

PHL349

Set Theory

by Brandon Papandrea

The following text is based on lecture notes written during the Winter 2026 offering of PHL349 - Set Theory at the University of Toronto, and are in large part based on the textbook *Elements of Set Theory* by Herbet Enderton, but also include notes based on the books *Conceptions of Sets and the Foundations of Mathematics* by Luca Incurvati, and *The Joy of Sets* by Kieith Delvin. The inention for these notes is to be a more polished and complete version of my own handwritten notes I take during class, and thus are not intended to be a primary source for learning set theory.

# Contents

---

<b>1</b>	<b>Lecture 1: Two Conceptions of Sets</b>	<b>1</b>
1.1	Why Should I Care About Sets? . . . . .	1
1.2	What is a Set? . . . . .	1
1.2.1	The Logician Conception of Sets . . . . .	1
1.2.2	The Generative Conception of Sets . . . . .	4
1.2.3	Which is Better? . . . . .	5
<b>2</b>	<b>Lecture 2: The ABCs of Set Theory</b>	<b>7</b>

# Lecture 1: Two Conceptions of Sets

---

In the opening lecture of this course, we seek to understand what sets are through a philosophical lens by considering two different ways of conceptualizing them. Each have their pros and cons.

## 1.1 WHY SHOULD I CARE ABOUT SETS?

Any discussion about sets, especially at the level of formal set theory, will eventually lead to someone asking “what on Earth is the point of all this?” This is reasonable to ask considering how most math students, even those seeking to do research in the subject, will likely never have to deal with sets in this manner.

The reason why we care about sets is, in short, because sets have a maximum unifying power in mathematics. There are two facts that make this true:

1. Every concept in math may be defined in terms of sets, and all mathematical statements can be expressed using the language of set theory (a first-order language that has a single binary relation, membership).
2. All of mainstream math can be developed within the theory of sets, which is known as **Zermelo-Fraenkel with Choice**, or just **ZFC**.

The first fact follows from the power of the first-order language, which we call  $\mathcal{L}_\varepsilon$ , and that such first-order languages are sufficient in expressing mathematical reasoning. The second fact comes from how ZFC was specially constructed to allow us to formalize mathematics. Mathematicians and philosophers like Dedekind, Cantor, Frege, Russell, Whitehead, and finally, Zermelo & Fraenkel, worked over decades to find this natural theory in the late 19th and early 20th centuries. The fact that ZFC was chosen and not some other theory has a complicated history that we won't go into.

## 1.2 WHAT IS A SET?

The question of the day is the most basic one can ask in set theory: what exactly *is* a set? Answering this is not at all easy. We will discuss two schools of thought. In one, we say sets are defined, while in the other, we say they are constructed.

### 1.2.1 THE LOGICIST CONCEPTION OF SETS

Georg Cantor defined sets as a “totality of definite elements that can be combined into a whole by a law.” What did he mean by “law?” To some, like Frege and Crispin Wright, a law is

a concept, or in other words, like a predicate in first-order logic. This forms the basis of the **logician** conception of a set.

The key idea in this school of thought is as follows: given any predicate/concept, there exist a set containing only those things to which the predicate applies. This set is called the **extension of the predicate**. In this way, set theory is just a part of logic, like all other parts of math.

This conception of sets is based around two fundamental principles that each answer two important questions:

1. **Existence:** What sets are there?
2. **Identity:** How can we distinguish between two sets?

We can formalize these questions, and hence this conception of sets, in logical terms. Let  $\mathcal{K}$  be a suitably powerful language, whose variables range over all sets and objects that are not sets (called individuals). Let  $S$  mean ‘is a set,’ and  $\in$  mean ‘is a member of.’ The logicist principle of set existence can then be expressed as follows: If  $\phi$  is a formula of  $\mathcal{K}$ , then there should be a set of just the things to which the predicate applies. In logic terms, if  $\phi(x)$  is some well-formed formula of  $\mathcal{K}$ , we write

$$\exists y[Sy \wedge \forall x(x \in y \iff \phi(x))]$$

in the logicist theory, this is an axiom. In broader terms, any instance of the *comprehension schema* show up in the logicist theory as *existence postulates*.

The second principle is an identity principle. Think of how one would differentiate between two sets of, say, pencil crayons. If you can show that one set contains a coloured pencil not found in the other set, then we know that the sets are different. From the logicist viewpoint, this idea - differences in membership - is the only way to distinguish between sets:

$$\forall x \forall y [(Sx \wedge Sy \wedge x \neq y) \implies \exists z ((z \in x \wedge z \notin y) \vee (z \notin x \wedge z \in y))]$$

We can also write this another way: if two sets have exactly the same elements, they’re the same set:

$$\forall x \forall y [(Sx \wedge Sy) \implies (\forall z (z \in x \iff z \in y) \implies x = y)]$$

This idea is called *extensionality*.

By combining the comprehension schema and the extensionality schema, we get a theory commonly referred to as **naïve set theory** (you’ll see in a second why we use the word naïve). Some of its axioms are as follows:

1.  $\exists y[Sy \wedge \forall x(x \in y \iff x \neq x)]$  (There is an empty set)
2.  $\forall z \forall w \exists y[Sy \wedge \forall x(x \in y \iff (x = z \vee x = w))]$  (There is a pair set)
3.  $\forall z \exists y[Sy \wedge \forall x(x \in y \iff \exists w(w \in z \wedge x \in w))]$  (There is a union set)
4.  $\exists y[Sy \wedge \forall x(x \in y \iff (Sx \wedge x = x))]$  (There is a universal set)

## Russell’s Paradox

Frege didn’t state his logicist theory like this. Instead he used a second-order language and made two logical assumptions:

1.  $\exists X \forall y (Xy \iff \varphi(y))$  for all second-order well-formed formulas  $\varphi$  (comprehension schema)
2.  $X \equiv Y \iff \forall z (Xz \iff Yz)$  (extensionality schema)

But, Frege made a very small yet fatal error. He made two additional assumptions about the existence of concept extensions. He assumed that every concept has an extension. More importantly, he assumed that distinct concepts have distinct extensions. This is known as **Law V**. comprehension and extensionality for sets then comes from defining a set as just the objects that are concept extensions, and membership in terms of extensions and predication: If we let  $Qxy$  mean ‘ $x$  is an extension of  $Y$ ,’ then

1.  $Sx := \exists Y [QxY]$  (an object is a set if it extends some concept)
2.  $x \in y := \exists Y [QyY \wedge Yx]$  (an element of a set is an object that satisfies the corresponding concept)

The effect of these assumptions is that there are as many objects (sets), as there are concepts (predicates). But this does not fall in line with Cantor’s theory that the power set of a set is larger than the set itself.

This leads us to the famous **Russell’s Paradox**, which shows that naïve set theory is inconsistent. We consider the well-formed formula

$$Uy \wedge \forall x [Rxy \iff (Ux \wedge \neg Rxx)]$$

We can break this up into two parts at the outside  $\wedge$  symbol:

$$Uy \quad \text{and} \quad \forall x [Rxy \iff (Ux \wedge \neg Rxx)]$$

So  $Uy$  is true. We instantiate the right side by setting all  $x$ ’s to be  $y$ ’s. Giving us

$$Ryy \iff (Uy \wedge \neg Ryy)$$

as  $Uy$  is true, we just get that

$$Ryy \iff \neg Ryy$$

a clear contradiction. This means that the negation of this entire formula is true:

$$\neg \exists y [Sy \wedge \forall x [x \in y \iff (Sx \wedge x \notin x)]]$$

but notice that this is an instance of the comprehension schema. We have that this instance of the scheme is both true and false, hence the theory is inconsistent. This idea was first shown to Frege via a letter written by Bertrand Russell in 1902.

There are a handful of ways people have thought of to fix this problem. The first way involves modifying our comprehension schema to make it more predicative. One may argue that, since we are defining a set using a quantifier that runs over all sets, which would include the set we’re defining, we are using circular logic. To remedy this, we can either ban the use of quantifiers in our scheme, or use some sort of level system: variables are put into levels  $0, 1, 2, \dots, n, \dots$ . Then a set of level  $n$  may only be defined using variables of levels below  $n$ . This is quite complicated, and leads to more complexity than its worth, at least for some.

Another way is to look back at Frege's assumption and modify it: instead of saying all concepts have extensions, we say that only "safe" ones do. What is "safe?" That takes some work, and many of done work in this area. Some, however, have argued that the idea that certain concepts have no extensions, called the **limitation of size** is unnatural.

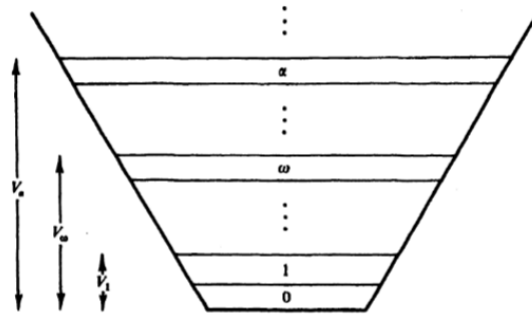
### 1.2.2 THE GENERATIVE CONCEPTION OF SETS

The second viewpoint we will discuss is to think of sets as not defined, but rather, *generated* in some iterative process. We can talk about this process using informal language, like 'stage,' 'is formed at,' 'earlier than,' etc.

We start with individuals, things that are not sets, and at Stage 0, form all possible collections of these individuals. If none exist, we only form the null set.

At Stage 1, we form all collections of individuals and sets formed at Stage 0. At Stage 1, we form all collections of individuals and sets formed at Stage 1. We keep going like this, and eventually reach Stage  $\omega$ , the stage immediately stages 0,1,2, and so on. We can still keep going, performing transfinite iteration, with stages  $\omega + 1, \omega + 2, \dots, 2\omega, \dots, 3\omega, \dots, \omega^2, \dots, \omega^\omega, \dots$

One can visualize this as an infinite large cone. At the base we have the collections of just individuals; Stage 0. Each level of the cone is a stage, which is built on top of the stages built previously.



Under the generative conception, every set is repeatedly formed at each stage. More importantly, no set is an element of itself, as a set is formed at some earliest stage, and its elements must have been formed earlier than that. Thus, there is no universal set. Using a similar argument, there are no sets  $x, y$  such that  $x \in y$  and  $y \in x$ . Thus, the set of all sets that do not contain themselves is just the set of all sets, which we know cannot exist. So, under the generative conception, Russell's Paradox is avoided!

### The Theory of Stages

Because we have introduced the idea of stages, we need to formalize the theory of stages in logic. Our language  $\mathcal{S}$  will have two types of variables:  $x, y, z, w, \dots$ , which range over sets, while  $r, s, t, \dots$  will range over stages. We also include two new relational symbols:  $E$  for 'is earlier than,' and  $F$  for 'is formed at.' We can now state some axioms:

1.  $\forall s \neg E s s$  ( $E$  is anti-reflexive)
2.  $\forall r \forall s \forall t [(E r s \wedge E s t) \implies E r t]$  ( $E$  is transitive)

3.  $\forall s \forall t [Est \vee Ets \vee s = t]$  ( $E$  is connected)
4.  $\exists s \forall t [s \neq t \implies Est]$  (There's an earliest stage)
5.  $\forall s \exists t [Est \wedge \forall r (Esr \implies (Etr \vee t = r))]$  (Each stage has a unique successor stage)
6.  $\exists s [\exists t Ets \wedge \forall t (Ets \implies \exists r (Etr \wedge Ers))]$  (There's a limit stage different from the first stage, but with no direct predecessor)
7.  $\forall x \exists s [Fxs \wedge \forall t (Fxt \implies t = s)]$  (Sets form at some unique stage)
8.  $\forall x \forall y \forall s \forall t [(y \in x \wedge Fxs \wedge Fyt) \implies Ets]$  (All elements of a set form before the set)
9.  $\forall x \forall s \forall t [(Fxs \wedge Ets) \implies \exists y \exists r (y \in x \wedge Fyr \wedge (t = r \vee Etr))]$  (Sets form immediately after their elements are )

That ninth axiom is important, as it imitates the comprehension schema: at every stage all definable sets of previously formed elements are formed. How are they defined? Using the language  $\mathcal{S}$ . Formally, if  $\chi(x)$  is an  $\mathcal{S}$  formula that has no free occurrence of  $y$ , then

$$\forall s \exists y \forall x [x \in y \iff (\chi(x) \wedge \exists t (Ets \wedge Fxt))]$$

is an axiom of stage theory, called the **specification axiom**.

We also need to take care of one other thing: induction. Since we introduced stages in an inductive way, there better be a formal way of defining induction on stages, like there is a way of formally defining induction on numbers. We omit this here because it's not entirely useful for our purposes, but the work has been done and can be found.

The long and short of our discussion on stage theory is that it has a surprising amount of power. In fact, stage theory can prove the axioms of the empty set, pairing, union (in a general form), power-sets, and the axioms of infinity and separation (which says that for a set  $x$  there is a subset of elements of  $x$  satisfying a formula). This is almost enough to prove all of ZFC. However, there are some things stage theory can't prove that are required for ZFC. For example, the replacement schema, which says that the image of a set under a function is also a set. The axiom of extensionality is taken to be analytic, in the sense that it is an essential property of the notion of a set. Most importantly, the axiom of choice cannot be derived. Even still, stage theory can derive almost all of set theory, so it is quite powerful.

### 1.2.3 WHICH IS BETTER?

Is one of these conceptions better than the other, or is there a compromise we have to make? Some have argued that neither of these viewpoints is sufficient in justifying ZFC, and so we will have to make choices about which viewpoint to use depending on circumstances. How do we decide? That's a question outside the scope of the course, but the curious reader can find answers.

A particularly interesting concept that arises in both of these conceptions, more so in the logicist approach because of the limitation of size, is the idea that the set-theoretic universe is **ineffable**. We can talk about sets, and we know there is some set of all sets, but can we say anything about the entire universe we are working in. More concretely, consider the cone from the generative conception. We can talk about individual sets in the cone or levels of the cone, but can we talk about the cone itself using the language of set theory? This idea actually has



some applications to theology, since God acts similarly to this set-theoretic universe. Many have argued that, while God cannot be described using language, language can be used to approach God. There is an analagouse idea in set theory: while it may not be possible to describe our universe using the language of set theory, we can approach it. There is work to be done in this direction.

## Lecture 2: The ABCs of Set Theory

---