Pre-Requisites

Let's cover some of the notation and concepts we'll need for later.

Section 1

Set Operations: Set Difference, Union, and Intersection

One of the useful characteristics of the very general and abstract definition of sets is that they allow for the combination of sets in ways that depend only upon the membership relation, and not on what the elements actually are. To realize this are the fundamental set theoretic operations of **set difference**, **union**, and **intersection**. These three important operations on sets can be intuitively understood with reference to **Venn Diagrams**, where circles represent sets and the shaded area within them denotes the elements contained within the set of interest.

Subsection 1.1

Set Difference and Complements

First, we have the **set difference** operation: given two sets, A and B, we denote the set that contains all elements of A that are not in B by

$$A - B = A \setminus B := \{x \in A \mid x \notin B\}.$$

Example 1

Let $A = \{0, 1, 2, 3\}$ and $B = \{3, 4, 5\}$. Then $A \setminus B = \{0, 1, 2\}$ and $B \setminus A = \{4, 5\}$.

Example 2

Let

$$A = \{n \mid n \text{ is an even natural number}\} \quad \text{and} \quad B = \{n \mid n \text{ is a natural number divisible by } 2\}.$$

Then $A \setminus B = \{n \mid n \text{ is a product of 2 and an odd number}\}$.

Closely related is an additional operator called **complement** (relative $\mathbb{U}$), denoted A^{c}, A', or $\overline{A}$, and is defined as the set of all elements of some pre-determined 'universe' $\mathbb{U}$ not in the A. That is, $A^{\mathsf{c}} = \mathbb{U} \setminus A$; this differs from the normal set difference only in the sense that $\mathbb{U}$ is a fixed set that depends upon the context and simplifies notation.

Example 3

The complement of

$$A = \{n \mid n \text{ is an even natural number}\}$$

relative to $\mathbb{N}$ is

$$\mathbb{N} \setminus A = A^{\mathsf{c}} = \{n \mid n \text{ is an odd natural number}\}.$$

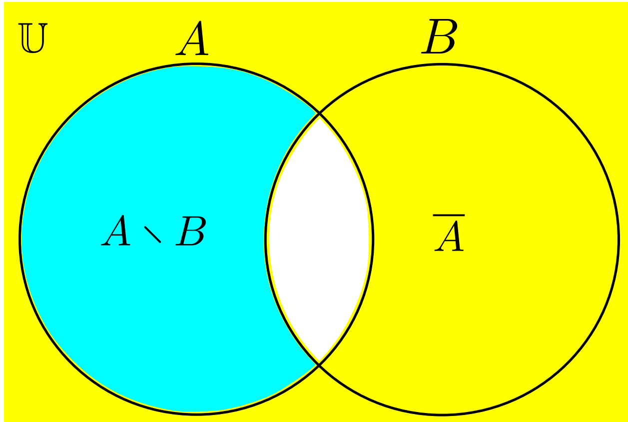

Figure 1: Here, the entire box represents the universe, the blue shaded area is $A \setminus B$, and the yellow shaded area is A^{c}.

Subsection 1.2

Union

The **union** operation, denoted $A \cup B$, is an operation which takes all the elements from A and B, and puts them into a single set, or

$$A \cup B := \{x \mid x \in A \text{ or } x \in B\}.$$

The union operation can be extended to an operation on any collection of sets by defining

$$\bigcup_{x \in X} x := \{y \mid \text{there exists } x \in X \text{ such that } y \in x\}.$$

Example 4

With $A = \{0, 2, 4\}$ and $B = \{1, 2, 4\}$, we have $A \cup B = \{0, 1, 2, 4\}$.

Example 5

The union of all subsets of a set X is X itself; symbolically, $\bigcup_{X \in \mathcal{P}(Y)} X = Y$.

Subsection 1.3

Intersection

The **intersection** operation, denoted $A \cap B$, is an operation which takes all the elements that are both in A and B, and puts them into a single set, or

$$A \cap B := \{x \mid x \in A \text{ and } x \in B\}.$$

Just like union, intersection can be extended to any collection of sets by defining

$$\bigcap_{x \in X} x := \{y \mid \text{for each } x \in X \text{ we have } y \in x\}.$$

If the intersection of two sets is the null set, then the sets are said to be **disjoint**. We write $A \sqcup B$ for the union of A and B when A and B are disjoint. Likewise, we write $\bigsqcup_{x \in X} x$ for the union $\bigcup_{x \in X} x$ when the elements of X are **pairwise disjoint**, meaning that any two distinct elements of X are disjoint; equivalently,

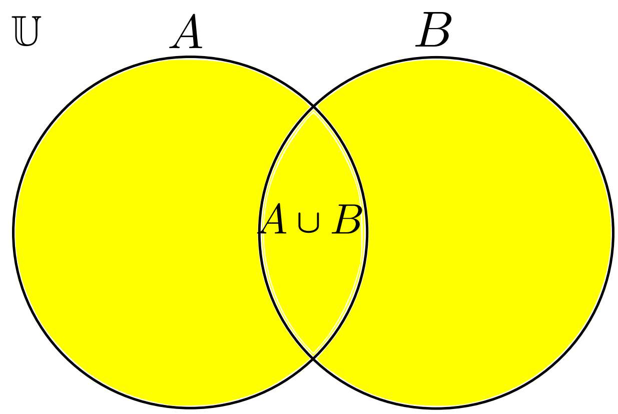

Figure 2: The Venn Diagram representation of $A \cup B$.

that $\bigcap_{x \in X} x = \varnothing$. We say that a union $\bigcup_{x \in X} x$ is an *disjoint* union when the elements of X are pairwise disjoint.

Example 6

With $A = \{0, 2, 4\}$ and $B = \{1, 2, 4\}$, we have $A \cap B = \{2, 4\}$.

Example 7

Let

$$A = \{n \mid n \text{ is a natural number divisible by } 2\}$$

and

$$B = \{n \mid n \text{ is a natural number divisible by } 3\}.$$

Then

$$A \cap B = \{n \mid n \text{ is a natural number divisible by } 6\}.$$

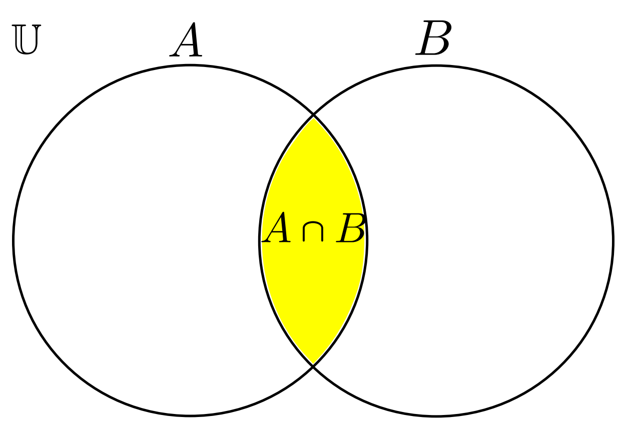

Figure 3: The Venn Diagram representation of $A \cap B$.

Section 2

Mappings Between Sets

In this section, we develop the definition of relations and functions, the latter of which can be informally thought of as an object that pairs an input and an output in a unique way. Naturally, we must develop some tools for describing these pairs:

Subsection 2.1

Ordered Pairs and Cartesian Products

Closely related to the concept of a set is that of the **ordered pair** , denoted (a, b), where a and b are known as the first and second **coordinates**, respectively, and also as the **abscissa** and **ordinate**, respectively, though we shall never use this vocabulary. Note that, as the name implies, order is important here: (a, b) is not equal to (b, a) in general, and in particular $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$; this is the characteristic property of the ordered pair, and it is the only property we need for our uses. One way of defining an ordered pair (a, b) in terms of sets is the following:

$$(a, b) := \{\{a\}, \{a, b\}\}.$$

Naturally, we check that this fulfills the properties we want: that $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$.

Lemma 2.1: Characteristic Property of Ordered Pairs

$(a, b) = (c, d)$ if and only if $a = c$ and $b = d$, where $(x, y) = \{\{x\}, \{x, y\}\}$.

We can extend this further to define ordered triplets, (a, b, c), by using a recursive definition: $(a, b, c) := ((a, b), c)$. Then the corresponding set is

$$(a, b, c) = \{\{(a, b)\}, \{(a, b), c\}\} = \{\{\{a\}, \{a, b\}\}\{\{a, \{a, b\}\}, c\}\}.$$

We can repeat this for any whole number n greater than two to produce an **ordered n-tuplet** , defining

$$(x_1, x_2, \dots, x_n) := ((x_1, x_2, \dots, x_{n-1}), x_n),$$

where the naming convention for elements follows in the same way as for the ordered pair; i.e. the first coordinate is x_1, the second coordinate is x_2, the third coordinate is x_3, etc.

In order to describe the act of pairing elements of the set A with the elements of the set B, we must first describe the set of all these pairs, called the **Cartesian product**

$$A \times B := \{(a, b) \mid a \in A \text{ and } b \in B\}$$

the set of all ordered pairs with first coordinate an element of A and second coordinate an element of B. Just as we generated the construction of the ordered pair to ordered n-tuplets, we can create Cartesian products of more than two sets using our general n-tuplets in our definition:

$$X_1 \times X_2 \times \cdots \times X_n = \{(x_1, x_2, \dots, x_n) \mid x_i \in X_i \text{ for each } 1 \le i \le n\}.$$

We denote the product $X \times \cdots \times X$, where X appears n-times, as X^n. Also, we define $X^0 = \{\varnothing\}$ for any set X; we shall see some motivation for this definition later.

Example 8

Suppose $A = \{1, 2\} = B$. Then

$$A \times B = \{(1, 1), (1, 2), (2, 1), (2, 2)\}.$$

Example 9

$\mathbb{R} \times \mathbb{R}$ is the set of all pairs of real-numbers, and is typically called the Cartesian plane.

Example 10

Given two sets A and B, we define their **disjoint union** by

$$A \sqcup B := (A \times \{0\}) \cup (B \times \{1\}).$$

Subsection 2.2

Relations and Functions

Intuitively, a function is a kind of object that one feeds in some element of a set A, and is given an element of a set B in return. The natural way to describe this mathematically is by associating to each element of A a single element of B, and ordered pairs give us precisely the tool needed to do this. Hence, we can describe a **(binary) relation between** A **and** B as a subset R of $A \times B$, or more formally as a triple (A, B, R) with $R \subset A \times B$; when $A = B$ we may instead say that (A, A, R) is a (binary) relation *on* A. We call R the **graph of the relation** (A, B, R), and if $\mathscr{R} = (A, B, R)$, then we write $R = \mathrm{Gr}(\mathscr{R})$.

A **function** is a relation (A, B, f) such that

(i) **well-defined** (or **functional**): if $(a, b), (a, b') \in f$ then $b = b'$, and

(ii) **left-total**: for every $a \in A$ there exists a $b \in B$ such that $(a, b) \in f$;

this captures the idea that a function *assigns* to every element of A a single value of B. We shall denote a function (A, B, f) by $f : A \to B$, and given $(a, b) \in f$ we write $f(a) = b$ or $f : a \mapsto b$. Sometimes we shall just say 'f' to stand for the full function '(A, B, f)', identifying a function with its graph, when the domain and codomain are understood (or irrelevant).

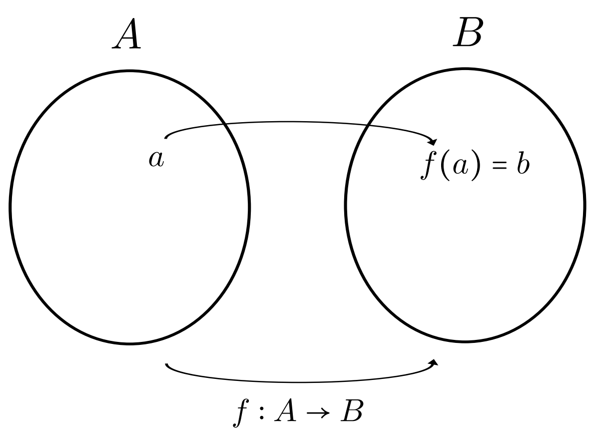

Figure 4: A visualization of a mapping between the sets A and B, as well as an element $a \in A$ being mapped to $b \in B$.

Example 11

Let $A = B$ be any non-empty set. Then take $R = \{(x, x) \in A \times B \mid x \in A\}$. Then (A, B, R) is a relation between A and B.

Example 12

Let $A = \{1,2\}$ and $B = \{1,2,3\}$. Define $f = \{(1,2),(2,3)\}$ and $g = \{(1,2),(1,3),(2,2)\}$. Then f is a function, but g is not because it is not well-defined (1 is mapped to both 2 and 3).

Example 13

Let $A = \{0,1,2\}$ and $f = \{(0,0),(1,0),(2,1)\}$. (A,A,f) is a function, for there is exactly one value corresponding to each $a \in \{0,1,2\}$. However, with $B = \{0,1\}$, the relation (A,B,f) is *also* a function.

Example 14

Consider the function from $\mathcal{P}(C) \times \mathcal{P}(C)$ to $\mathcal{P}(C)$ that assigns to each pair of sets $(A,B) \in \mathcal{P}(C) \times \mathcal{P}(C)$ their union $A \cup B$; more generally, the assignments $(A,B) \mapsto A \cap B$, $(A,B) \mapsto A \,\Delta\, B$, and $(A,B) \mapsto A \setminus B$ are functions.

Example 15

Let A be any set, and take $\mathrm{id}_A : A \to A$ to be defined by $\mathrm{id}_A : a \mapsto a$. This is a function called the **identity function on** A.

Example 16

Let A be any set and B any non-empty set. Take $b \in B$, and define $\mathrm{const}_b : A \to B$ by $\mathrm{const}_b : a \mapsto b$. If we need to be explicit about A and B, then we shall denote this constant function by $\mathrm{const}_{b,A,B}$.

Example 17

Suppose we have sets A and B such that $A \subset B$. Then the **inclusion map** $\iota : A \to B$ is defined by $\iota(a) = a$ for $a \in A$. When we want to be explicit about A and B, we shall denote this inclusion map by $\iota_{A \subset B}$.

Example 18

$\varnothing$ is a function from $\varnothing$ to any set A, for $\varnothing = \varnothing \times A$ so that $(\varnothing, \varnothing, A)$ is a relation for any A. Moreover, it satisfies left-totality and well-definedness vacuously, meaning that all its elements fulfill the required properties simply because there are no elements to consider. $(\varnothing, \varnothing, A)$ is called the **empty function into** A and is a special case of an inclusion $\iota_{\varnothing \subset A}$.

Example 19

Suppose A and B are sets. Then we define $\pi_1 : A \times B \to A$ by $\pi_1(a,b) = a$ and $\pi_2 : A \times B \to B$ by $\pi_2(a,b) = b$. These are functions called the **projection onto the first coordinate** and **projection onto the second coordinate**, respectively, and together are known as the **canonical projections**.

Example 20

Suppose A and B are sets. Then we define $\iota_1 : A \to A \amalg B$ by $\iota_1(a) = (a,0)$ and $\iota_2 : B \to A \amalg B$ by $\iota_2(b) = (b,1)$. These are functions called the **canonical injections**.

Now, we defined a relation as a *triple* (A,B,R); in other words, we made the sets A and B explicit in the definition. This is because, given a subset $R \subset A \times B$, there can exist C and D such that $R \subset C \times D$. Indeed, if $R \subset A \times B$ and $A \subset C$, $B \subset D$, then $R \subset C \times B$. It is for this reason that we must be explicit about A and B.

Usually, the distinction won't be wholly necessary, but there are a few technical reasons for why we should want this. One of these reasons has to do with considering the collection of all relations between X and Y, and so in order to distinguish between the elements of two collections of this form, it is necessary to specify A and B. A second reason regards the special case of functions and a concept known as *surjectivity*, which we define later.

In the case where we are considering a function $f : A \times B \to C$, we shall denote the image of an ordered pair (a, b) by $f(a, b)$ rather than $f((a, b))$ in order to simplify notation (for the additional parentheses don't really add anything); this notation naturally extends to larger products. A function $f : A^n \to B$ is said to be an **n-ary function**; 0-ary functions can be identified with an element of B, for a function $f : A^0 \to B$ is uniquely identified by an element of B and also does not depend upon A, since $A^0 = \{\varnothing\}$ for every set A.

We denote the **function space**, or the set of all functions $f : A \to B$, as B^A.

Example 21

The set 2^X is the set of all functions from X to the two-element set $2 \coloneqq \{0, 1\}$.

Given a relation (A, B, R), the set $\operatorname{dom} R \coloneqq A$ is called the **domain**, and the set $\operatorname{cod} R \coloneqq B$ is called the **codomain**. We also define

$$\operatorname{im} R \coloneqq \{b \in B \mid \text{there exists } a \in A \text{ such that } (a, b) \in R\}$$

called the **image** of R, and

$$\operatorname{im}^{-1} R \coloneqq \{a \in A \mid \text{there exists } b \in B \text{ such that } (a, b) \in R\}$$

called the **pre-image** of R. When R is a function, left-totality guarantees that $\operatorname{dom} R = \operatorname{im}^{-1} R$

We can generalize the above definitions to subsets of A and B, so that if $A' \subset A$ and $B' \subset A$, then

$$R[A'] \coloneqq \{b \in B \mid \text{there exists } a \in A' \text{ such that } (a, b) \in R\}$$

called the **image of A' under R** and

$$R^{-1}[B'] = \{a \in A \mid \text{there exists } b \in B' \text{ such that } (a, b) \in R\}.$$

Hence, we see that $\operatorname{im}^{-1} R = R^{-1}[B]$ and $\operatorname{im} R = R[A]$. Admittedly, we shall use this notation principally with functions. When A' and B' are singletons, $\{a\}$ and $\{b\}$, respectively, and R is a function, we drop the brackets to just write $f(a) = b'$ and $f^{-1}(b) = A'$ where the latter *set* is called the **fibre of b under f**.

Example 22

Let $f : \{0, 1, 2\} \to \{0, 1, 2\}$ be the function defined in Example 32. Then $\operatorname{dom} f = \{1, 2, 3\} = \operatorname{cod} f$, $\operatorname{im} f = \{0, 1\}$, $f^{-1}(2) = \varnothing$, and $f^{-1}(0) = \{0\}$.

Example 23

Let $f : \mathcal{P}(C)^2 \to \mathcal{P}(C)$ be the map defined by $f : (A, B) \mapsto A \cup B$. The $\operatorname{dom} f = \mathcal{P}(C)^2$, $\operatorname{cod} f = \mathcal{P}(C)$, $\operatorname{im} f = \mathcal{P}(C)$, and $f^{-1}[A] = \{(B, A \setminus B) \mid B \subset A\}$ for every $A \in \mathcal{P}(C)$.

Example 24

Consider $\mathrm{id}_A : A \to A$. Then $\operatorname{dom} \mathrm{id}_A = A = \operatorname{cod} \mathrm{id}_A = \operatorname{im} \mathrm{id}_A$ and $\mathrm{id}_A^{-1}(a) = \{a\}$ for every $a \in A$.

Example 25

Consider $\mathrm{const}_b : A \to B$. Then $\mathrm{dom}\,\mathrm{const}_b = A$, $\mathrm{cod}\,\mathrm{const}_b = B$, $\mathrm{im}\,\mathrm{const}_b = \{b\}$, and $\mathrm{const}_b^{-1}[C]$ is equal to A when $b \in C$ and $\varnothing$ otherwise.

Example 26

Consider $\iota : A \to B$ the inclusion map of A into B. Then $\mathrm{dom}\,\iota = A$, $\mathrm{cod}\,\iota = B$, $\mathrm{im}\,\iota = A$, and $\iota^{-1}(b)$ is equal to $\{b\}$ if $b \in A$ and $\varnothing$ otherwise.

Subsection 2.3

Operations On and Between Relations and Functions

Just as there are useful set-theoretic operations on and between sets, it is natural to ask what meaningful manipulations exist on and between functions:

The first is to focus on a function's particular action on a subset of the domain or of the codomain: given a function $f : A \to B$ and $A' \subset A$, the **restriction of f to A'**, denoted $f|_{A'}$, is the function $f|_{A'} : A' \to B$ defined by

$$f|_{A'}(x) \coloneqq f(x)$$

for $x \in A'$; all the restriction essentially does is *restrict* the domain that f acts on.

Given functions $f : A \to B$ and $f' : A' \to B$, then we say that f' is a **extension of f** if $A \subset A'$ and $f'|_A = f$.

Given $f : A \to B$ and $g : B \to C$, we can envision f as acting on the elements of A, and g as acting on the elements of B. Then it is rather natural to imagine acting on A via f immediately followed by g. Alternatively, given relation binary relations (A, B, R) and (B, C, S), then R is relating elements of A and B, while S is relating elements of B and C; in this case, we can imagine relating elements of A with elements of C by using B as a sort of intermediary. With these motivations in mind, for binary relations (A, B, R) and (B, C, S), we define their **composition** to be the relation $(A, C, S \circ R)$ where $S \circ R$ is defined by

$$S \circ R \coloneqq \{(a, c) \in A \times C \mid \text{there exists } b \in B \text{ such that } (a, b) \in R \text{ and } (b, c) \in S\}.$$

Given functions $f : A \to B$ and $g : B \to C$, their composition is a function as well:

Lemma 2.2

Given $f : A \to B$ and $g : B \to C$, their composition $g \circ f$ is a function defined by $g \circ f : a \mapsto g(f(a))$.

Composition is *associative*, or in other words, it does not matter which order we perform our compositions in:

Lemma 2.3

For any functions $f : A \to B$, $g : B \to C$, and $h : C \to D$, then $h \circ (g \circ f) = (h \circ g) \circ f$.

Example 27

Let $A = B = C$ be the set of all humans. Then let $(a, b) \in R$ if and only if a is a sibling of b (i.e. share the same parents), and $(b, c) \in S$ if and only if b is a child of c. Then $S \circ R = R$, but $(a, b) \in R \circ S$ if and only if b is an aunt or uncle of a.

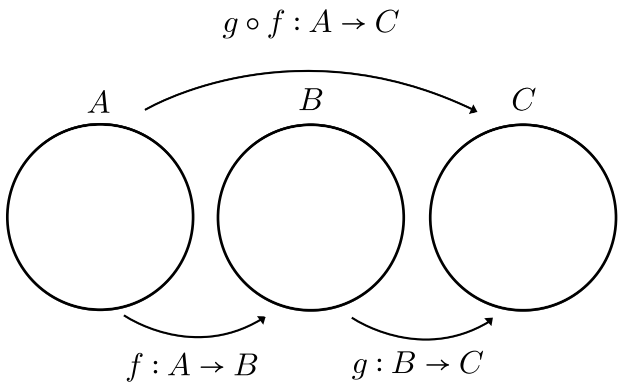

Figure 5: A visual representation of the composition of two functions.

EXAMPLE 28

If $f : \{0,1\} \to \{0,1\}$ is defined by $f = \{(0,1),(1,0)\}$ and $g : \{0,1\} \to \{0,1\}$ is defined by $g = \{(0,1),(1,1)\}$, then $g \circ f = \{(0,1),(1,1)\}$ and $f \circ g = \{(0,0),(1,0)\}$.

EXAMPLE 29

If $f : A \to B$ is any function, then $\mathrm{id}_B \circ f = f$ and $f \circ \mathrm{id}_A = f$. This is why id_B and id_A are called 'identity functions'.

EXAMPLE 30

Given $f : A \to B$, $A' \subset A$, and $\mathrm{im}\, f \subset B' \subset B$, then $f|_{A'} = f \circ \iota_{A' \subset A}$ and $\iota_{B' \subset B} \circ f|^{B'} = f$.

Given a relation (A, B, R), we define its **inverse** to be the relation (B, A, R^{-1}) where

$$R^{-1} := \{(b,a) \in B \times A \mid (a,b) \in R\}.$$

The use of R^{-1} mirrors the notation used with pre-images of sets under the relation R, for given $B' \subset B$, the pre-image of B' under R is exactly the image of B' under R^{-1}. In general, even when (A, B, f) is a function, (B, A, f^{-1}) need *not* be a function; when it is a function, we say that the "inverse of f exists". There are two conditions necessary for the inverse of a function to itself be a function:

LEMMA 2.4

Let $f : A \to B$ be a function. Then its inverse f^{-1} is a function if and only if for every $b \in B$ there exists a unique $a \in A$ such that $(a,b) \in \mathrm{Gr}\, f$.

When the inverse function exists, it acts as a proper inverse under function composition:

PROPOSITION 2.5

Let $f : A \to B$ be any function, and suppose that its inverse f^{-1} is a function. Then $f \circ f^{-1} = \mathrm{id}_B$ and $f^{-1} \circ f = \mathrm{id}_A$, and if there exists a function $g : B \to A$ such that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$, then $f^{-1} = g$.

Subsection 2.4

Injectivity, Surjectivity, Bijectivity, and Inverse Functions

Recall the properties that characterized a function: the relation (A, B, f) is a function if and only if it is

(i) left total – for every $a \in A$ there is some $b \in B$ such that $(a, b) \in f$, or in other words every input has an associated output – and

(ii) well-defined – given $(a, b), (a, b') \in f$, then $b = b'$, or in other words that the output is unique.

There are dual statements to left totality and well-definedness by switching the roles of the first and second coordinates. Dual to left totality is **surjectivity**: a relation (A, B, f) is *surjective* (or **onto**) if for every $b \in B$ there exists $a \in A$ such that $(a, b) \in f$. Dual to well-definedness is **injectivity**: a relation (A, B, f) is *injective* (or **one-to-one**) if $(a, b), (a', b) \in f$ implies $a = a'$. A surjective function is called a **surjection**, while an injective function is called a **injection**.

The ability to make the notion of surjectivity meaningful is only possible because we made the codomain explicit in the definition of a relation, and this is perhaps the most important motivation for doing this.

If a function is both injective and surjective, we say that it is **bijective**, or that the function is a **bijection**; another term for this is that the function is an *isomorphism*, or that the domain and codomain are *isomorphic*, though this term is generally used when some additional structure is preserved, such as when dealing with algebras or an ordering on the set, so we shall reserve the term 'isomorphism' for such cases.

We can reinterpret Lemma 2.4 in terms of bijectivity: a function $f : A \to B$ is a bijection if and only if f^{-1} is a function.

Example 31

Let $A = \{1, 2\}$ and $B = \{1, 2, 3\}$ and define $\operatorname{Gr} f = \{(1, 2), (2, 3)\}$. Then f is injective but not surjective.

Example 32

Let $A = \{0, 1, 2\}$ and $\operatorname{Gr} f = \{(0, 0), (1, 0), (2, 1)\}$. Then f is neither injective nor surjective.

Example 33

The map $f : \mathcal{P}(C) \times \mathcal{P}(C) \to \mathcal{P}(C)$ defined by $f(A, B) = A \cup B$ is always surjective, since given $A \in \mathcal{P}(C)$ we have $f(A, \varnothing) = A$, but in general is not injective, for if C is non-empty, then $C \cup \varnothing = \varnothing \cup C$.

Example 34

The identity map id_A is always a bijection with inverse id_A.

Example 35

Given sets A and B with $b \in B$, $\mathrm{const}_b : A \to B$ is only injective when A consists of only at most one element; likewise, it is surjective only when B has a single element (for B cannot be empty by hypothesis that $b \in B$).

Example 36

The inclusion map $\iota : A \to B$ is an injection, but is only surjective when $A = B$.

Lemma 2.6

Let $f : A \to B$ and $g : B \to C$ be functions. Then if f and g are injective, surjective, or bijective, then $g \circ f$ is injective, surjective, bijective, respectively.