Proof Workshop

Contents

3 Week 3: Set Theory

3.1 Sets

There are a few reasons for introducing set theory. First, set theory is the most commonly accepted foundation for mathematics, and the language of sets is ubiquitous. We have already used the term in the last few weeks. Secondly, set theory serves as a a rich playing field for proofs which do not rely on much prior knowledge. Third, set theory is intimately related to logic (although we will not be able to delve into their formal similarities, some of the exercises hint at their connection).

Definition 3.1. A **set** is a collection of objects.

We indicate that an object is a set by writing two brackets {} with something in the middle.

Example 3.2.

1. The set consisting of the numbers 1, 2, and 3 is a set, written $\{1, 2, 3\}$.
2. The set of all students in the class is a set.
3. The set of all integers is a set, denoted $\mathbb{Z}$.
4. The set of all real numbers is a set, denoted $\mathbb{R}$.

Definition 3.3. The objects contained in a set are called its **elements**. If an object x is an element of a set S, we write $x \in S$. Conversely, if an object x is **not** an element of a set S, we write $x \notin S$.

Example 3.4. We write $1 \in \{1, 2, 3\}$, $2 \in \{1, 2, 3\}$, and $3 \in \{1, 2, 3\}$.

We say that two sets are **equal** if they have precisely the same elements. In symbols,

$$X = Y \iff \forall x : (x \in X \Leftrightarrow x \in Y).$$

This is why people often say that a set cannot contain multiple elements; the sets $\{1, 2, 3\}$ and $\{1, 2, 2, 3\}$ are equal.

There is one particular set which shows up throughout mathematics and deserves its own name.

Definition 3.5. The **empty set**, written $\varnothing$ is the set with no elements.

Beyond saying what a set contains, we also have a way of comparing what two sets contain.

Definition 3.6. Suppose we have two sets S and T. If all of the elements of S are contained in T, we say S is a **subset** of T, and write $S \subset T$ or $S \subseteq T$. In symbols,

$$S \subseteq T \iff \forall x : ((x \in S) \Rightarrow (x \in T)).$$

If $S \subset T$ and $S \neq T$, we call S a **strict subset** of T, and write $S \subsetneq T$.

As with elements, if S is **not** a subset of T, we write $S \not\subset T$.

Example 3.7.

1. We have $\{1, 2\} \subset \{1, 2, 3\}$.
2. We also have $\{1, 2, 3\} \subset \mathbb{Z}$.
3. For any set S, we have $\varnothing \subset S$ (why?).

Sometimes, we wish to define a subset S of a set T which consists of precisely the elements x which satisfy a property P. We write

$$S = \{x \in T : P(x)\}.$$

This is called **set builder notation**.[1] Because S consists of all the elements of T which satisfy $P(x)$, we can think of S defined in this way as the "truth set" of P in T.

Example 3.8. The set of even integers can be written as

$$\{n \in \mathbb{Z} : n \text{ is even}\}.$$

[1] Sometimes, when we want to consider *all* objects which satisy a property P, we write $\{x : P(x)\}$ for this set. However, for technical reasons, this set may not actually exist! Consequently, this notation should mostly be avoided.

Remark 3.9. Up until now, we have largely been writing quantifiers with no specified set in mind. We will (for the most part) stop doing this now that we have introduced sets, because this is technically bad practice for the same reason that we do not specify *all* objects which satisfy a certain property. Thus, to notate statements like "for all odd integers x," we will say

$$\forall x \in \{n \in \mathbb{Z} : (\exists k \in \mathbb{Z} : n = 2k + 1)\}$$

instead of

$$\forall x : x \in \{n \in \mathbb{Z} : (\exists k \in \mathbb{Z} : n = 2k + 1)\}$$

In the same way that there are the addition and multiplication operations that we do on numbers, there are two key operations we can do on sets.

Definition 3.10. Let S and T be sets.

- The **union** of S and T, written $S \cup T$, is the set of all elements contained in either S or T (in set builder notation, $S \cup T = \{x : (x \in S) \vee (x \in T)\}$).
- The **intersection** of S and T, written $S \cap T$, is the set of all elements contained in both S and T (in set builder notation, $S \cap T = \{x : (x \in S) \wedge (x \in T)\}$).
- The **difference** between two sets S and T, written $S \backslash T$, is the set of all elements in S not contained in T (in set builder notation, $S \backslash T = \{x \in S : x \notin T\}$.

You can think of unions as the "or" of sets, intersections as the "and" of sets, and difference as the "not."

Remark 3.11. Union and intersection are *commutative*, which means that $S \cup T = T \cup S$ and $S \cap T = T \cap S$. Set difference, on the other hand, is not: usually, $S \backslash T \neq T \backslash S$.

Here is an example of a proof using sets.

Proposition 3.12. For any sets A, B, and C, we have

$$A \cap (B \cup C) = (A \cap B) \cup (B \cup C).$$

This is an analogue of the distributive law.

Proof. For notational convenience, write $E = A \cap (B \cup C)$ and $F = (A \cap B) \cup (A \cap C)$. Suppose $x \in E$. Then $x \in A$ and $x \in B \cup C$, which means $x \in B$ or $x \in C$ (and possibly both). Then either $x \in A \cap B$ or $x \in A \cap C$, so $x \in F$. It follows that $E \subset F$.

Now suppose $x \in F$. Then $x \in A \cap B$ or $x \in A \cap C$. Then we must have $x \in A$, and either $x \in B$ or $x \in C$. Thus $x \in E$, so $F \subset E$. Since $E \subset F$ and $F \subset E$, we know that $E = F$. □

The above argument is called a **double containment** argument, and illustrates the general approach of showing that two sets A and B are equal by showing that $A \subset B$ and $B \subset A$. We can use this method, combined with de Morgan's laws, to show all the usual properties of unions and intersections which we would expect (for example, that $S \cup T = T \cup S$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, etc.). Examples are given in the exercises.

3.2 Relations

Given a set T and a propositional statement depending on one variable x, we can form the "truth subset" $S = \{x \in T : P(x)\}$, consisting of all the elements $x \in S$ such that $P(x)$ is true. What about propositional statements consisting of multiple variables? The corresponding set-theoretic notion is known as a **relation.**

Definition 3.13. Let X, Y be two sets. A **relation** from X to Y is a subset $R \subseteq X \times Y$, i.e. some collection of ordered pairs $(x, y) \in X \times Y$. If R is a relation from X to Y, we will write xRy if $(x, y) \in R$.

Example 3.14.

1. Let $X = Y = \{\text{humans on earth}\}$. Then
$$R = \{(x, y) \in X \times Y : x \text{ is related to } y\}$$
is a relation.
2. Let $X = Y$. Then $R = \{(x, y) \in X \times X : x = y\}$ is a relation, consisting of the "diagonal" of the set $X \times X$.
3. Let $X = Y = \mathbb{R}$. Then $R = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x \leqslant y\}$ is a relation.
4. Let $X = \mathbb{N}\backslash\{0\}$ and $Y = \mathbb{N}$. Then
$$R = \{(a, b) \in X \times Y : a \mid b\}$$
is a relation. In other words, we have aRb if and only if $a \mid b$. R contains $(1, n)$ for any natural number n. It also contains the elements $(2, 4)$, $(2, 6)$, $(3, 9)$, but not $(5, 2)$.
5. Let $X = \{1, 2, 3\}$ and $Y = \{\triangle, \uparrow\}$. Then the subset
$$R = \{(1, \triangle), (2, \triangle), (3, \triangle), (2, \uparrow)\}$$
is a relation from X to Y.

There are many types of relations which appear in mathematics. There are three, however, which are most commonly seen: equivalence relations, partial orders, and functions. For the rest of this lecture, we will focus on functions.

3.3 Functions

Functions are commonly defined as follows:

Definition 3.15. A **function** f from a set S to a set T is a rule which assigns every element of S to a single element of T.

In other words, a function takes in an input from a set S and spits out an output in the set T. Practically, this is how one should always think of a function. However, we can give a more rigorous definition of a function using relations.

Definition 3.16. A **function** from a set S to a set T is a relation $f \subseteq S \times T$ which satisfies the following property: for each $x \in S$, there exists a unique $y \in T$ such that xfy. We write $f : S \to T$; if xfy, we write $f(x) = y$.

The relation-based definition above is how we can make precise the idea that "f assigns each element of x some element in Y." In this case, x and y are "related" if y is the element which is assigned to x.

If $f : S \to T$ is a function, we call S the **domain** of f and T the **co-domain** of f.

Remark 3.17. The condition defining a function is also known as the "vertical line test." This is because the graph of some function $f : \mathbb{R} \to \mathbb{R}$ always satisfies the following property: any vertical line passing through the page intersects the graph at precisely one point.

Remark 3.18. If we have a function $f : S \to T$, then f is defined on every element of S, and it can assign any given element of S to no more than one element of T. That is, f assigns each element of S to exactly one element of T.

For example, the function $f : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ sending a real number $x \geqslant 0$ to "the" real number y satisfying $y^2 = x$ is not a function, since for most values of x, there are two such y. The same rule, however, *does* define a function $f : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$, since each nonnegative real number has exactly one nonnegative square root.

More precisely, what we are saying is that the relation $f \subseteq \mathbb{R}_{\geqslant 0} \times \mathbb{R}$ defined by $f = \{(x, y) \in \mathbb{R}_{\geqslant 0} \times \mathbb{R} : y^2 = x\}$ is not a function, since when $x = 1$, we have both $(1, 1) \in f$ and $(1, -1) \in f$, violating the vertical line test. On the other hand, the relation $f \subseteq \mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0}$ defined by $f = \{(x, y) \in \mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0} : y^2 = x\}$ is a function.

Both perspectives of a function are helpful. The first is intuitive and provides a quick way to define and give examples of functions; however, it is technically not rigorous. On the other hand, the second is rigorous and helpful for giving formal proofs about functions, but it is less intuitive. We will switch between them when it is convenient.

Example 3.19.

1. We can define familiar functions $f, g : \mathbb{R} \to \mathbb{R}$ by setting $f(x) = x^2$ or $g(x) = \sin(x)$. In terms of relations, we are defining $f = \{(x, y) \in \mathbb{R} \times \mathbb{R} : y = x^2\}$ and $g = \{(x, y) \in \mathbb{R} \times \mathbb{R} : y = \sin(x)\}$.

2. We could also define a function $h : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ given by $h(x) = x^2$, where $\mathbb{R}_{\geqslant 0}$ denotes the set of real numbers greater than or equal to 0. Although f and h satisfy the same "formula," they are different functions, as they have different domains.

3. Not every function needs to be written as a numerical "formula." For instance, we could define a function $r : \{1, 2, 3\} \to \{\text{cow}, \text{pig}\}$ as follows:

x	$r(x)$
1	cow
2	pig
3	cow

   In other words, r is the relation

$$r = \{(1, \text{cow}), (2, \text{pig}), (3, \text{cow})\}.$$

4. With the same sets as the example above, the relation

$$r = \{(1, \text{cow}), (1, \text{pig}), (2, \text{pig}), (3, \text{cow})\}$$

   is not a function. This is because it fails the vertical line test. Similarly, the relation

$$r = \{(1, \text{cow}), (2, \text{pig})\}$$

   is not a function, since it doesn't assign a value to 3.

5. What is a function $\varnothing \to X$? What about a function $X \to \varnothing$?

Definition 3.20. We say that two functions $f, g : X \to Y$ are **equal** if they define the same subset of $X \times Y$, i.e. they are the same relation.

Proposition 3.21. Two functions $f, g : X \to Y$ are equal if and only if $f(x) = g(x)$ for all $x \in X$.

Proof. We have defined a function as a subset $f \subseteq X \times Y$ satisfying the condition that for all $x \in X$, there is a unique element $y \in Y$ such that $(x, y) \in f$; when $(x, y) \in f$, we write $y = f(x)$. We need to show that two functions $f, g \subseteq X \times Y$ are equal if and only if $f(x) = g(x)$ for all x.

Assume first that $f, g \subseteq X \times Y$ are equal. We need to show that $f(x) = g(x)$ for all $x \in X$. Suppose $x \in X$. By definition, $(x, f(x)) \in f$. Thus, $(x, f(x)) \in g$. By definition, this implies $f(x) = g(x)$.

Conversely, suppose $f(x) = g(x)$ for all $x \in X$. Suppose $(x, y) \in f$. Then by definition, $y = f(x)$. Since $f(x) = g(x)$, this implies $(x, g(x)) \in f$. But by definition, we have $(x, g(x)) \in g$. Hence, $(x, y) \in g$. This implies $f \subseteq g$. Similarly for the other direction, suppose $(x, y) \in g$. Then by definition, $y = g(x)$. Since $f(x) = g(x)$, this implies $(x, f(x)) \in g$. But by definition, $(x, f(x)) \in f$, so this implies $(x, y) \in f$. We have shown that $g \subseteq f$, so we conclude that $f = g$. □

Remark 3.22. We can only say that two functions are equal if they have the same domain and codomain!

Remark 3.23. The proposition above basically says that a function is entirely determined by "the rule which assigns elements," which justifies the first, nonrigorous definition. Thus, we can always specify a function on a set on a set X by specifying a formula $f(x)$ for each $x \in X$, provided that the formula is "well-defined," i.e. there is one and only one way to interpret $f(x)$.

Example 3.24.

1. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x$ and $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = (1 + x^2)/(1 + x^2) - 1 + x$ are equal, because $g(x) = (1 + x^2)/(1 + x^2) - 1 + x = 1 - 1 + x = x = f(x)$ for all $x \in \mathbb{R}$.
2. The function $f : \{1\} \to \mathbb{R}$ defined by $f(x) = x - 1$ and $g : \{1\} \to \mathbb{R}$ defined by $g(x) = \log x$ are equal, because for all $x \in \{1\}$, we have $x = 1$, and hence $f(x) = x - 1 = 0 = \log x = g(x)$.
3. Any two functions $\varnothing \to X$ are equal to each other (why?).

In other words, a function is determined purely by its domain, codomain, and how it assigns elements of its domain to elements of its codomain, *not* how we label it.

Definition 3.25. For any set X, the function $f : X \to X$ defined by setting $f(x) = x$ for every $x \in X$ is called the **identity function** on X, and is written id_X. In terms of relations, id_X is the relation

$$\mathrm{id}_X = \{(x, y) \in X \times X : y = x\}.$$

The case of f and h in Example 3.19 highlights another important attribute of functions.

Definition 3.26. Let $f : S \to T$ and $A \subseteq S$. The **restriction** of f to A, written $f|A$ or $f|_A$, is a function $f|_A : A \to T$ given by the same rule as f, but with domain A. In terms of relations, $f|_A$ is the relation

$$f|_A = f \cap (A \times T).$$

Example 3.27. The functions $f : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ given by $f(x) = h(x) = x^2$ from Example 3.19 are an example of a restriction: in particular, $h = f|_{\mathbb{R}_{\geqslant 0}}$.

Sometimes, when we want a function to be defined by multiple "formulas," we write it **piecewise**; for instance, the function r from Example 3.19 can be rewritten as

$$r(x) = \begin{cases} \text{cow} & x = 1 \vee x = 3 \\ \text{pig} & x = 2 \end{cases}$$

or in terms of relations,

$$r = \{(x, y) : (y = \text{cow} \wedge (x = 1 \vee x = 3)) \vee (y = \text{pig} \wedge x = 2)\}.$$

Definition 3.28. Let $f : X \to Y$ and $g : Y \to Z$ be functions. We define the **composition** of f and g, denoted $g \circ f$, to be the function defined by

$$(g \circ f)(x) = g(f(x))$$

for all $x \in X$. In terms of relations, $g \circ f : X \to Z$ is the relation defined by

$$g \circ f = \{(x, z) \in X \times Z : z = g(f(x))\}.$$

Example 3.29. If $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are defined by

$$f(x) = \begin{cases} x + 1 & x \leqslant -1 \\ \frac{x^2+3}{x+1} & x > -1 \end{cases}, \quad g(x) = 5x + 2,$$

then $g(f(x))$ is given by the formula

$$g(f(x)) = \begin{cases} 5(x + 1) + 2 & x \leqslant -1 \\ 5\left(\frac{x^2+3}{x+1}\right) + 2 & x > -1 \end{cases} = \begin{cases} 5x + 7 & x \leqslant -1 \\ \frac{5x^2+2x+17}{x+1} & x > -1 \end{cases}$$

Definition 3.30. Let $f : S \to T$. We say f is **injective** if whenever $f(x_1) = f(x_2)$, we have $x_1 = x_2$. Equivalently (via contrapositive), f is injective if $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$. This condition is also known as the "horizontal line test."

Definition 3.31. Let $f : S \to T$. We say f is **surjective** if for every $y \in T$, there exists an $x \in S$ such that $f(x) = y$.

Definition 3.32. If $f : S \to T$ is both injective and surjective, we say it is a **bijection** between S and T.

Remark 3.33. An other words, an injective function is one where every element in the codomain is hit at most once (but possibly not at all). A surjective function is one where every element in the codomain is hit at least once (but possibly multiple times). A bijective function is one where *every* element in the codomain is hit once, and exactly once.

Remark 3.34. One way to think about bijections is that a bijection describes when two sets S and T are "the same." A bijection $f : S \to T$ is essentially a rule which relabels the elements of S as elements of T, and the labeling scheme is a perfect one-to-one correspondence.

Example 3.35. The function $h : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ given in Example 3.19 as $h(x) = x^2$ is an example of an injection but not a surjection, since negative numbers don't have real square roots. Note that the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is **not** an injection (nor a surjection), since $-x$ and x map to the same element x^2. However, the function $h : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ given by $h(x) = x^2$ is both injective and surjective, hence bijective. This exhibits the need to consider the domain and codomain as part of the data of a function.

Example 3.36. An example of a surjective function which is not injective is $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3 - x$.

Example 3.37. The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$ is a bijection.

Example 3.38. For any set X, id_X is a bijection.

We will give an equivalent formulation of what it means for a function to be bijective.

Definition 3.39. We say a function $f : X \to Y$ has an **inverse** if there exists some function $g : Y \to X$ such that $g(f(x)) = x$ and $f(g(y)) = y$ for any $x \in X$ and $y \in Y$. In other words, $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. We write $g = f^{-1}$ if this is the case.

Proposition 3.40. $f : X \to Y$ is a bijection if and only if it has an inverse.

Proof. Suppose first that f has an inverse $g : Y \to X$. We wish to show that f is bijective. To show that f is injective, suppose $x, y \in X$ are such that $f(x) = f(y)$. Then $g(f(x)) = g(f(y))$. Since $g(f(x)) = x$ for all $x \in X$, we have $x = y$. Therefore, f is injective. To show that f is surjective, suppose $y \in Y$. We need to find some $x \in X$ such that $f(x) = y$. Take $x = g(y)$. Then $f(x) = f(g(y)) = y$, as desired.

Conversely, suppose f is bijective. We need to construct a function $g : Y \to X$ such that $f(g(y)) = y$ and $g(f(x)) = x$ for all $x \in X$ and $y \in Y$. Let $g : Y \to X$ be the relation defined by

$$g = \{(y, x) \in Y \times X : (x, y) \in f\}.$$

We need to show that (i) g satsifies the vertical line test, and (ii) g is an inverse of f. First, suppose $y \in Y$. We need to show that there is a unique $x \in X$ such that $(y, x) \in g$. Since f is surjective, there is an element $x \in X$ such that $f(x) = y$. Then $(x, y) = (x, f(x)) \in f$, so that $(y, x) \in g$. If x' is another element such that $(y, x') \in g$, then by definition, (x, y) and (x', y) are both in f. This means that $f(x) = y = f(x')$. Since f is injective, this means that $x = x'$, so that the element is unique.

Now we show that g is an inverse. For all $x \in X$, we have by definition that $(x, f(x)) \in f$, so that $(f(x), x) \in g$. This means that $g(f(x)) = x$. On the other hand, we have $(y, g(y)) \in g$ for all $y \in Y$ by definition, so that $(g(y), y) \in f$. This means that $f(g(y)) = y$ for all $y \in Y$. □

More interesting examples of functions and bijections are given in the exercises.

Remark 3.41. In order to show that a function $f : S \to T$ is a bijection, one can either show that f is both injective and surjective, or one can prove that f has an inverse by constructing a function $g : T \to S$ and verifying that $g(f(x)) = x$ and $f(g(y)) = y$ for all $x \in S$ and $y \in T$.

As is the case with logic, unions and intersections are not the only way to construct new sets from old sets. We can, for example, look at *sets of sets* or *sets of functions.* Given a function $f : X \to Y$, we can also look at *images and preimages* of subsets. Some of these constructions will be detailed in the exercises.

Definition 3.42. Let X, Y be sets. We define the **powerset of** X to be the set of all subsets of X, denoted $\mathcal{P}X$. We also denote the **set of all functions** $X \to Y$ by $\mathrm{Hom}(X, Y)$ or Y^X.

You will have the opportunity to explore the properties of these constructions in the exercises.