I. Foundations

UBC M320 Lecture Notes by Philip D. Loewen

Sets

2026-09-09

Russell's Paradox (1901). Imagine classifying sets by declaring

$$\text{a set } S \text{ is } \textbf{regular} \ \Leftrightarrow\ S \notin S.$$

Clearly every set is either regular or irregular, and it is impossible for a set to be both at once. Now let R be the collection of all regular sets. Classify R.

- Could R be regular? If so, write the definition: $R \notin R$. Here the RHS (R) contains all the regular sets, so the LHS (R) must be irregular. This can't happen.
- Could R be irregular? If so, we must have $R \in R$. But again, every set in the RHS (R) is regular, so the set on the LHS (R) must be regular. This is impossible, too.

Something terrible is happening here. The regular/irregular test seems to split the collection of all sets into distinct groups, yet R cannot logically belong to either one. What's wrong? *It's not the logic, it's the setup.* Our intuitive, colloquial idea of "sets" and "elements" is too elastic. We need strict foundations for even these fundamental concepts, in which "set" becomes a defined term that the collection R above does not deserve.

ZFC. Standard mathematics is built on the 9 "ZFC axioms", named after Zermelo, Fraenkel, and Choice. Here Zermelo and Fraenkel are people's surnames (for Ernst Zermelo, 1871–1953, and Abraham Fraenkel, 1891–1965), and Choice refers to a powerful axiom to be discussed later. Precise formulations and discussion are available on Wikipedia. In the ZFC universe, each and every "object" is literally a set, and the binary relation "$\in$" is the single fundamental source of the logical values True and False. It is implicit in the setup that at least one set exists; the explicit axioms then imply that there is a set $\emptyset$ with the property that $x \in \emptyset$ is false for every x. (Which x are active in the phrase "for every x"? All the objects in the ZFC universe, i.e., all sets.) We call $\emptyset$ the empty set, and allow the alternative notation $\{\,\}$. (Rudin writes 0, but we prefer the distinctive symbol shown here.)

Of course we can build composite assertions with T/F values using logical operations like "and", "or", and "not", but they all ultimately reduce to checking instances of set membership.

Here are 7 of the ZFC axioms. (We omit Replacement—Fraenkel's contribution, which we won't rely on, and Choice, to be discussed later. Also, the numbering scheme here is not universal; the names are closer to being standardized.) They refer to arbitrary sets A, B, and $\mathcal{F}$.

1. Extensionality: $A = B$ if and only if both $A \subseteq B$ and $A \supseteq B$. In symbols (which implicitly define the relations just mentioned),

$$A = B \iff [\forall a \in A,\ a \in B] \wedge [\forall b \in B,\ b \in A].$$

2. Pairing: $\{A, B\}$ is a set.
3. Union: $\bigcup \mathcal{F} = \bigcup_{F\in\mathcal{F}} F = \{x : \exists F \in \mathcal{F} \text{ s.t. } x \in F\}$ is a set.
4. Separation: Whenever ϕ is some logical formula (i.e., a finite-length expression for which $\phi(x)$ returns a T/F result for each $x \in A$),

$$\{x \in A : \phi(x)\} \text{ is a set.}$$

   (Experts call this an *axiom schema*, because every choice of ϕ generates its own axiom.)
5. Power set: $\mathcal{P}(A) = \{B : B \subseteq A\}$ is a set.
6. Infinity: There exists a set Ω such that $\emptyset \in \Omega$ and for each x in Ω, the set $x \cup \{x\}$ is also in Ω.
7. Regularity: Every nonempty set S contains an element x such that $S \cap x = \emptyset$.

Axioms 2–5 describe operations that build new sets from known ones. Axiom 4 gives us a license to form subsets using logical criteria; Axiom 5 implicitly ensures that *every* subcollection of a set is also a set. (This matters: even for the set of integers, no collection of criteria will be able to capture every last subset.) Axiom 6 asserts the existence of a set that lies just beyond the reach of the others (see "Numbers", below).

It should be noted that the set-building axioms of ZFC do not allow the collection of all sets to qualify for the term "set". Reasoning about this collection calls for new terminology and insights beyond the scope of our course.

Axiom 7, which we seldom use explicitly later, helps explain why Russell's paradox cannot arise in ZFC. Given any set A, take $B = A$ and use axiom 2 to build the set $S = \{A\}$. Clearly $S \neq \emptyset$, so there must be some x in S with $S \cap x = \emptyset$. The only choice for x is A, so $S \cap A = \emptyset$. This means every $x \in S$ has $x \notin A$. Again, the only x in S is $x = A$, so we have $A \notin A$. Thus every set in the ZFC universe is "regular" in the sense used to state Russell's paradox above. This makes the collection of regular sets the same as the collection of *all* sets, and that collection is *not a set*.

$\boxed{\text{2026-09-11}}$

Literalism. Precise thinking and communication require care with the *symbols* as well as the *words*. Critical distinctions include the following.

- The symbols $\in$ and $\subseteq$ have different meanings. ("$\in$" is irreducible, whereas $A \subseteq B$ is a quantified statement, namely, $\forall x \in A,\ x \in B$.) So, e.g., $x \in \emptyset$ is False for every set x (remember that sets are the only objects we consider), and $\emptyset \subseteq x$ is True for every set x.
- Braces carry meaning. If A is a set, then $\{A\}$ is also a set, and $A \neq \{A\}$. (Indeed, the inclusion $\{A\} \subseteq A$ requires that $A \in A$, and we saw earlier that there are no sets A with this property. But this inclusion is an essential part of the statement $A = \{A\}$, so that statement must be false.)

In building new sets from known ones, the following standard approaches are consequences of the axioms:

(i) Intersection: If $\mathcal{F} \neq \emptyset$, $\bigcap \mathcal{F} = \bigcap_{F \in \mathcal{F}} F = \{x : \forall F \in \mathcal{F},\ x \in F\}$ is a set.

(To get this from the axioms, pick any $F_0 \in \mathcal{F}$ and recognize the desired result as a *subset of* F_0, etc.)

(ii) Cartesian Products (Tuples): Given a finite collection of sets, say $A_1, A_2, \ldots, A_n$, the following construction produces a set:

$$A_1 \times A_2 \times \cdots \times A_n = \prod_{k=1}^{n} A_k = \{(x_1, x_2, \ldots, x_n) : \forall k \in \{1, \ldots, n\},\ x_k \in A_k\}.$$

In this set, generic elements $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ obey $x = y$ if and only if $x_k = y_k$ for each $k = 1, 2, \ldots, n$.
(Detail: When $n = 2$, treat tuple notation (x_1, x_2) as a convenient shorthand for the set $\{\{x_1\}, \{x_1, x_2\}\}$. Check that equality of sets corresponds exactly to the usage of "=" proposed for tuples. Then iterate: interpret the 3-tuple (x_1, x_2, x_3) as a reader-friendly reformatting for the set corresponding to the ordered pair $((x_1, x_2), x_3)$, etc.)

Practice. (i) Find nonempty sets A and X such that both $A \in X$ and $A \subseteq X$.

[Ans: $A = \{\emptyset\}$ and $X = \{\emptyset, A\} = \{\emptyset, \{\emptyset\}\}$ do the job.]

(ii) Suppose S is a set. Express the T/F statement, "the set S has exactly one element" as a logical test based on the axioms above.

[Ans: Use S to define $C = S \times S$, as above. Then let $D = \{(x, y) \in C : y = x\}$: this is a set, by the Axiom of Separation. Clearly, if S has more than one element, then $D \neq C$. So S has exactly one element exactly when $[C = D] \wedge [S \neq \emptyset]$.]

Further Study. Axiomatic systems and the logic that drives them come in several distinct variants, and these topics have their own interest. At UBC the Department of Philosophy is the home for this material: interested students can start exploring it in PHIL_V 222, *Enriched Symbolic Logic*.

Numbers

The axiomatic foundation for analysis rests on an exact matching between the numbers and calculations we use every day and a careful specification of *objects* and *operations*. Here the objects are, of course, sets, and the operations are … sets (of course).

Counting. Here is an infinite list of sets, all different, with suggestive notation for each one:

$$\begin{aligned}
0 &:= \emptyset, \\
1 &:= 0 \cup \{0\} = \{0\}, \\
2 &:= 1 \cup \{1\} = \{0, 1\}, \\
&\ \vdots \\
n+1 &:= n \cup \{n\} = \{0, 1, 2, \ldots, n\}, \\
&\ \vdots
\end{aligned}$$

These are the *von Neumann ordinals*; any set Ω as in the Axiom of Infinity, above, must contain each of them. As their names suggest, the sets have a natural correspondence to the nonnegative integers; the operation of addition can be expressed in terms of repeated application of the successor function. Gathering these objects into a legitimate set, $\omega = \{0, 1, 2, \ldots\}$, relies on the axioms as follows. Start with a specific set Ω of the type detailed in the Axiom of Infinity (#6). Use the Power Set axiom to form $\mathcal{P}(\Omega)$. From this set, use the Separation Axiom to define

$$\mathcal{F} = \{I \in \mathcal{P}(\Omega) \,:\, I \text{ has the properties required of } \Omega \text{ in Axiom 6}\}.$$

(Note that $\Omega \in \mathcal{F}$, so $\mathcal{F} \neq \emptyset$.) Then let $\omega = \bigcap \mathcal{F}$.

In our course, we will write $\mathbb{N} = \{1, 2, 3, \ldots\}$ and call this the set of natural numbers. (The Separation Axiom ensures that $\mathbb{N} = \{k \in \omega \,:\, k \neq 0\}$ is a set.) Unfortunately this terminology has never stabilized in the wider world of mathematics—we simply have to commit to the local definition for $\mathbb{N}$.

Standard constructions use the objects and ideas above to build the set of all integers $\mathbb{Z}$, and then to form the set of rational numbers $\mathbb{Q}$. They are nicely laid out in this 22-minute video. Once we have the rational numbers in hand, with all their familiar arithmetical and inequality properties, we will slow down and pay close attention to the construction of $\mathbb{R}$, the set of real numbers.

Functions and Mappings

Definition. Let X and Y be nonempty sets. We write $G\colon X \to Y$ when G is a subset of $X \times Y$ such that every $x \in X$ appears in exactly one pair $(x, y) \in G$. Any such G is called a *function* or *mapping* from X to Y. (The letter "G" reminds us of the word "graph".)

Culture. The terms "function" and "mapping" have the same meaning at the level of axioms, but they have different etymologies that lead to different default usages. "Function" tends to be preferred when the codomain Y is a set of numbers; "mapping" sounds more abstract and is preferred when Y is a set of vectors or a more general algebraic structure; and some stock phrases—like the "implicit function theorem" and "open mapping theorem" in MATH 321—come in packets that defy simple generalizations.

Notation. Given $G: X \to Y$, we overload* the symbol G as follows.

(a) Standard notation for the unique $y \in Y$ corresponding to x is $y = G(x)$;

(b) G induces a natural $\mathcal{G}: \mathcal{P}(X) \to \mathcal{P}(Y)$, defined by

$$\mathcal{G}(A) = \bigcup_{x \in A} \{G(x)\} = \{y \in Y \,:\, y = G(x) \text{ for some } x \in A\}, \qquad \forall A \subseteq X.$$

(c) G induces a natural $\mathcal{G}^{-1}: \mathcal{P}(Y) \to \mathcal{P}(X)$, defined by

$$\mathcal{G}^{-1}(B) = \{x \in X \,:\, G(x) \in B\}, \qquad \forall B \subseteq Y.$$

(Note that $\mathcal{G}^{-1}(B) = \emptyset$ is possible.) The set above is called *the preimage of B under G*.

It's typical to write G instead of $\mathcal{G}$ in cases (b) and (c). For (b) this is straightforward; further discussion for (c) appears below.

Example. Writing just $y = x^2$ triggers healthy reflexes in Calculus veterans. The equation "$y = x^2$" evokes the subset of the Cartesian plane $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ defined by selecting the compatible (x, y) pairs:

$$G = \left\{(x, y) \in \mathbb{R}^2 \,:\, y = x^2\right\} = \left\{(x, x^2) \,:\, x \in \mathbb{R}\right\} = \left\{(t, t^2) \,:\, t \in \mathbb{R}\right\}.$$

Our reflexes give the literal symbols "x" and "y" a load of baggage and expectations that we intend to stop taking for granted: the set G can be defined without ever mentioning either of those special letters. Beginners might infer from a statement like $G(x) = x^2$ that the inputs to G must be named x. That's nonsense, which is unfortunately reinforced in some advanced settings. Look how the standard wave equation $u_{tt} = c^2 u_{xx}$ privileges the special letters t and x when referring to the function u. ////

Definition. (Mapping Properties) Let $G: X \to Y$.

(a) To call G *one-to-one* (or 1-1, or *injective*) means this:

$$\begin{gathered} \forall x_1, x_2 \in X,\ G(x_1) = G(x_2) \implies x_1 = x_2 \\ \text{i.e.,} \quad \forall x_1, x_2 \in X,\ x_1 \neq x_2 \implies G(x_1) \neq G(x_2). \end{gathered}$$

(b) To call G *onto* (or *surjective*) means that $G(X) = Y$, i.e.,

$$\forall y \in Y,\ \exists x \in X \text{ s.t. } y = G(x).$$

(c) To call G a *one-to-one correspondence* (or *bijection*) means that G is **both** 1-1 and onto.

* In computing, function overloading lets one name carry several definitions, with the type of the argument determining which one applies. Here our "overloaded" symbol G returns an *element* of Y when its input is an *element* of X, and a *subset* of Y when its input is a *subset* of X.

Notation. If $G: X \to Y$ is a bijection, the induced map $\mathcal{G}^{-1}: \mathcal{P}(Y) \to \mathcal{P}(X)$ mentioned above carries single-element input sets to single-element outputs:

$$\forall x \in X,\, y \in Y, \qquad y = G(x) \iff \mathcal{G}^{-1}\left(\{y\}\right) = \{x\}\,.$$

This makes it possible to define $G^{-1}: Y \to X$ by declaring $G^{-1}(y) = x$ if and only if $y = G(x)$. This special situation is what we mean when we say, "G is invertible, and G^{-1} is the inverse of G." (Using the same symbol G^{-1} for two different things seldom causes trouble; the distinct terms "preimage" and "inverse" can add clarity if necessary.) We have $G(G^{-1}(y)) = y$ for all $y \in Y$, and $G^{-1}(G(x)) = x$ for all $x \in X$.

Let's switch to a more familiar function name for the next example.

Example. Consider $f: \mathbb{R} \to \mathbb{R}$, defined by $f(x) = \dfrac{x}{\sqrt{1+x^2}}$.

(i) f is not onto.
Clearly $|f(x)| < 1$ for every real x, so any y with $|y| \geq 1$ cannot be expressed as $f(x)$. (Sketch.)

(ii) f is 1-1.
To prove this, assume $f(x) = f(y)$ and deduce $x = y$. Indeed,

$$x\sqrt{1+y^2} = y\sqrt{1+x^2} \implies x^2 + x^2y^2 = y^2 + x^2y^2 \implies y = \pm x.$$

If $y = x$, there is nothing more to do. So suppose $y = -x$: in this case $f(y) = f(-x) = -x/\sqrt{1+(-x)^2} = -f(x)$. Together with $f(x) = f(y)$, this entails $f(x) = -f(x)$, so $f(x) = 0$, and that implies $x = 0$. Thus $y = -x = 0$ also, so indeed $y = x$.

(iii) The mapping f puts $\mathbb{R}$ into one-to-one correspondence with the interval $(-1, 1)$. That is, the function $g: \mathbb{R} \to (-1, 1)$ defined by $g(x) = f(x)$ for all $x \in \mathbb{R}$ is invertible. ////