MATH 113: DISCRETE STRUCTURES
FUNCTIONS

1. FUNCTIONS AS ASSIGNMENTS

In an earlier handout, we defined a function $f : A \to B$ (with domain the set A and codomain the set B) to be a subset $f \subseteq A \times B$ such that for every $a \in A$ there is a unique pair $(a, b) \in f$. This is the *graph* interpretation of functions: think of A as the "horizontal axis" and B as the "vertical axis." (If A and B are (subsets of) $\mathbb{R}$, you can literally do this!) The function condition is then the "vertical line test" — each "vertical line" through some $a \in A$ hits exactly one graphed point (a, b).

It is typical to think of functions as *assignments* rather than as particular subsets of a Cartesian product. When $(a, b) \in f : A \to B$, we say that $b = f(a)$ and think of f "sending" a to b. The function condition then says that each $a \in A$ gets sent to precisely one $b \in B$.[1]

Example 1.1. Consider the set $f = \{(1, 3), (2, 3), (3, 4)\} \subseteq \{1, 2, 3\} \times \{1, 2, 3, 4\}$. This is a function for which $f(1) = 3$, $f(2) = 3$, and $f(3) = 4$.

Notation. We will sometimes write $f : a \mapsto b$ when $f(a) = b$ and read this statement as "f maps a to b." It is important that "$\mapsto$" is not the same as "$\to$": $f : A \to B$ tells us that f is a function with domain A and codomain B, while $f : a \mapsto b$ says that $f(a) = b$. For the function from Example 1.1, we could write $f : 1 \mapsto 3,\ 2 \mapsto 3,\ 3 \mapsto 4$.

Example 1.2. In calculus, you may have considered a function $\mathbb{R} \to \mathbb{R}$ given by a formula such as $f(x) = x^3 + \sin x$. This is still a perfectly reasonably function because each $x \in \mathbb{R}$ is sent to one $f(x) \in \mathbb{R}$ (namely, $x^3 + \sin x$). As a graph, this function is $\{(x, x^3 + \sin x) \ : \ x \in \mathbb{R}\}$.

Example 1.3. Not all functions have reasonable formulas. For instance, there is a function $g : \mathbb{R} \to \mathbb{R}$ which takes x to x if the first nonzero digit of x is 1 and otherwise takes x to 0. Weird, but still a function.[2]

Example 1.4. Here's an interesting way to use a function: Given a set X and subset $A \subseteq X$, let's build a function which specifies the points of A. We define the *indicator function* for A to be $\chi_A : X \to \{0, 1\}$ given by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

A couple of comments: first, χ is the Greek letter "chi." Second, the formula above is an example of a *piecewise definition*: we partition the domain into disjoint subsets whose union is all of X (in this case, A and $X \smallsetminus A$), and then give a formula or rule describing what the function does to elements in each subset.

Note that we can reconstruct A from χ_A as all $x \in X$ such that $\chi_A(x) = 1$, *i.e.*,

$$A = \{x \in X \ : \ \chi_A(x) = 1\}.$$

You've actually seen this trick before! We were secretly using indicator functions to enumerate subsets, producing a one-to-one correspondence between subsets of X and functions $X \to \{0, 1\}$.

[1] Note that for a given $b \in B$, more than one a can go to b. The point here is that (1) $f(a)$ takes some value in B, and (2) it only takes one, instead of multiple, values in B.

[2] Worse yet, "most" functions between infinite sets are not describable by any written rule whatsoever, but we will not pursue this perversity further.

2. Composition

Let's now explore how functions interact with each other via composition.

Definition 2.1. Suppose $f : A \to B$ and $g : B \to C$ are functions and the codomain of f equals the domain of g. Then we define the *composite* of g with f to be the function $g \circ f : A \to C$ by the equation $(g \circ f)(a) = g(f(a))$.

The composite $g \circ f$ "does f first" and then "does g." We can express this graphically with a picture called a commutative diagram:

$$\begin{tikzcd} & B \arrow[dr, "g"] & \\ A \arrow[ur, "f"] \arrow[rr, "g\circ f"'] & & C. \end{tikzcd}$$

Here the arrows go from domain to codomain and are labelled by the corresponding function. If we start with $a \in A$, then the arrow labelled f takes a to $f(a)$. Continuing this path, the arrow labelled g takes $f(a)$ to $g(f(a))$. Meanwhile, the arrow labelled $g \circ f$ takes a to $g(f(a))$ by definition. Since both paths do the same thing to every $a \in A$, we say that it "commutes."

The exact shape of a commutative diagram doesn't matter. If someone told us that the diagram

$$\begin{tikzcd} Y \arrow[r, "K"] & Z \\ X \arrow[u, "J"] \arrow[ur, "L"'] & \end{tikzcd}$$

commutes, we would know that $K(J(x)) = L(x)$ for each $x \in X$; in other words, $L = K \circ J$ when that diagram commutes.

We can compose more than two functions as well, as long as domains and codomains match up properly. For instance, $h \circ g \circ f : A \to D$ makes sense as long as $f : A \to B$, $g : B \to C$, and $h : C \to D$ for some sets A, B, C, and D; we have $(h \circ g \circ f)(a) = h(g(f(a)))$. We leave it as an exercise to the reader to (a) check that $h \circ g \circ f = h \circ (g \circ f) = (h \circ g) \circ f$, and (b) draw a commutative diagram describing this triple composite. Property (a) has a name: composition is *associative*.

Every set A supports a special function $\mathrm{id}_A : A \to A$, called the *identity function* on A, which interacts in a special way with composition. This function simply takes a to a for each $a \in A$, *i.e.*, $\mathrm{id}_A : a \mapsto a$ or $\mathrm{id}_A(a) = a$. If $f : A \to B$ is a function, let's consider the composite $f \circ \mathrm{id}_A$. Well, $(f \circ \mathrm{id}_A)(a) = f(\mathrm{id}_A(a)) = f(a)$ for every $a \in A$, so $f \circ \mathrm{id}_A = f$. Similarly, $\mathrm{id}_B \circ f = f$. (Note that we had to change id_A to id_B so that domains and codomains would match up!) We see then that composition with the identity function *does nothing* to the other function. This distinguishes identity functions amongst all functions with the same domain and codomain.

3. Special types of functions

We now explore functions with special properties, namely *injections*, *surjections*, and *bijections*.

3.1. Injections.

An injection is a function which does not hit the same value twice. We formalize this idea in the following definition.

Definition 3.1. A function $f : A \to B$ is *injective* (or is an *injection*) if $f(x) = f(y)$ (for $x, y \in A$) if and only if $x = y$.

Meditate on this definition for a while if it seems funny. The point is that f does not duplicate values in the codomain, so an equality between values ($f(x) = f(y)$) is only possible when $x = y$.

Let's briefly return to our graph interpretation of functions. An injection hits each value in the codomain at most once. This is also referred to as the *horizontal line test*: when we draw a horizontal line through any $b \in B$, we hit at most one point of the form (a, b) in the graph.

You may have learned in middle school that functions passing the horizontal line test have inverses. This fact remains true in the current context, although we must be careful with the domain of our inverse function, requiring the following definition.

Definition 3.2. The *image* of a function $f : A \to B$ is the set

$$\mathrm{im}(f) = \{b \in B \ : \text{ there exists } a \in A \text{ such that } f(a) = b\}.$$

In other words, the image of f consists of all the elements of B that are "hit" by the function. For instance, the image of the function $f : \{1, 2, 3\} \to \{1, 2, 3, 4\}$ from Example 1.1 is $\{3, 4\}$. The image of the function from Example 1.3 is

$$\{x \in \mathbb{R} \ : \text{ the first nonzero digit of } x \text{ is } 1\} \cup \{0\}.$$

When a function $f : A \to B$ is injective, it has an *inverse* function $f^{-1} : \mathrm{im}(f) \to A$; this is the unique function satisfying the equalities $f(f^{-1}(b)) = b$ for each $b \in \mathrm{im}(f)$ and $f^{-1}(f(a)) = a$ for each $a \in A$. It is tempting then to write that $f \circ f^{-1} = \mathrm{id}_{\mathrm{im}(f)}$ and $f^{-1} \circ f = \mathrm{id}_A$, but we should recognize that there is a slight mismatch between domains and codomains. If we replace $f : A \to B$ with $\tilde{f} : A \to \mathrm{im}(f)$ taking the same values ($\tilde{f}(a) = f(a)$ for all $a \in A$), then its completely legitimate to write $\tilde{f} \circ f^{-1} = \mathrm{id}_{\mathrm{im}(f)}$ and $f^{-1} \circ \tilde{f} = \mathrm{id}_A$.

3.2. **Surjections.** Given the terminology we've already introduced, surjections are easy to define.

Definition 3.3. A function $f : A \to B$ is *surjective* (or is a *surjection*) if $\mathrm{im}(f) = B$.

In other words, surjections hit everything in their codomain. Of course, when we define a function, we have some choice regarding the codomain. For instance, we could consider the assignment on real numbers $x \mapsto x^2$ to have codomain $\mathbb{R}$ or codomain $[0, \infty) = \{x \in \mathbb{R} \ : \ x \geq 0\}$. In the first instance, the function is not surjective, but in the latter case it is (because every nonnegative real number has a square root [in fact, two square roots]).

Example 3.4. Suppose $A \subsetneq X$ is a nonempty proper subset of X. Then the indicator function $\chi_A : X \to \{0, 1\}$ is surjective. (Why? What if $A = \varnothing$ or X?)

3.3. **Bijections.** Finally, we come to bijections, also called one-to-one correspondences.

Definition 3.5. A function is *bijective* (or is a *bijection*) if it is both injective and surjective.

Suppose $f : A \to B$ is bijective. Then it is injective with $\mathrm{im}(f) = B$, so it has an inverse function of the form $f^{-1} : B \to A$ satisfying $f \circ f^{-1} = \mathrm{id}_B$ and $f^{-1} \circ f = \mathrm{id}_A$. (We don't need to replace f with $\tilde{f}$ because $\mathrm{im}(f)$ is all of B.) In fact, a function has such an inverse if and only if it is bijective.

Theorem 3.6. *A function $f : A \to B$ is bijective if and only if there exists a function $g : B \to A$ (called a [two-sided] inverse of f) such that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$.*

Proof. We have already seen that if f is bijective, then such a g exists. Suppose now that $f : A \to B$ is a function and there exists $g : B \to A$ such that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$. We need aim to show that f is bijective, and will first show that it is injective. Suppose that there are $x, y \in A$ such that $f(x) = f(y)$. Applying g to this equality, we get $g(f(x)) = g(f(y))$, and since $g \circ f = \mathrm{id}_A$, this becomes $x = y$. Hence f is injective.

We now show that f is surjective. Given $b \in B$, let $a = g(b)$. Then $f(a) = f(g(b)) = b$, so f is surjective. Since f is injective and surjective, it is in fact a bijection, as desired. $\square$

Bijections are incredibly useful in combinatorics. Every combinatorial problem can be reframed as trying to determine the cardinality of a set. The following theorem tells us that bijections preserve cardinality, so a good way to "count" is to produce a bijection between the set we would like to count, and a set with a known number of elements.

Theorem 3.7. *There exists a bijection $f : A \to B$ if and only if $|A| = |B|$.*

Proof. Just kidding! This is actually the *definition* of cardinality. A *cardinal number* is actually an equivalence class of sets up to bijection.

Nonetheless, I will still try to explain why this makes sense in the case where both sets are finite. Suppose that $|A| = n = |B|$. By counting the n elements of A and B, we produce bijections $a : \{1, 2, \ldots, n\} \to A$ and $b : \{1, 2, \ldots, n\} \to B$. You should check that $f = b \circ a^{-1}$ is a bijection $A \to B$.

Now suppose that A is finite of cardinality n and there exists a bijection $f : A \to B$. Counting A again produces a bijection $a : \{1, 2, \ldots, n\} \to A$. Convince yourself that $f \circ a : \{1, 2, \ldots, n\} \to B$ counts B, so $|B| = n$ as well. $\square$