in w) \big)$$

Using the new *"set comprehension notation,"* we can rewrite this axiom as follows.

$$\forall x \exists y \Big(y = \Big\{ z \;\Big|\; \exists w (w \in x \wedge z \in w) \Big\} \Big)$$

Since $\{z \mid \exists w (w \in x \wedge z \in w)\}$ is $\cup x$ by definition, this can be further rewritten.

$$\forall x \exists y (y = \cup x)$$

Each of the three statements above are equivalent reformulations of the same axiom, which asserts that the iterated union over the elements of a set is a set. ⌟

[1] An alternative, maybe more familiar, notation for $\cup x$ is given below.

$$\bigcup_{y \in x} y \coloneqq \cup x = \Big\{ z \;\Big|\; \exists y (y \in x \wedge z \in y) \Big\}$$

If $x \coloneqq \{a,b,c\}$, we get the following.

$$\cup x = \bigcup_{y \in x} y = \bigcup_{y \in \{a,b,c\}} y = a \cup b \cup c$$

This is similar to doing something like

$$\sum_{i=1}^{3} 2^i = 2^1 + 2^2 + 2^3 = 14$$
$$\sum_{y \in \{2^1, 2^2, 2^3\}} y = 2^1 + 2^2 + 2^3 = 14$$

where we iterate the binary $\bigcirc + \bigcirc$ operation over the elements of a set like $\{2,4,8\}$ and accumulate the result. Since binary $+$ on natural numbers produces a natural number, the result 14 we got above was a natural number; when we implement the same idea with the binary $\cup$ operation, which returns a set in general, the answer we get is a generic set.

With this axiom, we not only know that our more general notion of *"union,"* given by iterating the binary union operation over the elements of a set, exists, but we can actually recover specifically the pairwise union between just two sets.

Theorem 8.1: Existence of Pairwise Unions.
$\forall x \forall y \exists z (z = x \cup y)$.

Proof. Let x and y be arbitrary sets. By *Axiom 2*, we know there exists some set p such that $p = \{x, y\}$. By *Axiom 3*, we know there exists some set u such that $u = \cup\{x, y\}$. The claim is that $\cup\{x, y\} = x \cup y$. We will prove this by showing $\cup\{x, y\}$ has all of the same elements as $x \cup y$ and then appealing to *Axiom 1*. For this, let t be a set.

First, suppose $t \in \cup\{x, y\}$. By definition, this means there exists z such that $z \in \{x, y\}$ and $t \in z$. We know $z \in \{x, y\}$ is equivalent to $z = x \vee z = y$ by definition, so saying $t \in z$ is equivalent to saying $t \in x \vee t \in y$. But that means that $t \in x \cup y$ by definition. This shows $t \in \cup\{x, y\} \Rightarrow t \in x \cup y$.

Conversely, suppose $t \in x \cup y$, meaning $t \in x \vee t \in y$ by definition. This implies there exists some w such that $t \in w$ and $w = x \vee w = y$, implying $t \in w \wedge w \in \{x, y\}$. But this is precisely what it means to say $\exists z (z \in \{x, y\} \wedge t \in z)$, so $t \in \cup\{x, y\}$.
This proves $t \in x \cup y \Rightarrow t \in \cup\{x, y\}$.

Therefore, we have shown $\forall z (z \in \cup\{x, y\} \Leftrightarrow z \in x \cup y)$. We now invoke *Axiom 1* and conclude $\cup\{x, y\} = x \cup y$. This shows $\exists z (z = x \cup y)$ as desired.

QED

Lemma 8.1.
The union between two sets shares two familiar properties with other operations.

1. $\forall x \forall y (x \cup y = y \cup x)$ *Commutativity*
2. $\forall x \forall y \forall z (x \cup (y \cup z) = (x \cup y) \cup z)$ *Associativity*

As a result, the order in which we write or evaluate unions does not matter. ⌟

8.5 How Can We Decompose Sets?

Now that we know how to build larger sets out of smaller ones, we should describe how to tear sets apart into their constituent components. While the most important relation in set theory is $\in$ elementhood, suppose we have two sets $A := \{1, 3, 5\}$ and $B := \{0, 1, 2, 3, 4, 5\}$. While $A \notin B$, which means A itself is not an element of B, we can see that all of the *elements* of A are elements of B. We can't say $A \in B$, but we can say that A is *"contained in"* B in some weaker sense—A is something like a *"component"* of B. We formalize this important concept in our next definition.

Definition 8.4: Subsets.
Given any two sets x and y, we say that x *is a subset of* y when every element of x is an element of y. We denote this with the $\subseteq$ symbol, defined formally as follows.

$x \subseteq y$

$$x \subseteq y \;\; :\Leftrightarrow \;\; \forall z \big(z \in x \Rightarrow z \in y\big)$$ ⌟

The $\subseteq$ subset relation between sets is defined based on the $\in$ elementhood predicate, but its role in understanding set theory (and mathematics in general) is just as significant.[1] Most importantly, the $\subseteq$ subset relation has all of the properties of a *partial order*, which is the prototype for things like the $\leqslant$ order on $\mathbb{N}$.

[1] In fact, *Axiom 1* can be restated using the definition we gave of subsets. $\forall x \forall y (x = y \Leftrightarrow (x \subseteq y \wedge y \subseteq x))$

Theorem 8.2: The Subset Relation Is a Partial Order.
The following three universally quantified properties hold for the $\subseteq$ relation.

1. $\forall x(x \subseteq x)$ *Reflexivity*
2. $\forall x \forall y((x \subseteq y \land y \subseteq x) \Rightarrow x = y)$ *Antisymmetry*
3. $\forall x \forall y \forall z((x \subseteq y \land y \subseteq z) \Rightarrow x \subseteq z)$ *Transitivity*

This makes the $\subseteq$ subset relation (*the canonical* example of) a *partial order.* ⌟

Definition 8.5: Some More Binary Set Operations.
For the purposes of these definitions, let x and y be arbitrary sets. There are several important binary operations on sets that, for various reasons, we might like to define.

The *intersection* between two sets x and y is an operation *"dual"*[1] to the union between two sets, containing only those things that are elements of *both* sets.

$x \cap y$

$$x \cap y \coloneqq \left\{ z \;\middle|\; z \in x \land z \in y \right\}$$

The *difference* between x and y, sometimes called the *relative complement of y in x*, is the set of all things that are elements of x but not elements of y.[2]

$x \setminus y$

$$x \setminus y \coloneqq \left\{ z \;\middle|\; z \in x \land z \notin y \right\}$$

Notice that the difference operation defined above is *not* commutative—there are examples of sets for which $x \setminus y \neq y \setminus x$. The *symmetric difference,* however, *is* commutative, and produces the set of all things that are in x or in y but *not in both.*

$x \triangle y$

$$x \triangle y \coloneqq \left\{ z \;\middle|\; z \in x \Leftrightarrow z \notin y \right\}$$ ⌟

As a quick aside, it should be clear that $x \triangle y = y \triangle x$ for any x and y. In fact, we can rewrite the symmetric difference using previously defined operations as follows.

$$x \triangle y = (x \cup y) \setminus (x \cap y) = (x \setminus y) \cup (y \setminus x)$$

The proof is kind of tedious and not very insightful, but the fact itself can be useful.

Now, with some thought, we should be able to see that $x \cap y$ and $x \setminus y$ are both subsets of x, and that $x \triangle y$ is a subset of $x \cup y$. We document how these operations interact through the subset relation in our next lemma.

Lemma 8.2.
The following statements relating the basic binary set operations all hold universally.

1. For any sets x and y, we have $x \cap y \subseteq x$ and $x \subseteq x \cup y$.
2. For any sets x and y, we know $x \setminus y \subseteq x$.
3. For any sets x and y, we can see $x \setminus y \subseteq x \triangle y$ and $x \triangle y \subseteq x \cup y$. ⌟

We definitely want to know that, when we carry out these operations on sets that exist, the results we obtain all also exist, and all of them can be described as particular kinds of subsets.[3] In general, if we are given a set, it should feel reasonable to allow ourselves to say that any definable subset of that set exists—if we are taking elements from a set that already exists, we should be allowed to filter them out according to some property that we can express as a well-formed formula. This is precisely what our next axiom lets us do: *separate* from an existing set a subset of its elements who satisfy some property we can formally express.

[1] What we mean by this will be a bit more clear when we analyze *"power sets"* later, but it is not a coincidence that disjunctions $\lor$ show up in the union $\cup$ definitino, conjunctions $\land$ show up in the intersection $\cap$ definition, and the symbols for $\lor$ and $\cup$ are *literally* just $\land$ and $\cap$ upside-down.

[2] An alternative notation that you may encounter in other places for $x \setminus y$ is $x - y$. We heavily discourage the use of that notation; $x - y$ is not very popular, and is far too easily confused with *"subtraction"* between numbers.

[3] To be more explicit, they can all be described as subsets of sets *that we know exist* either by assumption (like x, who we have assumed is a set for these examples) or thanks to a theorem (like $x \cup y$, whose existence we proved).

Axiom Schema 4: Separation.
For any well-formed formula φ with at most one free variable, the following holds.

$$\forall x \exists y \forall z \Big(z \in y \Leftrightarrow \big(z \in x \wedge \varphi(z) \big) \Big)$$

We can rewrite this equivalently using set comprehension notation as shown below.

$$\forall x \exists y \Big(y = \big\{ z \bigm| z \in x \wedge \varphi(z) \big\} \Big)$$

⌟

Using this axiom, we can prove that each of the operations we defined previously actually produce sets that exist when applied to sets that exist.

Theorem 8.3: Existence of Intersections and Differences.
The binary operations we defined before produce sets when applied to sets.

$$\forall x \forall y \exists z (z = x \cap y) \qquad \forall x \forall y \exists z (z = x \setminus y) \qquad \forall x \forall y \exists z (z = x \triangle y)$$

These all follow by applying *Axiom Schema 4* to $x \cup y$ and using *Axiom 1*. ⌟

What the *Axiom Schema of Separation* lets us do is filter out the elements of a set by searching for those elements that satisfy some property (which we express with a well-formed formula). But what happens if we don't get any results? For example, what if we let $\varphi(x) \coloneqq$ "$x < 0$" and we consider $\{x \mid x \in \mathbb{N} \wedge \varphi(x)\}$, which is exactly the set $\{x \mid x \in \mathbb{N} \wedge x < 0\}$. The *Axiom Schema of Separation* tells us immediately that this set exists, but if n is an element of this set, that would mean that n is a *negative* natural number. There are no such natural numbers! So this set has no elements!

Definition 8.6: Emptiness.
We say that a set x is *empty* if and only if, by definition, the statement $\forall y (y \notin x)$ is satisfied, meaning that x has no elements. ⌟

Theorem 8.4: Empty Sets Exist.
$\exists x \forall y (y \notin x)$.

Proof. Recall from *our proposal for Axiom 0* that we believe $\mathbb{N}$ exists (even if we don't fully understand what that means yet). Then, by *Axiom Schema 4*, we know that there exists a set x such that $x = \{z \mid z \in \mathbb{N} \wedge z \neq z\}$. We will now prove that x is empty. Towards a contradiction, suppose that x is *not* empty, so there exists a set t such that $t \in x$. This means $t \in \mathbb{N}$ and $t \neq t$ by definition, contradicting the fact that $t = t$. ↯ Therefore, we conclude that $\forall y (y \notin x)$, and so an empty set exists.

QED

Definition 8.7: The Empty Set.
We use the notation $\varnothing$ to refer to the collection of all objects that are not equal to themselves. More precisely, we define $\varnothing \coloneqq \{z \mid z \neq z\}$. As it turns out, we are justified in calling this *the empty set* because it *exists,* it is *empty,* and it is *unique.* ⌟

$\varnothing$

Lemma 8.3: The Empty Set is Empty.
$\forall x (x \notin \varnothing)$. ⌟

Corollary 8.1: The Empty Set Exists.
$\exists x (x = \varnothing)$. ⌟

Lemma 8.4: The Empty Set is Unique.
$\forall x (\forall y (y \notin x) \Rightarrow x = \varnothing)$. ⌟

The empty set is fundamentally important to the theory of sets. In some sense,[1] we use $\varnothing$ as the fundamental building block upon which we build all of our other sets. As a result, we should be careful to note how any new definition or concept we introduce in set theory interacts with the empty set.

[1] Which can be made very precise.

Theorem 8.5.
The following statements regarding the empty set all hold.

1. $\forall x(\varnothing \subseteq x)$ and $\forall x(x \subseteq \varnothing \Rightarrow x = \varnothing)$.
2. $\forall x(x \cup \varnothing = x)$ and $\cup\varnothing = \varnothing$.
3. $\forall x(x \cap \varnothing = \varnothing)$.
4. $\forall x(x \setminus \varnothing = x)$.
5. $\forall x(\varnothing \setminus x = \varnothing)$.
6. $\forall x(x \triangle \varnothing = x)$. ⌟

Now that we can separate off from a given set A any definable subset of A that we want,[2] we should immediately feel inclined to collect *all* of a set's subsets into a set of their own. This refers to what we call a *"power set":* the set of all subsets of a given set.

[2] Notice the word *"definable"* in that sentence. The *Axiom Schema of Separation* requires us to write down a *well-formed formula* that we use to filter out elements from A. This means that if we want to filter the elements of A, the property use for filtering *must be expressible in the language of the first-order logic.* It might seem like the first-order logic is capable of expressing anything, but this is certainly not the case. We will get a better understanding of that later when we discuss the cardinalities of sets.

Definition 8.8: Power Set.
Given a set x, we define the *power set of* x to be the set containing all of the subsets of x. We use the notation $\mathbb{P}(x)$ to refer to this set, defined by $\mathbb{P}(x) \coloneqq \{z \mid z \subseteq x\}$. ⌟

$\mathbb{P}(x)$

Axiom 5: Power.
Given a set x, the collection of all subsets of x is a set. We say this formally as follows.

$$\forall x \exists y \forall z(z \in y \Leftrightarrow z \subseteq x)$$

Using set comprehension notation, the above is equivalent to the expression below.

$$\forall x \exists y\Big(y = \Big\{z \;\Big|\; z \subseteq x\Big\}\Big)$$

This is asserting the existence of what we've defined as the *power set* of a given set x.

$$\forall x \exists y\big(y = \mathbb{P}(x)\big)$$

All of these statements are equivalent to each other using different notation. ⌟

Lemma 8.5.
For every set x, we have that $\varnothing \in \mathbb{P}(x)$ and $x \in \mathbb{P}(x)$. ⌟

Corollary 8.2.
For every set x, we have that $\cap\mathbb{P}(x) = \varnothing$ and $\cup\mathbb{P}(x) = x$. ⌟