Discrete Mathematics

Lecture Notes

on

Sets and Functions

M. Andrew Moshier

October 2014

Overview

The mathematical universe consists of various things: numbers, functions, graphs, lists and so on. A *set* is a collection of things. For example, the collection of all natural numbers is a set. A *function* is a correlation of the members of one set with members of another set. These two abstract concepts (sets and functions) form a conceptual framework in which virtually all of mathematics can be built. So an understanding of sets and functions is key to a rigorous approach to most other parts of mathematics. This conceptual framework can itself be put on a formal, precise footing called the Category of Sets and Functions.

In these lectures, we build up the Category of Sets and Functions, so that we can use these things as the basic building blocks of everything else we do.

Version: 0:0.0 (November 8, 2014) Format: print

Contents

Contents **2**

1 Sets **3**
1.1 Set Basics 4
1.2 Subsets and Extensionality 6

2 Functions **9**
2.1 Internal Diagrams 10
2.2 Extensionality 12

3 Building Sets and Functions **15**
3.1 Elements, Pointers and Constant Functions 16
3.2 Solution Sets, Subsets, Characteristic Functions 17
3.3 Product Sets and Functions of Two Variables 18
3.4 Function Sets, Partial Evaluation 19
3.5 Powersets 21

4 The Set of Natural Numbers **23**
4.1 Sequences and Successions 24
4.2 Primitive Recursion 24

5 Monomophisms, Epimorphisms and Isomorphisms **27**

1

Sets

Goals

Lecture

- Describe informally the category of sets.
- Define list set notation.
- Introduce the idea of a subset.
- Introduce the axiom of extensionality for sets and some of its consequences.

Study

- Demonstrate ability to determine equality of sets.
- Develop facility in basic set theoretic notation.

Sets are the mathematician's way of thinking about *collections* of objects. Examples will be the set of natural numbers, the set of pairs of natural numbers, the set of lists of natural numbers, and so on. *Functions* are the mathematicians way of thinking about operations, such as successor, addition, summation, and so on.

Mathematicians also use functions to model attributes of the things in a collection, like "the color of", "the mass of", "the location of", "the father of", "the favorite book of the person to the left of" and so on. There is little sense in saying these are "operations", but they have a similar behavior. For each potential object $\mathfrak{x}$ of the right sort (a wooden building block, for example), *the color of* $\mathfrak{x}$ is a specific color. Similarly, for a natural number $\mathfrak{n}$, "the successor of $\mathfrak{n}$" is another specific natural number. So "the color of" is a function from the set of wooden building blocks to the set of colors; "the successor of" is a function from the set of natural numbers to set of natural numbers. Likewise, addition constitutes a function from the set of pairs of natural numbers to the set of natural numbers. In English we might say "the sum of $\mathfrak{m}$ and $\mathfrak{n}$", but we usually write $\mathfrak{m}+\mathfrak{n}$. Either way, any pair of natural numbers has a sum. So "the sum of" is an attribute of pairs of natural numbers, just like "the color of" is an attribute of wooden building blocks.

Taken together, sets and functions constitute a fundamental structure in contemporary mathematics called the *Category of Sets and Functions*. This is a slight lie. Actually, there are many different categories of sets that differ in subtle ways. But for most mathematics, the differences are irrelevant. So in practice, it is safe to talk as if there is just one category of sets. The Category of Sets and Functions sometimes abbreviated as **Set**.

To understand sets and functions as they are used in every day mathematics, we need to answer some questions:

- What do we mean by saying that a set is a collection?
- What do we mean by saying that two sets are equal?
- What do we mean by saying that a function behaves like an attribute?
- What do we mean by saying that two functions are equal?

The answers to these leads to some basic principles for reasoning about sets and functions. Other principles allow us to construct sets and functions with specific behaviors. We could be more formal and present these principles as *axioms*, but the word "axiom" has a special connotation in mathematics that we do not need here. Nevertheless, everything we say in these lectures could be presented in terms of formal axioms.

1.1 Set Basics

A set consists of things that are "in" the set. All other things are "not in" the set. For example, later we will see that the natural numbers constitute a set $\mathbb{N}$. So 0 is in $\mathbb{N}$, 1 is in $\mathbb{N}$, and so on, but $\frac{1}{2}$ is not in $\mathbb{N}$. We make this precise and introduce notation for the idea.

Basic Vocabulary 1.1

A *set* is a mathematical entity A with the following feature. For any thing x, either x *is in* A or x *is not in* A. We write $x \in A$ if x is in A and $x \notin A$ if x is not in A.

The symbol $\in$ is used in mathematics exclusively to indicate membership in a set. You will not see it used in any other way.

For variety, all of the following phrases mean the same thing:

- x is in A
- x is an *element of* A
- x is a *member of* A
- A *contains* x
- x *belongs to* A

Basic Vocabulary 1.1 describes how we can talk about sets and elements, and how to use the notation of membership, but does not tell us that any sets actually exist. We will remedy that in the next sections. Our first remedy is to make room for finite sets.

Principle 1.1

[Finite Sets] For any list $L = [a_0, \ldots, a_{n-1}]$, there is a set, denoted by $\{a_0, \ldots, a_{n-1}\}$, so that $x \in \{a_0, \ldots, a_{n-1}\}$ if and only if $x = a_i$ for some $i < n$. More precisely,

- $x \notin \{\}$ for any x (so $\{\}$ is said to be *empty*);
- $x \in \{a_0, \ldots, a_n\}$ if and only if $x = a_0$ or $x \in \{a_1, \ldots, a_n\}$.

Example 1.1

Here are some examples of sets built from finite lists:

- $\{\}$ – an empty set;
- $\{1, 2, 5\}$ – a set consisting of three elements;
- $\{\{\}\}$ – a set consisting of one element, which is $\{\}$;
- $\{1, 2, 4, \{1, 2\}\}$ – a set consisting of four elements, 1, 2, 4 and the set $\{1, 2\}$.
- $\{4, 5, \{\}, [\,]\}$ – a set consisting of four elements. Note that the set $\{\}$ and the list $[\,]$ are not the same things.
- $\{1, 2, 3, 4, 3, 2, 1\}$ – a set consisting of four elements, listing an element twice is redundant.

The study of finite sets is surprisingly complex, and comprises a large part of the branch of mathematics called *combinatorics*. We will touch on some basics of combinatorics later in the course.

Various infinite sets of numbers also exist. All of these are indeed sets (that is, their existence follows from general principles of set theory), but we will not try to justify that explicitly, except for the set of natural numbers.

Definition 1.1

The following sets are denoted by special symbols:

$$\begin{aligned}
\mathbb{N} &= \text{the set of natural numbers} \\
\mathbb{Z} &= \text{the set of integers} \\
\mathbb{Q} &= \text{the set of rational numbers} \\
\mathbb{R} &= \text{the set of real numbers} \\
\mathbb{C} &= \text{the set of complex numbers}
\end{aligned}$$

Exercises 1.1

1. Let $A = \{1, \{2, 3\}, 4\}$. Determine which of the following assertions are true.
   a) $1 \in A$
   b) $2 \in A$
   c) $\{\} \in A$
   d) $\{2, 3\} \in A$
   e) $A \in A$
2. In the following examples of sets with elements following a pattern, write an expression for the same set that makes the pattern clearer.
   a) $\{0, 2, 4, \ldots, 100\}$
   b) $\{1, 2, 4, 8, \ldots, 256\}$
   c) $\{0, 1, 3, 6, 10, \ldots, 55\}$

1.2 Subsets and Extensionality

Sets are meant to be bare collections. For a set A, some things are in A, some are not. And that's all we can say. Unlike a list, a set has no "initial" element. For example, the set $\{1, 2, 3\}$ should be the same as the set $\{2, 3, 1\}$, because both have the same elements. This is an important difference between lists and sets: $[1, 2, 3]$ and $[2, 3, 1]$ are *not* the same lists because order matters in lists. To make this precise, we need to be clear about when sets are equal. To do this, we introduce an important definition.

Definition 1.2

For sets A and B, we say that A *is a subset of* B provided that every element of A is an element of B. We also write this as $A \subseteq B$, and say that A *is included in* B. We may also write $B \supseteq A$ to mean the same thing, and say that B *is a superset of* A.

If A is *not* a subset of B, we write $A \not\subseteq B$. If $A \subseteq B$ and $B \not\subseteq A$, then A is called a *proper subset of* B. To indicate that A is a *proper* subset of B, we may write $A \subsetneq B$.

Saying $A \subseteq B$ is exactly the same as saying that for any x, if $x \in A$ then $x \in B$.

Example 1.2

Here are some examples and counter-examples of the subset relation.

- $\{1, 2, 3\} \subseteq \{0, 1, 2, 3\}$
- $\{\} \subseteq \{0\}$
- $A \subseteq A$ for any set A because, trivially, every element of A is an element of A
- $\{\} \subseteq A$ for any set A because every element of $\{\}$ (there are none) is an element of A
- $\{1, 2, 3\} \not\subseteq \{0, 2, 3\}$ because $1 \in \{1, 2, 3\}$ but $1 \notin \{0, 2, 3\}$
- $\{1, 2, 3\} \subseteq \{2, 3, 1\}$

Exercises 1.2

For each of the following pairs of sets, determine whether or not the first is a subset of the second. Explain your answer.

1. $\{0, 1\}$ and $\{1, 0\}$
2. $\{a, b, c, d\}$ and $\{a, b, d, e, c\}$
3. $\{\}$ and $\{\{\}\}$
4. $\{0, 3, 6, 10\}$ and $\{10, 9, 8, 7, 6, 5, 4, 2, 1, 0\}$

We can summarize two useful properties of $\subseteq$ as follows.

- [Reflexivity] For any set A, $A \subseteq A$. We say $\subseteq$ is *reflexive*.
- [Transitivity] For any sets A, B and C, if $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$. We say $\subseteq$ is *transitive*.

Another familiar example of a reflexive, transitive relation is $\leq$ on the natural numbers. In fact there are many examples of reflexive transitive relations throughout mathematics.

The relation $\leq$ is also *anti-symmetric*, meaning that if $m \leq n$ and $n \leq m$ then $m = n$. Suppose $A \subseteq B$ and $B \subseteq A$. Then, by definition A and B have the exact same elements. By our understanding of sets as collections, A and B must be equal. So we state this as another axiom.

Principle 1.2

[**The Axiom of Set Extensionality**] For sets A and B, if $A \subseteq B$ and $B \subseteq A$, then $A = B$. In other words, $\subseteq$ is anti-symmetric.

Based on this, we can already establish a useful fact: there is exactly one empty set. To set the tone for what follows, we make this a formal claim.

Lemma 1.1

There is exactly one empty set.

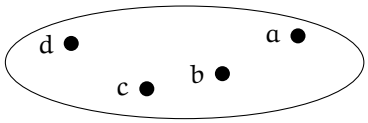
Proof: We have already noted that the set built from an empty list $\{\}$ has no elements. So there is at least one empty set.

Suppose E is a set with no elements. Then $E \subseteq \{\}$ because every element of E (there are none) is an element of $\{\}$. Similarly, $\{\} \subseteq E$ because every element of $\{\}$ (again, there are none) is an element of E. So by Principle 1.2 $E = \{\}$. $\square$

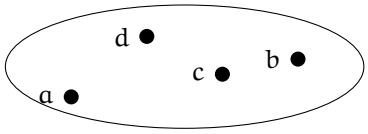
Definition 1.3

The set $\{\}$ is also denoted by $\emptyset$.

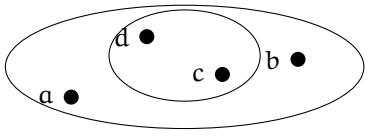
Set extensionality makes precise the idea that a set by itself does not have any structure other than what members it possesses. To emphasize this, sometimes it is useful to depict a set with elements scattered about something like



with the elements scattered about. Evidently, a re-arrangement of the elements does not change the depicted set. So



is the same set. Depicting a subset of a set is a simple matter of drawing a smaller boundary around some of the elements as in the following.



Exercises 1.2

Draw depictions of the following sets

5. $\{1, 4, 5, 2, 3\}$
6. $\{1, 2, 3, \ldots, 23\}$
7. $\{a, b, c, d, e\}$ and $\{c, e, f, g\}$ on the same diagram
8. $\{a, e, b, c, e\}$ [sic]
9. $\{1, 3, 6, 7\}$ and $\{1, 3, 5, 6, 7, 9\}$ on the same diagram
10. $\{\bot, \top$
11. $\{\bullet\}$
12. $\{\top, \bot, 3, 5, 1, \bullet\}$

2

Functions

Goals

Lecture

- Introduce basic structure of functions
- Define the identity functions and function composition
- Introduce internal diagrams of functions.

Study

- Be able to determine equality of functions
- Use internal diagrams to depict function composition

Functions (perhaps in your calculus courses) are often talked about as *operations*. For example,

$$f(x) = x^2 - 1$$

can be seen as an operation that transforms a number x into its square. But it can also be seen as an attribute (the "square of x"). The "operational" view is informal, and often useful. As we will see, though, it gets an important aspect of functions wrong because two entirely different operations may define the same function.

Informally, a function "takes" an argument from a given set as input and "produces" an output in a given set. So the function f defined by $f(x) = x^2 + 2x + 1$ might "take" the natural number 2 and "produce" the natural number 10. That is, $f(3) = 3^2 + 1 = 10$. We begin by introducing the vocabulary of functions.

Basic Vocabulary 2.1

- For a set A and a set B, there are things called *functions from* A *to* B, with a function f from A to B being writing $f\colon A \to B$ or $A \xrightarrow{f} B$.
- For $f\colon A \to B$, the set A is called the *domain* of f and B is called the *codomain* of f.
- For any function $f\colon A \to B$ and every element $a \in A$, f and a determine an element of B, written $f(a)$, and read "f of a".

A functions may sometimes also be called a *map*, a *transformation*, or an *operation*. As we will see, however, *operation* is somewhat misleading.

Often, a function $f\colon A \to B$ is *defined* by a rule, just as they are familiar in other parts of mathematics. We typically, write such rules by giving the function a name (very often f) and spelling out the rule at the same time. So we write things like

$$f(x) = x^2 + 4x + 2$$

to define a function $f\colon \mathbb{R} \to \mathbb{R}$ (recall that $\mathbb{R}$ is the set of real numbers). But sometimes it is useful to have a rule without giving it a name. To do that, we use the "maps to" arrow $\mapsto$. So we might define the same function f by saying that f *is given by the rule* $x \mapsto x^2 + 4x + 2$. We will not go so far as to write $f = (x \mapsto x^2 + 4x + 2)$ because this more easily understood by writing $f(x) = x^2 + 4x + 2$. The rule $x \mapsto x^2 4x + 2$ is the same as the rule $y \mapsto y^2 + 4y + 2$. The variable only serves as a place holder, so its particular name does not matter.

There are two fundamental (trivial) types of rules that can be used to build functions.

Axiom 2.1

- For any set A, there is a function $\mathsf{id}_A\colon A \to A$ defined by the rule $x \mapsto x$. This is called the *identity* function on A
- For any two functions $f\colon A \to B$ and $g\colon B \to C$, there is a function $g \circ f\colon A \to C$ defined by the rule $x \mapsto g(f(x))$. This is called the composition of g and f (or sometimes "g following f").

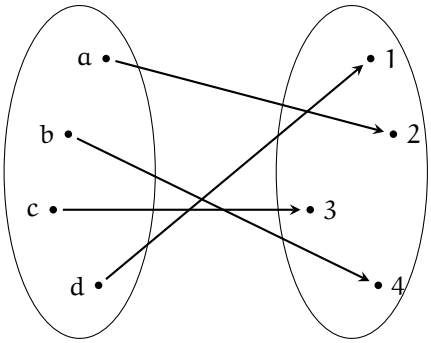
Notice that $g \circ f$ is only defined when the *domain* of g matches the codomain of f. Also, be careful about definitions like $f(x) = 1/x$. This does not define a function on the real numbers because $f(0)$ is undefined. In other words, to be a function, f must determine an element of the codomain for each element of the domain.

Exercises 2.0

Suppose $f\colon A \to B$, $g\colon B \to C$ and $h\colon C \to D$ are functions. Then $h \circ (g \circ f)$ and $(h \circ g) \circ f$ are functions from A to D. Do you think they are equal? Explain your answer in a few clearly written sentences.

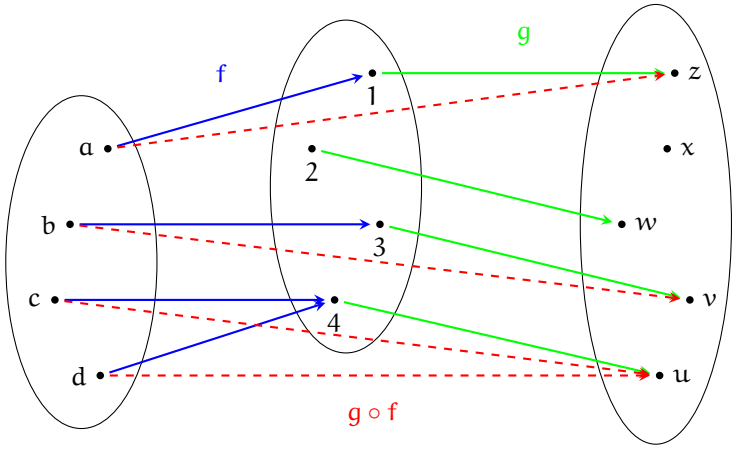
2.1 Internal Diagrams

To depict a function on small sets, we can use the simple internal diagrams of the last section. For example,



depicts a function from the set $\{a, b, c, d\}$ to the set $\{1, 2, 3, 4\}$.

Composition can be illustrated using internal diagrams. For example,

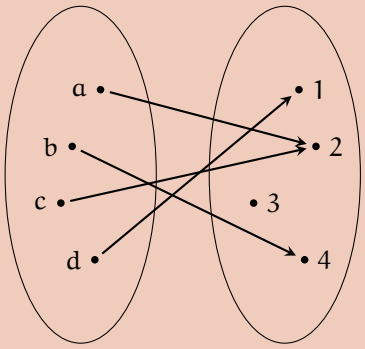


Exercises 2.1

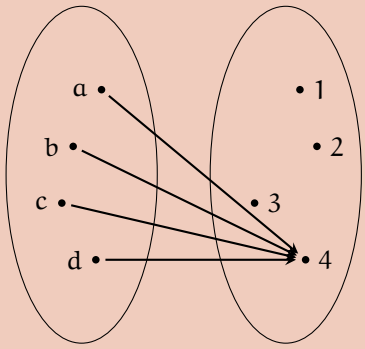
Use internal diagrams for the following exercises.

1. Depict four different functions from the set $\{1, 2, 3\}$ to the set $\{\bot, \top\}$. [Draw four different diagrams.]
2. Depict all of the functions from $\{\bullet\}$ to $\{a, b, c\}$
3. Depict all of the functions from $\{a, b, c\}$ to $\{\bullet\}$
4. Are there any functions from $\{a, b\}$ to $\emptyset$?
5. Are there any functions from $\emptyset$ to $\{a, b\}$? If there are, how many?
6. For each of the following diagrams, determine whether or not the diagram depicts a function. If not, explain why not.

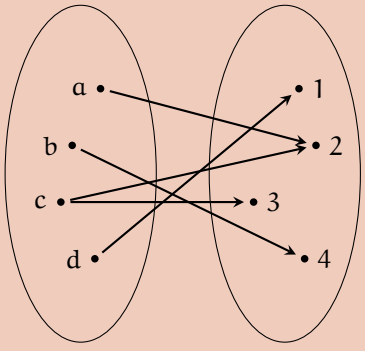
a)



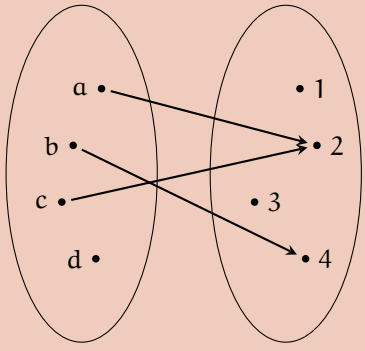
c)



b)



d)



7. Let $A = \{1, 2, 3\}$. Let $B = \{a, b, c, e\}$ and let $C = \{\bot, \top\}$. Depict some functions $f\colon A \to B$, $g\colon B \to C$, and $g \circ f$.
8. Think about how you might depict a function $h\colon A \to A$ using only one picture of the set A. Describe what you would do, and provide an example.
9. Suppose $f\colon \mathbb{R} \to \mathbb{R}$ is given by the rule $x \mapsto x^2$, suppose $g\colon \mathbb{R} \to \mathbb{R}$ is given by the rule $x \mapsto x - 1$. Write rules for $f \circ g$ and $g \circ f$ without using the symbols f and g. Explain whether or not it is the case that $f \circ g = g \circ f$.

2.2 Extensionality

Just like sets, we need a way to say when two functions are equal. Let $\mathbb{N}$ denote the set of natural numbers. Then define $f\colon \mathbb{N} \to \mathbb{N}$ by

$$f(n) = n^2 + 2n + 1.$$

Likewise, define $g\colon \mathbb{N} \to \mathbb{N}$ by

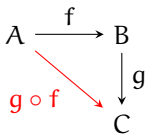
$$g(n) = (n + 1)^2.$$

Evidently, for each $n \in \mathbb{N}$, it is true that $f(n) = g(n)$. So even though f and g are defined by different *operations*, the two functions yield the same results. As with sets, this leads to an axiom for equality of functions.

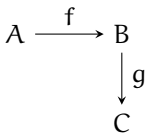
Axiom 2.2

[**The Axiom of Function Extensionality**] For functions $f\colon A \to B$ and $g\colon A \to B$, if it is the case that $f(x) = g(x)$ for all $x \in A$, then $f = g$. Note that equality of functions only makes sense when the two functions share the same domain and the same codomain.

If we not concerned about the detailed internals sets, but only with how functions interact, then function can be depicted very simply as $A \xrightarrow{f} B$. So a composition of functions can be depicted as in

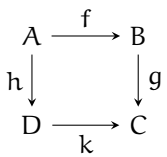


We do not really need to draw $g \circ f$ as a separate arrow because the *path* from A to B to C is already implicitly a depiction of $g \circ f$. So the simpler diagram



shows the same information, namely, that $f\colon A \to B$ and $g\colon B \to C$ are functions and therefore, $g \circ f$ is too.

Now a diagram such as this



depicts two composite functions $g \circ f$ and $k \circ h$, but $g \circ f$ and $k \circ k$ may not be equal. We say that the diagram *commutes* or that it is a *commutative diagram* if $g \circ f = k \circ h$. In other words, saying that a certain diagram commutes *is* an assertion that certain functions are equal.

Exercises 2.2

1. For each of the following pairs of functions $\mathbb{N} \to \mathbb{N}$, determine whether they are equal and explain why or why not.
   a) $f(n) = 2n + 3$ and $g(m) = 2m + 3$
   b) $f(n) = 2^{n+1} - 1$ and $g(n) = \sum_{i=0}^{n} 2^i$
   c) $f(n) = n^2 + 5n + 6$ and $g(n) = (n + 3)(n + 2)$
   d) $f(n) = n^4 - 10n^3 + 35n^2 + 50n + 24$ and $g(n) = 24$
2. Let $\mathbb{R}$ denote the set of all real numbers. Let $f(x) = \tan(x)$. Explain why this does *not* define a function from $\mathbb{R}$ to $\mathbb{R}$.
3. Suppose the following functions exist: $f\colon A \to B$, $g\colon B \to C$, $a\colon A \to D$, $b\colon C \to D$. Draw a commutative diagram asserting that $b \circ g \circ f = a$.
4. Suppose the following functions exist: $f\colon C \to A$, $g\colon C \to B$, $h\colon C \to P$, $p\colon P \to A$ and $q\colon P \to B$. Draw a commutative diagram asserting that $f = p \circ h$ and $g = q \circ h$.

3

Building Sets and Functions

Goals

Lecture

- Characterize and define
  - Pointer and constant functions
  - Solution sets
  - Characteristic functions
  - Products of sets
  - Exponents of sets
- Introduce the idea of a *universal* construction.

Study

- Be able to calculate membership in various constructed sets
- Learn to use universal constructions to define functions.

So far, we have been able to think mainly about finite sets and a few informally defined functions on, say, the real numbers or the natural numbers. To fill out our understanding of sets, we need to be able to build sets for specific purposes.

Two finite sets will play particularly important roles. The first, which we denote by $\mathbb{1}$, is set with one element; the second, which we denote by $\mathbb{2}$, is a set with two elements. It does not matter at all *what* elements are in these because, as we will soon see, any two sets of the same size are interchangeable. What 'interchangeable' means is discussed later. What 'same size' means is obvious for finite sets, but not at all apparent for infinite ones. We discuss the general situation later as well.

For the time being, we merely need to agree on a fixed set with one element and a fixed set with two elements. The particular choices I make here will be clearer as we put them to use.

Definition 3.1

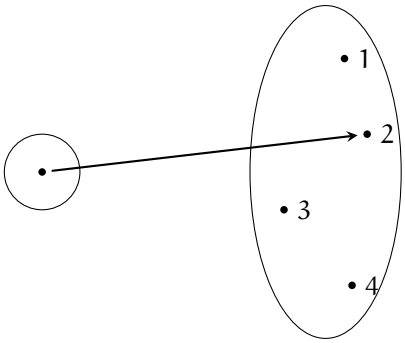
Let $\bullet$, $\bot$ and $\top$ be fixed symbols. Then define

$$\mathbb{1} = \{\bullet\}$$
$$\mathbb{2} = \{\bot, \top\}$$

The single element of $\mathbb{1}$ is intended to look like a generic point in an internal diagram. The element $\top$ is meant to remind you of the letter 'T' (short for 'True') and $\bot$ is meant to be the opposite of $\top$ (that is, 'False').

3.1 Elements, Pointers and Constant Functions

Suppose we are told that $p\colon \mathbb{1} \to A$ is a function. Since $\bullet \in \mathbb{1}$, this function determines an element of A, namely $p(\bullet)$. A picture of the situation might be this:



Since $\mathbb{1}$ has only a single element, it can "point" only to a single element of A. So we might refer to a function $\mathbb{1} \to A$ as a *pointer* into A. Each pointer determines an element of A. And conversely, it should be possible to point to any element of A. This leads to our first axiom guaranteeing that certain functions exist.

Principle 3.1

For any set A, and any $a \in A$, there is a pointer $\hat{a}\colon \mathbb{1} \to A$ so that $\hat{a}(\bullet) = a$.

In effect, this principle claims that elements of a set A and functions $\mathbb{1} \to A$ are interchangible: from $a \in A$ we get $\hat{a}\colon \mathbb{1} \to A$; from $p\colon \mathbb{1} \to A$. we get $p(\bullet)$.

This principle also justifies the drawing of internal diagrams for pointers because it means that any such diagram does depict an actual function. Other principles that we will discuss later justify our use of internal diagrams for all finite sets.

We can specify $\hat{a}$ by the rule $x \mapsto a$. That is, since $\bullet$ is an element of $\mathbb{1}$, $\hat{a}(\bullet) = a$. And since $\bullet$ is the only element of $\mathbb{1}$, $\hat{a}x = a$ is true for *every* element of $\mathbb{1}$.

Suppose $f\colon A \to \mathbb{1}$ and $g\colon A \to \mathbb{1}$ are functions, that is, their *codomain* is $\mathbb{1}$ instead their *domain*. Then $f(a) = g(a)$ is true for every $a \in A$ because $\bullet$ is the only possible value for $f(a)$ and $g(a)$. So $f = g$ by the Principle of Function Extensionality. In other words, there is at most one function from A to $\mathbb{1}$. But the rule $x \mapsto \bullet$ is as simple a rule as one can imagine. This leads to another axiom.

Definition 3.2

A set T is *terminal* if it is the case that for any set A there is exactly one function from A to T.

Axiom 3.1

The set $\mathbb{1}$ is a terminal set. We denote the unique function from A to $\mathbb{1}$ by $\diamond_A\colon A \to \mathbb{1}$.

The rule defining $\diamond_A$ must be

$$x \mapsto \bullet$$

because no other rule is possible.

Using $\diamond_A$ and $\hat{b}$ for an element $b \in B$, we can now define constant functions. That is $\hat{b} \circ \diamond_A$ is a function from A to B defined by

$$(\hat{b} \circ \diamond_A)(x) = \hat{b}(\diamond_A(x)) = \hat{b}(\bullet) = b.$$

In short, this is the function sending any element of A to the constant result b.

Exercises 3.1

1. Show that for any pointer $p\colon \mathbb{1} \to A$, it is the case that $\widehat{p(\bullet)} = p$.
2. Show that any set with exactly one element is a terminal set.
3. Suppose that $f\colon A \to B$ is a function. Show that for every $a \in A$, $\widehat{f(a)} = f \circ \hat{a}$.

3.2 Solution Sets, Subsets, Characteristic Functions

Suppose we are given two functions that are "parallel": $f\colon A \to B$ and $g\colon A \to B$. Then for some value $a \in A$, it might be the case that $f(a) = g(a)$. We may call such a value a *particular solution to the equation* $f(x) = g(x)$. It might be the case that there are no particular solutions. For example, there are no natural numbers n such that $n + 1 = n$. On the other hand, there might be many particular solutions. For example, let $f(x) = x^3$ and let $g(x) = 6x^2 - 11x + 6$ both as functions on the natural numbers. Then it is easy to check that 1, 2 and 3 solve the equation. In fact, these three are the only particular solutions. We generalize as follows.

Definition 3.3

For two functions $A \overset{f}{\underset{g}{\rightrightarrows}} B$, a *solution* is a function $s\colon C \to A$ so that

$$f \circ s = g \circ s.$$

Thus for example, if $a \in A$ is a particular solution then the pointer $\hat{a}$ is a solution.

For functions $f\colon A \to B$ and $g\colon A \to B$, an *equalizer* is solution $e\colon E \to A$ so that for any solution $s\colon C \to A$, there is exactly one function $h\colon C \to E$ so that

$$e \circ h = k.$$

Principle 3.2

For functions $f\colon A \to B$ and $g\colon A \to B$, the collection of all particular solutions to the equation $f(x) = g(x)$ form a set, denoted by $\{x \in A \mid f(x) = g(x)\}$. The function $i\colon \{x \in A \mid f(x) = g(x)\} \to A$ given by the rule $x \mapsto x$ (called an *inclusion map*) is an equalizer for f and g.

If $s\colon C \to A$ is a solution of $f(x) = g(x)$, that is, $f \circ s = g \circ s$, then the function $\check{s}\colon C \to \{x \in A \mid f(x) = g(x)\}$ given by the rule $x \mapsto s(x)$ is the unique function for which $s = i \circ \check{s}$.

This axiom tells us three main things. First, we can form a subset of A by specifying an equation $f(x) = g(x)$ for any two functions $A \overset{f}{\underset{g}{\rightrightarrows}} B$, and picking out the particular solutions. Second, a subset formed in this way "embeds" in the given set A by its inclusion map i. Third, for any solution s, the function into the set of particular solutions is defined by the same rule as s.

Suppose $c \in C$ and $f\colon A \to C$ is a function, then we can form the equalizer of f and the constant function $\hat{c} \circ \diamond_A$. This is more easily written we $\{x \in A \mid f(x) = c\}$. Since it common to pick out sets like this, special notation is in order.

Definition 3.4

For a function $f\colon A \to C$, and a value $c \in C$,

$$f^{-1}(c) = \{x \in A \mid f(x) = c\}.$$

In this case, $f^{-1}(c)$ is called the *inverse image of* c *with respect to* f.

Evidently, in Definition 3.4, the set $f^{-1}(c)$ is a subset of A. It would be good to know that any subset of A can be described as an inverse image. This is where the set 2 plays a role.

Definition 3.5

A *subset classifier* is a set S with a distinguished element $t \in S$ so that for any set A and any subset $B \subseteq A$, there is exactly one function $k\colon T \to A$ for which $B = k^{-1}(t)$. That is, B is uniquely defined as the inverse image of t with respect to a function into T.

Principle 3.3

The set 2 with the distinguished element $\top$ is a subset classifier. For subset $B \subseteq A$, the function corresponding to B, called the *characteristic function of* B, is denoted by κ_B. In other words, κ_B is the unique function for which $B = \kappa_B^{-1}(\top)$.

For $B \subseteq A$, the characteristic function is defined by the rule

$$x \mapsto \begin{cases} \top & \text{if } x \in B \\ \bot & \text{otherwise} \end{cases}$$

Just as Principle 3.1 asserts that elements of A and functions $1 \to A$ are interchangeable, Principle 3.3 asserts that the subsets of A and the functions $A \to 2$ are interchangeable.

Exercises 3.2

1. Draw a depiction of $A = \{a, b, c, d, e, f, g\}$ and its subset $B = \{a, c, e, g\}$ in the same internal diagram. Now depict the characteristic map for B as a subset of A.
2. Define two functions $\mathbb{N} \overset{f}{\underset{g}{\rightrightarrows}} \mathbb{N}$ so that the set of particular solutions fo $f(x) = g(x)$ is $\{1, 5\}$.
3. Consider the functions $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \sin(x) + \cos(x)$, and $s\colon \mathbb{N} \to \mathbb{R}$ defined by $s(n) = 2\pi n^2$. Is s a solution for the equation $f(x) = -1$? What is the set of all particular solutions?

3.3 Product Sets and Functions of Two Variables

We should be able to deal with functions of more than one argument, such as a function $f(x, y) = x + y$. To account for such functions, we take our cue from Descartes. He studied the plane in terms of a coordinate system consisting of the so-called x-axis and y-axis. Once we have decided where to place the axes (as long as they do not run in parallel), a pair $(2, 3)$ determines a point on the plane, and any point p in the plane determines a pair. So Descartes realized that we might as well just say that the plane actually

is the collection of all pairs of real numbers. What makes this work is that points in the plane *project* onto the two axes. Products of sets generalize this idea.

Definition 3.6

For sets A and B, a *table* consists of two functions $A \xleftarrow{f} C \xrightarrow{g} B$. Note that the two functions have the same domain. We may call the two functions *legs*.

For sets A and B, a *product of* A *and* B is a table $A \xleftarrow{p} P \xrightarrow{q} B$ so that for any table $A \xleftarrow{f} C \xrightarrow{g} B$ there is exactly one function $h\colon C \to P$ for which $f = p \circ h$ and $g = q \circ h$. For a product, the functions p and q are called the *projections*.

Principle 3.4

For sets A and B, the collection of all pairs (a, b) where $a \in A$ and $b \in B$ is a set, denoted by $A \times B$. The functions $\pi_0\colon A \times B \to A$ and $\pi_1\colon A \times B \to B$ defined by the rules $\pi_0(a, b) = a$ and $\pi_1(a, b) = b$ are projections. For $f\colon C \to A$ and $g\colon C \to B$, the unique function required by the product may be denoted by $\langle f, g\rangle$.

For $f\colon C \to A$ and $g\colon C \to B$, the function $\langle f, g\rangle\colon C \to A \times B$ is defined by the rule $\langle f, g\rangle(x) = (f(x), g(x))$.

Suppose we are given two unrelated functions $f\colon A \to B$ and $g\colon C \to D$. We can now form a single function from $A \times C$ to $B \times D$ by combining f and g "independently". That is, define $f \times g = \langle f \circ \pi_0, g \circ \pi_1\rangle$. Calculating concretely in terms of elements $(f \times g)(x, y) = (f(x), g(y))$.

Exercises 3.3

1. For the sets $A = \{a, b, c\}$ and $B = \{1, 2, 3, 4\}$, calculate $A \times B$ and $B \times A$
2. What is $\emptyset \times A$?
3. Calculate $\{4, a, 0\} \times 2$.
4. Describe in plain English what are the elements of $\mathbb{N} \times \mathbb{N}$.
5. Suppose A is a finite set with m elements and B is a finite set with n elements. How many elements are in $A \times B$? Describe in plain English why it makes sense to refer to $A \times B$ as a "product."

3.4 Function Sets, Partial Evaluation

A function from A to B might depend on a parameter from C. For example, the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by the rule $f(x) = \sin(x + c)$ depends on the constant c. There is a related function $g\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by $g(c, x) = \sin(x + c)$. Evidently, g describes the same behavior as f, but it makes the parameter explicit. This leads to the following definition.

Definition 3.7

For sets A and B, a *parametric function from* A *to* B is a function $\mathtt{g}\colon \mathrm{C} \times \mathrm{A} \to \mathrm{B}$ for some set C. The set C may be called the *set of parameters*.

Suppse $\mathtt{f}\colon \mathrm{D} \times \mathrm{A} \to \mathrm{B}$ is a parametric function with parameter set D and $\mathtt{k}\colon \mathrm{C} \to \mathrm{D}$ is a function. Then we can form another parametric function with parameters in C by composition:

$f \circ (k \times \mathsf{id}_A)$. We may refer to k as a *change of parameters* function because k transforms the parametric function with parameters in D into a parametric function with parameters in C. Specifically, $f \circ (k \times \mathsf{id}_A)$ is given by the rule $(c, a) \mapsto f(k(c), a)$.

An *evaluation map* for A and B is a parameteric function $a\colon F \times A \to B$, so that for any parametric function $g\colon C \times A \to B$ there is exactly one change of parameters $h\colon C \to F$ so that $g = a \circ (h \times \mathsf{id}_A)$. In that case, F is called a *function set* or an *exponential*.

Axiom 3.2

For sets A and B, the collection of all functions from A to B, denoted by B^A, is a set. Moreover, there is an evaluation map $\alpha\colon B^A \times A \to B$ defined by the rule $(f, x) \mapsto f(x)$. The unique change of parameters function corresponding to $f\colon C \times A \to B$ does not have a completely standard name. Some mathematicians honor the twentieth century logician names Haskell Curry by refering to this as 'currying'. For these lectures, we follow that tradition and write $\mathsf{curry}[f]$ for the unique function satisfiying $f = \alpha \circ (\mathsf{curry}[f] \times \mathsf{id}_A)$.

Calculating how $\mathsf{curry}[f]\colon C \to B^A$ must behave, we see that $\mathsf{curry}[f](c)$ is the function from A to B given by the rule $x \mapsto f(c, x)$. So for any $c \in C$ and $a \in A$, $\mathsf{curry}[f](c)(a) = f(c, a)$.

3.5 Powersets

The exponential 2^A for any set A is the set of all characteristic maps on A. So the elements of 2^A correspond to subsets of A, and vice versa. For technical reasons, it makes sense also to suppose that the actual collection of subsets of a set also form an exponential.

Principle 3.5

For any set A, the collection of subsets of A is a set, denoted by $\mathcal{P}(A)$, and called the *power set of* A. Moreover, the parametric function $\ni\colon \mathcal{P}(A) \times A \to 2$ defined by

$$\ni(B, x) = \begin{cases} \top, & \text{if } x \in B \\ \bot & \text{otherwise} \end{cases}$$

is an evaluation map.

For a function $f\colon C \times A \to 2$, then the rule $x \mapsto \{a \in A \mid f(c, a) = \top\}$ determines a function from C to $\mathcal{P}(A)$.

This principle means that any function $k\colon C \times A \to 2$ determines a function $h\colon C \to \mathcal{P}(A)$ satisfying $a \in h(c)$ if and only if $k(c, a) = \top$. But this same k also determines a subset of $C \times A$ by $k^{-1}(\top)$. So evidently functions $C \to \mathcal{P}(A)$, functions $(C \times A) \to 2$, functions $C \to 2^A$ and subsets of $C \times A$ are all interchangeable. In Lecture **??** we will discuss this interchange more thoroughly.

Exercises 3.5

1. Write out $\mathcal{P}(\{a, b, c\})$
2. Write out $\mathcal{P}(\emptyset)$
3. Is it the case that $\emptyset \in \mathcal{P}(A)$ for any set A? Explain.

4. Write out $\mathcal{P}(2 \times 2)$ and $\mathcal{P}(\mathcal{P}(2))$. Pay attention to writing them in a systematic way, so that it is clear you have actually listed everything.
5. I claim that $\mathcal{P}(\emptyset)$ is a terminal set (Definition 3.2). Justify the claim.
6. I claim that $\{\emptyset\} \in \mathcal{P}(\mathcal{P}(\emptyset))$ is a subset classifier (Definition 3.5). Justify the claim.

4

The Set of Natural Numbers

Goals

Lecture:

- Re-introduce the natural numbers as a set
- Introduce sequences and recursively defined sequences
- Relate recursion to proofs by induction

Study:

- Be able to define simple functions by recursion
- Be able to prove explain how induction and recursion are related

We have used $\mathbb{N}$ informally to denote the set of natural numbers. It is time that we make the structure of the natural numbers explicit, and investigate what role the natural numbers play in our theory of sets and functions

Natural numbers provide a precise picture of counting and of putting things in an order: first, second, third, and so on. Now that we have sets and functions as can consider a function $a\colon \mathbb{N} \to A$ to be an *infinite sequence*: $a(0)$, $a(1)$, $a(2)$, ... When we do that, we sometimes write a_0, a_1, a_2, ... instead, but the point is still that a itself is a function.

Much of what we will discuss in this lecture will have feel of computer programming about it. This is because there is a sense in which natural numbers are the main objects of calculation. We will want to understand, for example, how to define $n!$ (n factorial) as a function from $\mathbb{N}$ to $\mathbb{N}$ by specifying how it behaves. In particular, $0! = 1$ and $(n^\frown)! = n^\frown \cdot$ characterizes factorial by spelling out how to calculate it by recursion. For example,

$$\begin{aligned} 4! &= 4 \cdot 3! \\ &= 4 \cdot (3 \cdot 2!) \\ &= 4 \cdot (3 \cdot (2 \cdot 1!)) \\ &= 4 \cdot (3 \cdot (2 \cdot (1 \cdot 0!))) \\ &= 4 \cdot (3 \cdot (2 \cdot (1 \cdot 1))) \qquad\qquad = 24 \end{aligned}$$

We will actually think of $\mathbb{N}$ as being something like "the simplest set on which we can use recursion".

It is quite common to think about a sequence in which a_{n+1} is functionally related to a_n. For example, in the sequence $1, 2, 4, 8, \ldots$, each successive entry is double its predecessor. The initial entry is 1. Indeed, if we know just those two facts – the initial entry is 1, and each subsequent entry is double its predecessor – then we know how the entire sequence behaves. Also, we know how to calculate the n^{th} entry, by recursion just like factorial.

The most basic sequence, of course, is $0, 1, 2, \ldots$. Its initial entry is 0 and each subsequent entry is the successor of its predecessor. That really is about as basic as it gets. So we can think of this sequence as a sort of prototype for all others, including sequences in other sets. It turns out that this will characterize the set of natural numbers. And it will lead to a fairly general scheme for defining recursive functions.

4.1 Sequences and Successions

Let us make the informal word *sequence* official.

Definition 4.1

A *sequence in* A is a function $a\colon \mathbb{N} \to A$. For a sequence a, we may write a_k instead of $a(k)$, but these mean exactly the same thing.

As we studied in previous lectures, the basic vocabulary of natural numbers is that (i) there is a starting natural number 0 and (ii) for each natural number n there is a next one, $n^\frown$. To put the successor in the language of sets and functions, we can stipulate that successor is a function $\mathsf{suc}\colon \mathbb{N} \to \mathbb{N}$ given by the rule $n \mapsto n^\frown$. So $\mathbb{N}$ is not only a set. It comes with functions $\mathbb{1} \xrightarrow{\hat{0}} \mathbb{N} \xrightarrow{\mathsf{suc}} \mathbb{N}$.

Suppose $\mathbb{1} \xrightarrow{\hat{b}} A \xrightarrow{r} A$ is a similar structure. Then we ought to be able to define a sequence in$A$, i.e., a function $a\colon \mathbb{N} \to A$, so that $a_0 = b$, $a_1 = r(b)$, $a_2 = r(r(b))$, and so on. In general, a_k should be determined by starting with b and repeatedly applying r a total of k times.

Definition 4.2

A *simple recurrence* is a set with functions $\mathbb{1} \xrightarrow{\hat{b}} C \xrightarrow{r} C$.

A *countably infinite set* is a simple recurrence $\mathbb{1} \xrightarrow{\hat{z}} N \xrightarrow{s} N$ so that for any other simple recurrence $\mathbb{1} \xrightarrow{\hat{b}} C \xrightarrow{r} C$, there is exactly one function $f\colon N \to C$ so that $f \circ \hat{z} = \hat{b}$ and $f \circ s = r \circ f$.

Simple recurrences go by lots of other names in the literature. You will recognize them later when you see them.

The principle we are interested in here is that simple recurrences determine sequences. In the next section, we discuss why we have called these "simple".

Principle 4.1

The collection of natural numbers is a set, denoted by $\mathbb{N}$. Moreover, $\mathbb{1} \xrightarrow{\hat{0}} \mathbb{N} \xrightarrow{\mathsf{suc}} \mathbb{N}$. makes $\mathbb{N}$ a countably infinite set.

From a simple recurrence $\mathbb{1} \xrightarrow{\hat{b}} C \xrightarrow{r} C$, the corresponding unique sequence in C may be denoted by $\mathsf{rec}[b, r]$. So $\mathsf{rec}[b, r]\colon \mathbb{N} \to A$ is defined by

$$\mathsf{rec}[b, r](0) = b$$
$$\mathsf{rec}[b, r](n^\frown) = r(\mathsf{rec}[b, r](n))$$

Thus every simple recurrence in C determines a sequence in C. On the other hand, it is not the case that every sequence is determined by a simple recurrence. Take for example, the sequence $0, 1, 0, 2, 0, 3, \ldots$. This can not be defined by giving an initial entry (0) and specifying successive entries based only on the predecessors. After all, the entries 1, 2, 3 and so on all have the same preceding entry.

4.2 Primitive Recursion

Evidently, addition, multiplication, factorial, and other familiar functions should be definable using Principle 4.1. But there are problems to overcome: Addition is not a sequence, at least not in an obvious way. And factorial is not obviously definable by a simple recurrence because we would need a function $r\colon \mathbb{N} \to \mathbb{N}$ so that $n^\frown! = r(n!)$ for all n. What that function might be is not at all clear. If, in place of r,

we could use a function that depends on n as well as on $n!$, we could define factorial recursively the usual way.

Although addition is not a sequence, it is is a parametric function $+\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ (where the first $\mathbb{N}$ is treated as the parameter set). If we could some how extend simple recursion to parametric functions, then addition would also be definable (by adapting the definition we gave in earlier lectures).

Putting things together, we consider a scheme that generalizes simple recursion to permit (i) dependence on n at each stage of the recursion and (ii) dependence on a parameter.

Definition 4.3

A *primitive recurrence in* A *(with parameters in* C*)* consists of two functions $C \xrightarrow{b} A \xleftarrow{r} \mathbb{N} \times C \times A$.

A *parametric sequence in* A *with parameters in* C is a function $a\colon C \times \mathbb{N} \to A$. For a parametric sequence, we might write $a_c(n)$ instead of $a(c, n)$.

Lemma 4.1

Any primitive recurrence $C \xrightarrow{b} A \xleftarrow{r} C \times \mathbb{N} \times A$, there is a unique function $f\colon C \times \mathbb{N} \to A$ satisfying:

$$f(c, 0) = b(c)$$
$$f(c, k^\frown) = r(c, k, f(c, k))$$

Proof: □

The "predecessor" function is defined by the scheme $\mathsf{pred}(0) = 0$ and $\mathsf{pred}(n^\frown) = n$.

Exercises 4.2

1. For a given function f, find a primitive recurrence that defines the function $n \mapsto \sum_{i=0}^{n-1} f(i)$. That is, result will be the sequence 0, $f(0)$, $f(0) + f(1)$, $f(0) + f(1) + f(2)$, ...
2. Find a primitive recurrence that defines the function from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$ given by the rule $(m, n) \mapsto m^n$.
3. Define the operation of *monus* $m \dot{-} n$ to be $m - n$ when $m \geq n$ and to be 0 otherwise by a primitive recurrence.

5

Monomophisms, Epimorphisms and Isomorphisms

Goals

Lecture

- Introduce special kinds of functions: monomorphism, epimorphism, injection, surjection.
- Investigate the relations between these.

Study

- Learn to recognize injections and surjections by internal behavior
- Be able to illustrate simple, small examples of injections, non-injections, surjections, non-surjections.

Recall that we required that if a natural number has a predecessor, it has exactly one. We wrote that as an axiom: $m^{\frown} = n^{\frown}$ implies $m = n$. Putting this in terms of functions: $\mathsf{suc}(m) = \mathsf{suc}(n)$ implies $m = n$.

Also, recall that we proved, using this axiom, that addition is *cancellative*: $m + p = n + p$ implies $m = n$, and that multiplication by a positive natural number is cancellative: $m \cdot p^{\frown} = n \cdot p^{\frown} = m = n$.

The general pattern of these examples is that if two expression that differ only by m and n are equal, then $m = n$. This leads to some definitions.

Definition 5.1

For a function $g\colon B \to C$, say that

- g is an *injection* if it is the case that $g(x) = g(y)$ implies $x = y$ for every $x, y \in A$;
- g is a *monomorphism* if for every $f_0, f_1\colon A \to B$, $g \circ f_0 = g \circ f_1$ implies $f_0 = f_1$; and
- g is an *epimorphism* if for every $h_0, h_1\colon C \to D$, it is the case that $h_0 \circ g = h_1 \circ g$ implies $h_0 = h_1$.

So an injection is a function with the property that application is cancellative. A monomorphism is a function with the property that composition on the left is cancellative. An epimorphism is a function with property that composition on the right is cancellative.

An injection is also said to be *injective* or *one-to-one*. A monomorphism is also said to be *left-cancellable*. An epimorphism is also said to be *right-cancellable*.

Example 5.1

1. Consider the function $f\colon \mathbb{N} \to \mathbb{N}$ defined as $f(n) = 2 \cdot n + 2$. Suppose $f(m) = f(n)$, then $2 \cdot m + 2 = 2 \cdot n + 2$. But addition is cancellative, so $2 \cdot m = 2 \cdot n$. And multiplication by a positive natural number is cancellative, $m = n$. Thus f is an injection.
2. Consider $g\colon \mathbb{R} \to \mathbb{R}$ defined as $f(x) = x^2$. This is not injective because $f(1) = f(-1)$, but $1 \neq -1$.

Other stuff