Chapter 2

Basic Set Theory

> **A set is a Many that allows itself to be thought of as a One.**
> - *Georg Cantor*

This chapter introduces set theory, mathematical induction, and formalizes the notion of mathematical functions. The material is mostly elementary. For those of you new to abstract mathematics elementary does not mean *simple* (though much of the material is fairly simple). Rather, elementary means that the material requires very little previous education to understand it. Elementary material can be quite challenging and some of the material in this chapter, if not exactly rocket science, may require that you adjust you point of view to understand it. The single most powerful technique in mathematics is to adjust your point of view until the problem you are trying to solve becomes simple.

Another point at which this material may diverge from your previous experience is that it will require proof. In standard introductory classes in algebra, trigonometry, and calculus there is currently very little emphasis on the discipline of *proof*. Proof is, however, the central tool of mathematics. This text is for a course that is a students formal introduction to tools and methods of proof.

2.1 Set Theory

A *set* is a collection of distinct objects. This means that $\{1, 2, 3\}$ is a set but $\{1, 1, 3\}$ is not because 1 appears twice in the second collection. The second collection is called a *multiset*. Sets are often specified with curly brace notation. The set of even integers can be written:

$$\{2n : \text{n is an integer}\}$$

The opening and closing curly braces denote a set, $2n$ specifies the members of the set, the colon says "such that" or "where" and everything following the colon are conditions that explain or refine the membership. All correct mathematics can be spoken in English. The set definition above is spoken "The set of twice n where n is an integer".

The only problem with this definition is that we do not yet have a formal definition of the integers. The integers are the set of whole numbers, both positive and negative: $\{0, \pm 1, \pm 2, \pm 3, \ldots\}$. We now introduce the operations used to manipulate sets, using the opportunity to practice curly brace notation.

Definition 2.1 *The* **empty set** *is a set containing no objects. It is written as a pair of curly braces with nothing inside* $\{\}$ *or by using the symbol* $\emptyset$.

As we shall see, the empty set is a handy object. It is also quite strange. The set of all humans that weigh at least eight tons, for example, is the empty set. Sets whose definition contains a contradiction or impossibility are often empty.

Definition 2.2 *The* **set membership symbol** $\in$ *is used to say that an object is a member of a set. It has a partner symbol* $\notin$ *which is used to say an object is not in a set.*

Definition 2.3 *We say two sets are* **equal** *if they have exactly the same members.*

Example 2.1 *If*

$$S = \{1, 2, 3\}$$

then $3 \in S$ *and* $4 \notin S$*. The set membership symbol is often used in defining operations that manipulate sets. The set*

$$T = \{2, 3, 1\}$$

is equal to S *because they have the same members: 1, 2, and 3. While we usually list the members of a set in a "standard" order (if one is available) there is no requirement to do so and sets are indifferent to the order in which their members are listed.*

Definition 2.4 *The* **cardinality** *of a set is its* size*. For a finite set, the cardinality of a set is the number of members it contains. In symbolic notation the size of a set* S *is written* $|S|$*. We will deal with the idea of the cardinality of an infinite set later.*

Example 2.2 Set cardinality
For the set $S = \{1, 2, 3\}$ *we show cardinality by writing* $|S| = 3$

We now move on to a number of *operations* on sets. You are already familiar with several operations on numbers such as addition, multiplication, and negation.

Definition 2.5 *The* **intersection** *of two sets* S *and* T *is the collection of all objects that are in both sets. It is written* $S \cap T$*. Using curly brace notation*

$$S \cap T = \{x : (x \in S) \text{ and } (x \in T)\}$$

The symbol *and* in the above definition is an example of a Boolean or logical operation. It is only true when both the propositions it joins are also true. It has a symbolic equivalent $\wedge$. This lets us write the formal definition of intersection more compactly:

$$S \cap T = \{x : (x \in S) \wedge (x \in T)\}$$

Example 2.3 Intersections of sets
Suppose $S = \{1, 2, 3, 5\}$,
$T = \{1, 3, 4, 5\}$*, and* $U = \{2, 3, 4, 5\}$.
Then:

$S \cap T = \{1, 3, 5\}$,

$S \cap U = \{2, 3, 5\}$*, and*

$T \cap U = \{3, 4, 5\}$

Definition 2.6 *If* A *and* B *are sets and* $A \cap B = \emptyset$ *then we say that* A *and* B *are* disjoint*, or* disjoint sets*.*

Definition 2.7 *The* **union** *of two sets* S *and* T *is the collection of all objects that are in either set. It is written* $S \cup T$*. Using curly brace notion*

$$S \cup T = \{x : (x \in S) \text{ or } (x \in T)\}$$

The symbol *or* is another Boolean operation, one that is true if either of the propositions it joins are true. Its symbolic equivalent is $\vee$ which lets us re-write the definition of union as:

$$S \cup T = \{x : (x \in S) \vee (x \in T)\}$$

Example 2.4 Unions of sets.
Suppose $S = \{1, 2, 3\}$, $T = \{1, 3, 5\}$, *and* $U = \{2, 3, 4, 5\}$.
Then:

$S \cup T = \{1, 2, 3, 5\}$,

$S \cup U = \{1, 2, 3, 4, 5\}$*, and*

$T \cup U = \{1, 2, 3, 4, 5\}$

When performing set theoretic computations, you should declare the domain in which you are working. In set theory this is done by declaring a universal set.

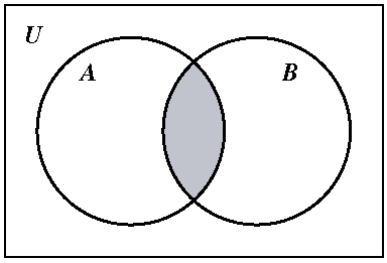
Definition 2.8 *The* **universal set***, at least for a given collection of set theoretic computations, is the set of all possible objects.*

If we declare our universal set to be the integers then $\{\frac{1}{2}, \frac{2}{3}\}$ is not a well defined set because the objects used to define it are not members of the universal set. The symbols $\{\frac{1}{2}, \frac{2}{3}\}$ do define a set if a universal set that includes $\frac{1}{2}$ and $\frac{2}{3}$ is chosen. The problem arises from the fact that neither of these numbers are integers. The universal set is commonly written $\mathcal{U}$. Now that we have the idea of declaring a universal set we can define another operation on sets.

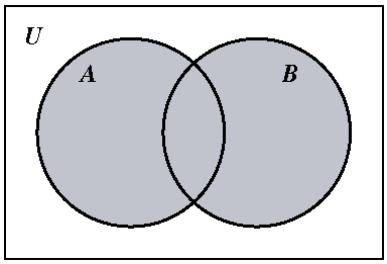
2.1.1 Venn Diagrams

A Venn diagram is a way of depicting the relationship between sets. Each set is shown as a circle and circles overlap if the sets intersect.

Example 2.5 *The following are Venn diagrams for the intersection and union of two sets. The shaded parts of the diagrams are the intersections and unions respectively.*



$A \cap B$



$A \cup B$

Notice that the rectangle containing the diagram is labeled with a U representing the universal set.

Definition 2.9 *The* **compliment** *of a set S is the collection of objects in the universal set that are not in S. The compliment is written S^c. In curly brace notation*

$$S^c = \{x : (x \in \mathcal{U}) \wedge (x \notin S)\}$$

or more compactly as

$$S^c = \{x : x \notin S\}$$

however it should be apparent that the compliment of a set always depends on which universal set is chosen.

There is also a Boolean symbol associated with the complementation operation: the *not* operation. The notation for not is $\neg$. There is not much savings in space as the definition of compliment becomes

$$S^c = \{x : \neg(x \in S)\}$$

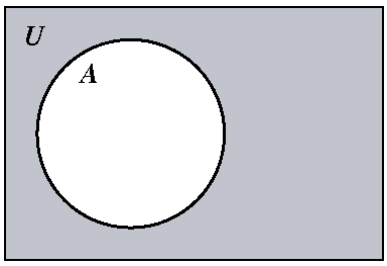
Example 2.6 Set Compliments

(i) Let the universal set be the integers. Then the compliment of the even integers is the odd integers.

(ii) Let the universal set be $\{1, 2, 3, 4, 5\}$, then the compliment of $S = \{1, 2, 3\}$ is $S^c = \{4, 5\}$ while the compliment of $T = \{1, 3, 5\}$ is $T^c = \{2, 4\}$.

(iii) Let the universal set be the letters $\{a, e, i, o, u, y\}$. Then $\{y\}^c = \{a, e, i, o, u\}$.

The Venn diagram for A^c is



A^c

We now have enough set-theory operators to use them to define more operators quickly. We will continue to give English and symbolic definitions.

Definition 2.10 *The* **difference** *of two sets S and T is the collection of objects in S that are not in T. The difference is written $S - T$. In curly brace notation*

$$S - T = \{x : x \in (S \cap (T^c))\},$$

or alternately

$$S - T = \{x : (x \in S) \wedge (x \notin T)\}$$

Notice how intersection and complementation can be used together to create the difference operation and that the definition can be rephrased to use Boolean operations. There is a set of rules that reduces the number of parenthesis required. These are called **operator precedence rules**.

(i) Other things being equal, operations are performed left-to-right.

(ii) Operations between parenthesis are done first, starting with the innermost of nested parenthesis.

(iii) All complimentations are computed next.

(iv) All intersections are done next.

(v) All unions are performed next.

(vi) Tests of set membership and computations, equality or inequality are performed last.

Special operations like the set difference or the symmetric difference, defined below, are not included in the precedence rules and thus always use parenthesis.

Example 2.7 Operator precedence
Since complementation is done before intersection the symbolic definition of the difference of sets can be rewritten:

$$S - T = \{x : x \in S \cap T^c\}$$

If we were to take the set operations

$$A \cup B \cap C^c$$

and put in the parenthesis we would get

$$(A \cup (B \cap (C^c)))$$

Definition 2.11 *The* **symmetric difference** *of two sets S and T is the set of objects that are in one and only one of the sets. The symmetric difference is written $S\Delta T$. In curly brace notation:*

$$S\Delta T = \{(S - T) \cup (T - S)\}$$

Example 2.8 Symmetric differences
Let S be the set of non-negative multiples of two that are no more than twenty four. Let T be the non-negative multiples of three that are no more than twenty four. Then

$$S\Delta T = \{2, 3, 4, 8, 9, 10, 14, 15, 16, 20, 21, 22\}$$

Another way to think about this is that we need numbers that are positive multiples of 2 or 3 (but not both) that are no more than 24.

Another important tool for working with sets is the ability to compare them. We have already defined what it means for two sets to be equal, and so by implication what it means for them to be unequal. We now define another comparator for sets.

Definition 2.12 *For two sets S and T we say that S* **is a subset of** *T if each element of S is also an element of T. In formal notation $S \subseteq T$ if for all $x \in S$ we have $x \in T$.*

If $S \subseteq T$ then we also say T contains S which can be written $T \supseteq S$. If $S \subseteq T$ and $S \neq T$ then we write $S \subset T$ and we say S is a *proper* subset of T.

Example 2.9 Subsets
If $A = \{a, b, c\}$ then A has eight different subsets:

$$\emptyset \qquad \{a\} \qquad \{b\} \qquad \{c\}$$

$$\{a, b\} \quad \{a, c\} \quad \{b, c\} \quad \{a, b, c\}$$

Notice that $A \subseteq A$ and in fact each set is a subset of itself. The empty set $\emptyset$ is a subset of every set.

We are now ready to prove our first proposition. Some new notation is required and we must introduce an important piece of mathematical culture. If we say "A if and only if B" then we mean that either A and B are both true or they are both false in any given circumstance. For example: "an integer x is even if and only if it is a multiple of 2". The phrase "if and only if" is used to establish *logical equivalence*. Mathematically, "A if and only if B" is a way of stating that A and B are simply different ways of saying the same thing. The phrase "if and only if" is abbreviated iff and is represented symbolically as the double arrow $\Leftrightarrow$. Proving an iff statement is done by independently demonstrating that each may be deduced from the other.

Proposition 2.1 *Two sets are equal if and only if each is a subset of the other. In symbolic notation:*

$$(A = B) \Leftrightarrow (A \subseteq B) \wedge (B \subseteq A)$$

Proof:

Let the two sets in question be A and B. Begin by assuming that $A = B$. We know that every set is

a subset of itself so $A \subseteq A$. Since $A = B$ we may substitute into this expression on the left and obtain $B \subseteq A$. Similarly we may substitute on the right and obtain $A \subseteq B$. We have thus demonstrated that if $A = B$ then A and B are both subsets of each other, giving us the first half of the iff.

Assume now that $A \subseteq B$ and $B \subseteq A$. Then the definition of subset tells us that any element of A is an element of B. Similarly any element of B is an element of A. This means that A and B have the same elements which satisfies the definition of set equality. We deduce $A = B$ and we have the second half of the iff. □

A note on mathematical grammar: the symbol □ indicates the end of a proof. On a paper turned in by a student it is usually taken to mean "I think the proof ends here". Any proof should have a □ to indicate its end. The student should also note the lack of calculations in the above proof. If a proof cannot be read back in (sometimes overly formal) English then it is probably incorrect. Mathematical symbols should be used for the sake of brevity or clarity, not to obscure meaning.

Proposition 2.2 De Morgan's Laws *Suppose that S and T are sets. DeMorgan's Laws state that*

(i) $(S \cup T)^c = S^c \cap T^c$, *and*

(ii) $(S \cap T)^c = S^c \cup T^c$.

Proof:

Let $x \in (S \cup T)^c$; then x is not a member of S or T. Since x is not a member of S we see that $x \in S^c$. Similarly $x \in T^c$. Since x is a member of both these sets we see that $x \in S^c \cap T^c$ and we see that $(S \cup T)^c \subseteq S^c \cap T^c$. Let $y \in S^c \cap T^c$. Then the definition of intersection tells us that $y \in S^c$ and $y \in T^c$. This in turn lets us deduce that y is not a member of $S \cup T$, since it is not in either set, and so we see that $y \in (S \cup T)^c$. This demonstrates that $S^c \cap T^c \subseteq (S \cup T)^c$. Applying Proposition 2.1 we get that $(S \cup T)^c = S^c \cap T^c$ and we have proven part (i). The proof of part (ii) is left as an exercise. □

In order to prove a mathematical statement you must prove it is always true. In order to disprove a mathematical statement you need only find a single instance where it is false. It is thus possible for a false mathematical statement to be "true most of the time". In the next chapter we will develop the theory of prime numbers. For now we will assume the reader has a modest familiarity with the primes. The statement "Prime numbers are odd" is false once, because 2 is a prime number. All the other prime numbers are odd. The statement is a false one. This very strict definition of what makes a statement true is a convention in mathematics. We call 2 a *counter example*. It is thus necessary to find only one counter-example to demonstrate a statement is (mathematically) false.

Example 2.10 Disproof by counter example *Prove that the statement $A \cup B = A \cap B$ is false.*

Let $A = \{1, 2\}$ and $B = \{3, 4\}$. Then $A \cap B = \emptyset$ while $A \cup B = \{1, 2, 3, 4\}$. The sets A and B form a counter-example to the statement.

Problems

Problem 2.1 *Which of the following are sets? Assume that a proper universal set has been chosen and answer by listing the names of the collections of objects that are sets. Warning: at least one of these items has an answer that, while likely, is not 100% certain.*

(i) $A = \{2, 3, 5, 7, 11, 13, 19\}$

(ii) $B = \{A, E, I, O, U\}$

(iii) $C = \{\sqrt{x} : x < 0\}$

(iv) $D = \{1, 2, A, 5, B, Q, 1, V\}$

(v) E is the list of first names of people in the 1972 phone book in Lawrence Kansas in the order they appear in the book. There were more than 35,000 people in Lawrence that year.

(vi) F is a list of the weight, to the nearest kilogram, of all people that were in Canada at any time in 2007.

(vii) G is a list of all weights, to the nearest kilogram, that at least one person in Canada had in 2007.

Problem 2.2 *Suppose that we have the set* $U = \{n : 0 \leq n < 100\}$ *of whole numbers as our universal set. Let* P *be the prime numbers in* U*, let* E *be the even numbers in* U*, and let* $F = \{1, 2, 3, 5, 8, 13, 21, 34, 55, 89\}$*. Describe the following sets either by listing them or with a careful English sentence.*

(i) E^c,

(ii) $P \cap F$,

(iii) $P \cap E$,

(iv) $F \cap E \cup F \cap E^c$, and

(v) $F \cup F^c$.

Problem 2.3 *Suppose that we take the universal set* $\mathcal{U}$ *to be the integers. Let* S *be the even integers, let* T *be the integers that can be obtained by tripling any one integer and adding one to it, and let* V *be the set of numbers that are whole multiples of both two and three.*

(i) Write S, T, and V using symbolic notation.

(ii) Compute $S \cap T$, $S \cap V$ and $T \cap V$ and give symbolic representations that do not use the symbols S, T, or V on the right hand side of the equals sign.

Problem 2.4 *Compute the cardinality of the following sets. You may use other texts or the internet.*

(i) Two digit positive odd integers.

(ii) Elements present in a sucrose molecule.

(iii) Isotopes of hydrogen that are not radioactive.

(iv) Planets orbiting the same star as the planet you are standing on that have moons. Assume that Pluto is a minor planet.

(v) Elements with seven electrons in their valence shell. Remember that Ununoctium was discovered in 2002 so be sure to use a relatively recent reference.

(vi) Subsets of $S = \{a, b, c, d\}$ with cardinality 2.

(vii) Prime numbers whose base-ten digits sum to ten. Be careful, some have three digits.

Problem 2.5 *Find an example of an infinite set that has a finite complement, be sure to state the universal set.*

Problem 2.6 *Find an example of an infinite set that has an infinite complement, be sure to state the universal set.*

Problem 2.7 *Add parenthesis to each of the following expressions that enforce the operator precedence rules as in Example 2.7. Notice that the first three describe sets while the last returns a logical value (true of false).*

(i) $A \cup B \cup C \cup D$

(ii) $A \cup B \cap C \cup D$

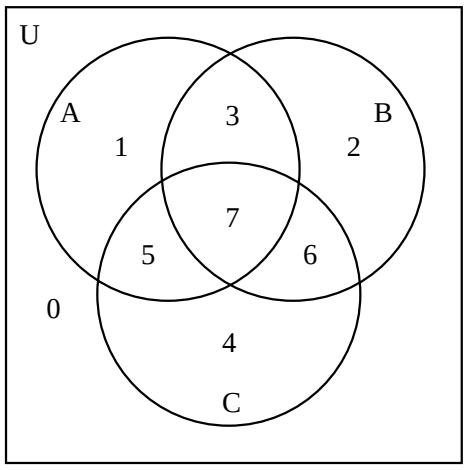
(iii) $A^c \cap B^c \cup C$

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

Problem 2.8 *Give the Venn diagrams for the following sets.*

(i) $A - B$ (ii) $B - A$ (iii) $A^c \cap B$

(iv) $A \Delta B$ (v) $(A \Delta B)^c$ (vi) $A^c \cup B^c$



Problem 2.9 *Examine the Venn diagram above. Notice that every combination of sets has a unique number in common. Construct a similar collection of four sets.*

Problem 2.10 *Read Problem 2.9. Can a system of sets of this sort be constructed for any number of sets? Explain your reasoning.*

Problem 2.11 *Suppose we take the universal set to be the set of non-negative integers. Let E be the set of even numbers, O be the set of odd numbers and $F = \{0, 1, 2, 3, 5, 8, 13, 21, 34, 89, 144, ...\}$ be the set of Fibonacci numbers. The Fibonacci sequence is $0, 1, 1, 2, 3, 5, 8, \ldots$ in which the next term is obtained by adding the previous two.*

(i) Prove that the intersection of F with E and O are both infinite.

(ii) Make a Venn diagram for the sets E, F, and O, and explain why this is a Mickey-Mouse problem.

Problem 2.12 *A binary operation $\odot$ is* **commutative** *if $x \odot y = y \odot x$. An example of a commutative operation is multiplication. Subtraction is non-commutative. Determine, with proof, if union, intersection, set difference, and symmetric difference are commutative.*

Problem 2.13 *An* **identity** *for an operation $\odot$ is an object i so that, for all objects x, $i \odot x = x \odot i = x$. Find, with proof, identities for the operations* set union *and* set intersection.

Problem 2.14 *Prove part (ii) of Proposition 2.2.*

Problem 2.15 *Prove that*

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

Problem 2.16 *Prove that*

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

Problem 2.17 *Prove that*

$$A\Delta(B\Delta C) = (A\Delta B)\Delta C$$

Problem 2.18 *Disprove that*

$$A\Delta(B \cup C) = (A\Delta B) \cup C$$

Problem 2.19 *Consider the set $S = \{1, 2, 3, 4\}$. For each $k = 0, 1, \ldots, 4$ how many k element subsets does S have?*

Problem 2.20 *Suppose we have a set S with $n \geq 0$ elements. Find a formula for the number of different subsets of S that have k elements.*

Problem 2.21 *For finite sets S and T, prove*

$$|S \cup T| = |S| + |T| - |S \cap T|$$

2.2 Mathematical Induction

Mathematical induction is a technique used in proving mathematical assertions. The basic idea of induction is that we prove that a statement is true in one case and then also prove that if it is true in a given case it is true in the next case. This then permits the cases for which the statement is true to cascade from the initial true case. We will start with the mathematical foundations of induction.

We assume that the reader is familiar with the symbols $<$, $>$, $\leq$ and $\geq$. From this point on we will denote the set of integers by the symbol $\mathbb{Z}$. The non-negative integers are called the *natural numbers*. The symbol for the set of natural numbers is $\mathbb{N}$. Any mathematical system rests on a foundation of axioms. Axioms are things that we simply assume to be true. We will assume the truth of the following principle, adopting it as an axiom.

The well-ordering principle: Every non-empty set of natural numbers contains a smallest element.

The well ordering principle is an axiom that agrees with the common sense of most people familiar with the natural numbers. An empty set does not contain a smallest member because it contains no members at all. As soon as we have a set of natural numbers with some members then we can order those members in the usual fashion. Having ordered them, one will be smallest. This intuition agreeing with this latter claim depends strongly on the fact the integers are "whole numbers" spaced out in increments of one. To see why this is important consider the smallest positive distance. If such a distance existed, we could cut it in half to obtain a smaller distance - the quantity contradicts its own existence. The well-ordering principle can be used to prove the correctness of induction.

Theorem 2.1 Mathematical Induction I *Suppose that $P(n)$ is a proposition that it either true or false for any given natural numbers n. If*

(i) $P(0)$ is true and,

(ii) when $P(n)$ is true so is $P(n+1)$

Then we may deduce that $P(n)$ is true for any natural number.

Proof:

Assume that (i) and (ii) are both true statements. Let S be the set of all natural numbers for which $P(n)$ is false. If S is empty then we are done, so assume that S is not empty. Then, by the well ordering principle, S has a least member m. By (i) above $m \neq 0$ and so $m-1$ is a natural number. Since m is the smallest member of S it follows that $P(m-1)$ is true. But this means, by (ii) above, that $P(m)$ is true. We have a contradiction and so our assumption that $S \neq \emptyset$ must be wrong. We deduce S is empty and that as a consequence $P(n)$ is true for all $n \in \mathbb{N}$. □

The technique used in the above proof is called *proof by contradiction*. We start by assuming the logical opposite of what we want to prove, in this case that there is some m for which $P(m)$ is false, and from that assumption we derive an impossibility. If an assumption can be used to demonstrate an impossibility then it is false and its logical opposite is true.

A nice problem on which to demonstrate mathematical induction is counting how many subsets a finite set has.

Proposition 2.3 Subset counting. *A set S with n elements has 2^n subsets.*

Proof:

First we check that the proposition is true when $n = 0$. The empty set has exactly one subset: itself. Since $2^0 = 1$ the proposition is true for $n = 0$. We now assume the proposition is true for some n. Suppose that S is a set with $n+1$ members and that $x \in S$. Then $S-\{x\}$ (the set difference of S and a set $\{x\}$ containing only x) is a set of n elements and so, by the assumption, has 2^n subsets. Now every subset of S either contains x or it fails to. Every subset of S that does not contain x is a subset of $S - \{x\}$ and so there are 2^n such subsets of S. Every subset of S that contains x may be obtained in exactly one way from one that does not by taking the union with $\{x\}$. This means that the number of subsets of S containing or failing to contain x are equal. This means there are 2^n subsets of S containing x. The total number of subsets of S is thus $2^n + 2^n = 2^{n+1}$. So if we assume the proposition is true for n we can demonstrate that it is also true for $n + 1$. It follows by mathematical induction that the proposition is true for all natural numbers. □

The set of all subsets of a given set is itself an important object and so has a name.

Definition 2.13 *The set of all subsets of a set S is called the* **powerset** *of S. The notation for the powerset of S is $\mathcal{P}(S)$.*

This definition permits us to rephrase Proposition 2.3 as follows: the power set of a set of n elements has size 2^n.

Theorem 2.1 lets us prove propositions that are true on the natural numbers, starting at zero. A small modification of induction can be used to prove statements that are true only for those $n \geq k$ for any integer k. All that is needed is to use induction on a proposition $Q(n-k)$ where $Q(n-k)$ is logically equivalent to $P(n)$. If $Q(n-k)$ is true for $n-k \geq 0$ then $P(n)$ is true for $n \geq k$ and we have the modified induction. The practical difference is that we start with k instead of zero.

Example 2.11 *Prove that $n^2 \geq 2n$ for all $n \geq 2$.*

Notice that $2^2 = 4 = 2 \times 2$ so the proposition is true when $n = 2$. We next assume that $P(n)$ is true for some n and we compute:

$$\begin{aligned} n^2 &\geq 2n \\ n^2 + 2n + 1 &\geq 2n + 2n + 1 \\ (n+1)^2 &\geq 2n + 2n + 1 \\ (n+1)^2 &\geq 2n + 1 + 1 \\ (n+1)^2 &\geq 2n + 2 \\ (n+1)^2 &\geq 2(n+1) \end{aligned}$$

To move from the third step to the fourth step we use the fact that $2n > 1$ when $n \geq 2$. The last step is $P(n+1)$, which means we have deduced $P(n+1)$ from $P(n)$. Using the modified form of induction we have proved that $n^2 \geq 2n$ for all $n \geq 2$.

It is possible to formalize the procedure for using mathematical induction into a three-part process. Once we have a proposition $P(n)$,

(i) First demonstrate a *base case* by directly demonstrating $P(k)$,

(ii) Next make the *induction hypothesis* that $P(n)$ is true for some n,

(iii) Finally, starting with the assumption that $P(n)$ is true, demonstrate $P(n+1)$.

These steps permit us to deduce that $P(n)$ is true for all $n \geq k$.

Example 2.12 *Using induction, prove*

$$1+2+\cdots+n = \frac{1}{2}n(n+1)$$

In this case $P(n)$ is the statement

$$1+2+\cdots+n = \frac{1}{2}n(n+1)$$

Base case: $1 = \frac{1}{2}1(1+1)$, *so $P(1)$ is true.* **Induction hypothesis:** *for some n,*

$$1+2+\cdots+n = \frac{1}{2}n(n+1)$$

Compute:

$$\begin{aligned} 1+2+\cdots+(n+1) &= 1+2+\cdots+n+(n+1) \\ &= \frac{1}{2}n(n+1)+(n+1) \\ &= \frac{1}{2}\left(n(n+1)+2(n+1)\right) \\ &= \frac{1}{2}\left(n^2+3n+2\right) \\ &= \frac{1}{2}(n+1)(n+2) \\ &= \frac{1}{2}(n+1)((n+1)+1) \end{aligned}$$

and so we have shown that if $P(n)$ is true then so is $P(n+1)$. We have thus proven that $P(n)$ is true for all $n \geq 1$ by mathematical induction.

We now introduce *sigma notation* which makes problems like the one worked in Example 2.12 easier to state and manipulate. The symbol $\sum$ is used to add up lists of numbers. If we wished to sum some formula $f(i)$ over a range from a to b, that is to say $a \leq i \leq b$, then we write :

$$\sum_{i=a}^{b} f(i)$$

On the other hand if S is a set of numbers and we want to add up $f(s)$ for all $s \in S$ we write:

$$\sum_{s \in S} f(s)$$

The result proved in Example 2.12 may be stated in the following form using sigma notation.

$$\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$$

Proposition 2.4 *Suppose that c is a constant and that $f(i)$ and $g(i)$ are formulas. Then*

(i) $\sum_{i=a}^{b}(f(i)+g(i)) = \sum_{i=a}^{b} f(i) + \sum_{i=a}^{b} g(i)$

(ii) $\sum_{i=a}^{b}(f(i)-g(i)) = \sum_{i=a}^{b} f(i) - \sum_{i=a}^{b} g(i)$

(iii) $\sum_{i=a}^{b} c \cdot f(i) = c \cdot \sum_{i=a}^{b} f(i)$.

Proof:

Part (i) and (ii) are both simply the associative law for addition: $a+(b+c) = (a+b)+c$ applied many times. Part (iii) is a similar multiple application of the distributive law $ca+cb = c(a+b)$. □

The sigma notation lets us work with indefinitely long (and even infinite) sums. There are other similar notations. If $A_1, A_2, \ldots, A_n$ are sets then the intersection or union of all these sets can be written:

$$\bigcap_{i=1}^{n} A_i$$

$$\bigcup_{i=1}^{n} A_i$$

Similarly if $f(i)$ is a formula on the integers then

$$\prod_{i=1}^{n} f(i)$$

is the notation for computing the product $f(1) \cdot f(2) \cdot \cdots \cdot f(n)$. This notation is called **Pi** notation.

Definition 2.14 *When we solve an expression involving $\sum$ to obtain a formula that does not use $\sum$ or "..." as in Example 2.12 then we say we have found a* **closed form** *for the expression. Example 2.12 finds a closed form for $\sum_{i=1}^{n} i$.*

At this point we introduce a famous mathematical sequence in order to create an arena for practicing proofs by induction.

Definition 2.15 *The* **Fibonacci numbers** *are defined as follows. $f_1 = f_2 = 1$ and, for $n \geq 3$, $f_n = f_{n-1} + f_{n-2}$.*

Example 2.13 *The Fibonacci numbers with four or fewer digits are: $f_1 = 1$, $f_2 = 1$, $f_3 = 2$, $f_4 = 3$, $f_5 = 5$, $f_6 = 8$, $f_7 = 13$, $f_8 = 21$, $f_9 = 34$, $f_{10} = 55$, $f_{11} = 89$, $f_{12} = 144$, $f_{13} = 233$, $f_{14} = 377$, $f_{15} = 610$, $f_{16} = 987$, $f_{17} = 1597$, $f_{18} = 2584$, $f_{19} = 4181$, and $f_{20} = 6765$.*

Example 2.14 *Prove that the Fibonacci number f_{3n} is even.*

Solution:

Notice that $f_3 = 2$ and so the proposition is true when $n = 1$. Assume that the proposition is true for some $n \geq 1$. Then:

$$\begin{aligned} f_{3(n+1)} &= f_{3n+3} && (2.1)\\ &= f_{3n+2} + f_{3n+1} && (2.2)\\ &= f_{3n+1} + f_{3n} + f_{3n+1} && (2.3)\\ &= 2 \cdot f_{3n+1} + f_{3n} && (2.4) \end{aligned}$$

but this suffices because f_{3n} is even by the induction hypothesis while $2 \cdot f_{3n+1}$ is also even. The sum is thus even and so $f_{3(n+1)}$ is even. If follows by induction that f_{3n} is even for all n. □

Problems

Problem 2.22 *Suppose that $S = \{a, b, c\}$. Compute and list explicitly the members of the powerset, $\mathcal{P}(S)$.*

Problem 2.23 *Prove that for a finite set X that*

$$|X| \leq |\mathcal{P}(X)|$$

Problem 2.24 *Suppose that $X \subseteq Y$ with $|Y| = n$ and $|X| = m$. Compute the number of subsets of Y that contain X.*

Problem 2.25 *Compute the following sums.*

(i) $\sum_{i=1}^{20} i$,

(ii) $\sum_{i=10}^{30} i$, and

(iii) $\sum_{i=-20}^{21} i$.

Problem 2.26 *Using mathematical induction, prove the following formulas.*

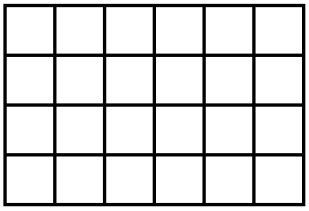
(i) $\sum_{i=1}^{n} 1 = n$,

(ii) $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$, and

(iii) $\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}$.

Problem 2.27 *If $f(i)$ and $g(i)$ are formulas and c and d are constants prove that*

$$\sum_{i=a}^{b}(c \cdot f(i) + d \cdot g(i)) = c \cdot \sum_{i=a}^{b} f(i) + d \cdot \sum_{i=a}^{b} g(i)$$



Problem 2.28 *Suppose you want to break an $n \times m$ chocolate bar, like the 6×4 example shown above, into pieces corresponding to the small squares shown. What is the minimum number of breaks you can make? Prove your answer is correct.*

Problem 2.29 *Prove by induction that the sum of the first n odd numbers equals n^2.*

Problem 2.30 *Compute the sum of the first n positive even numbers.*

Problem 2.31 *Find a closed form for*

$$\sum_{i=1}^{n} i^2 + 3i + 5$$

Problem 2.32 *Let $f(n,3)$ be the number of subsets of $\{1,2,\ldots,n\}$ of size 3. Using induction, prove that $f(n,3) = \frac{1}{6}n(n-1)(n-2)$.*

Problem 2.33 *Suppose that we have sets $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ such that $X_i \subseteq Y_i$. Prove that the intersection of all the X_i is a subset of the intersection of all the Y_i:*

$$\bigcap_{i=1}^{n} X_i \subseteq \bigcap_{i=1}^{n} Y_i$$

Problem 2.34 *Suppose that $S_1, S_2, \ldots S_n$ are sets. Prove the following generalization of DeMorgan's laws:*

(i) $\left(\bigcap_{i=1}^{n} S_i\right)^c = \bigcup_{i=1}^{n} S_i^c$, *and*

(ii) $\left(\bigcup_{i=1}^{n} S_i\right)^c = \bigcap_{i=1}^{n} S_i^c$.

Problem 2.35 *Prove by induction that the Fibonacci number f_{4n} is a multiple of 3.*

Problem 2.36 *Prove that if r is a real number $r \neq 1$ and $r \neq 0$ then*

$$\sum_{i=0}^{n} r^i = \frac{1 - r^{n+1}}{1 - r}$$

Problem 2.37 *Prove by induction that the Fibonacci number f_{5n} is a multiple of 5.*

Problem 2.38 *Prove by induction that the Fibonacci number f_n has the value*

$$f_n = \frac{\sqrt{5}}{5} \cdot \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{\sqrt{5}}{5} \cdot \left(\frac{1-\sqrt{5}}{2}\right)^n$$

Problem 2.39 *Prove that for sufficiently large n the Fibonacci number f_n is the integer closest to*

$$\frac{\sqrt{5}}{5}\left(\frac{1+\sqrt{5}}{2}\right)^n$$

and compute the exact value of f_{30}. Show your work (i.e. don't look the result up on the net).

Problem 2.40 *Prove that $\frac{n(n-1)(n-2)(n-3)}{24}$ is a whole number for any whole number n.*

Problem 2.41 *Consider the statement "All cars are the same color." and the following "proof".*

Proof:

We will prove for $n \geq 1$ that for any set of n cars all the cars in the set have the same color.

- *Base Case: n=1 If there is only one car then clearly there is only one color the car can be.*
- *Inductive Hypothesis: Assume that for any set of n cars there is only one color.*
- *Inductive step: Look at any set of $n+1$ cars. Number them: $1, 2, 3, \ldots, n, n+1$. Consider the sets $\{1, 2, 3, \ldots, n\}$ and $\{2, 3, 4, ..., n+1\}$. Each is a set of only n cars, therefore for each set there is only one color. But the n^{th} car is in both sets so the color of the cars in the first set must be the same as the color of the cars in the second set. Therefore there must be only one color among all $n+1$ cars.*
- *The proof follows by induction.* □

What are the problems with this proof?

2.3 Functions

In this section we will define functions and extend much of our ability to work with sets to infinite sets. There are a number of different types of functions and so this section contains a great deal of terminology.

Recall that two finite sets are the same size if they contain the same number of elements. It is possible to make this idea formal by using functions and, once the notion is formally defined, it can be applied to infinite sets.

Definition 2.16 *An* **ordered pair** *is a collection of two elements with the added property that one element comes first and one element comes second. The set containing only x and y (for $x \neq y$) is written $\{x, y\}$. The ordered pair containing x and y with x first is written (x, y). Notice that while $\{x, x\}$ is not a well defined set, (x, x) is a well defined ordered pair because the two copies of x are different by virtue of coming first and second.*

The reason for defining ordered pairs at this point is that it permits us to make an important formal definition that pervades the rest of mathematics.

Definition 2.17 *A* **function** f *with* **domain** S *and* **range** T *is a set of ordered pairs* (s,t) *with first element from* S *and second element from* T *that has the property that every element of* S *appears exactly once as the first element in some ordered pair. We write* $f : S \to T$ *for such a function.*

Example 2.15 *Suppose that* $A = \{a, b, c\}$ *and* $B = \{0, 1\}$ *then*

$$f = \{(a, 0), (b, 1), (c, 0)\}$$

is a function from A *to* B*. The function* $f : A \to B$ *can also be specified by saying* $f(a) = 0$*,* $f(b) = 1$ *and* $f(c) = 0$*.*

The set of ordered pairs $\{(a, 0), (b, 1)\}$ *is not a function from* A *to* B *because* c *is not the first coordinate of any ordered pair. The set of ordered pairs* $\{(a, 0), (a, 1), (b, 0), (c, 0)\}$ *is not a function from* A *to* B *because* a *appears as the first coordinate of two different ordered pairs.*

In calculus you may have learned the *vertical line rule* that states that the graph of a function may not intersect a vertical line at more than one point. This corresponds to requiring that each point in the domain of the function appear in only one ordered pair. In set theory, all functions are required to state their domain and range when they are defined. In calculus functions had a domain that was a subset of the real numbers and you were sometimes required to identify the subset.

Example 2.16 *This example contrasts the way functions were treated in a typical calculus course with the way we treat them in set theory.*

Calculus: *find the domain of the function*

$$f(x) = \sqrt{x}$$

Since we know that the square root function exists only for non-negative real numbers the domain is $\{x : x \geq 0\}$*.*

Set theory: *the function* $f = \sqrt{x}$ *from the non-negative real numbers to the real numbers is the set of ordered pairs* $\{(r^2, r) : r \geq 0\}$*. This function is well defined because each non-negative real number is the square of some positive real number.*

The major contrasts between functions in calculus and functions in set theory are:

(i) The domain of functions in calculus are often specified only by implication (you have to know how all the functions used work) and are almost always a subset of the real numbers. The domain in set theory must be explicitly specified and may be any set at all.

(ii) Functions in calculus typically had graphs that you could draw and look at. Geometric intuition driven by the graphs plays a major role in our understanding of functions. Functions in set theory are seldom graphed and often don't have a graph.

A point of similarity between calculus and set theory is that the range of the function is not explicitly specified. When we have a function $f : S \to T$ then the range of f is a subset of T.

Definition 2.18 *If* f *is a function then we denote the domain of* f *by* $dom(f)$ *and the range of* f *by* $rng(f)$

Example 2.17 *Suppose that* $f(n) : \mathbb{N} \to \mathbb{N}$ *is defined by* $f(n) = 2n$*. Then the domain and range of* f *are the integers:* $dom(f) = rng(f) = \mathbb{N}$*. If we specify the ordered pairs of* f *we get*

$$f = \{(n, 2n) : n \in \mathbb{N}\}$$

There are actually two definitions of range that are used in mathematics. The definition we are using, the set from which second coordinates of ordered pairs in a function are drawn, is also the definition typically using in computer science. The other definition is the set of second coordinates that actually appear in ordered pairs. This set, which we will define formally later, is the *image* of the function. To make matters even worse the set we are calling the range of a function is also called the *co-domain*. We include these confusing terminological notes for students that may try and look up supplemental material.

Definition 2.19 *Let X, Y, and Z be sets. The* **composition** *of two functions $f : X \to Y$ and $g : Y \to Z$ is a function $h : X \to Z$ for which $h(x) = g(f(x))$ for all $x \in X$. We write $g \circ f$ for the composition of g with f.*

The definition of the composition of two functions requires a little checking to make sure it makes sense. Since *every* point must appear as a first coordinate of an ordered pair in a function, every result of applying f to an element of X is an element of Y to which g can be applied. This means that h is a well-defined set of ordered pairs. Notice that the order of composition is important - if the sets X, Y, and Z are distinct there is only one order in which composition even makes sense.

Example 2.18 *Suppose that $f : \mathbb{N} \to \mathbb{N}$ is given by $f(n) = 2n$ while $g : \mathbb{N} \to \mathbb{N}$ is given by $g(n) = n + 4$. Then*

$$(g \circ f)(n) = 2n + 4$$

while

$$(f \circ g)(n) = 2(n + 4) = 2n + 8$$

We now start a series of definitions that divide functions into a number of classes. We will arrive at a point where we can determine if the mapping of a function is reversible, if there is a function that exactly reverses the action of a given function.

Definition 2.20 *A function $f : S \to T$ is* **injective** *or* **one-to-one** *if no element of T (no second coordinate) appears in more than one ordered pair. Such a function is called an* **injection**.

Example 2.19 *The function $f : \mathbb{N} \to \mathbb{N}$ given by $f(n) = 2n$ is an injection. The ordered pairs of f are $(n, 2n)$ and so any number that appears as a second coordinate does so once.*

The function $g : \mathbb{Z} \to \mathbb{Z}$ given by $g(n) = n^2$ is not an injection. To see this notice that g contains the ordered pairs $(1, 1)$ and $(-1, 1)$ so that 1 appears twice as the second coordinate of an ordered pair.

Definition 2.21 *A function $f : S \to T$ is* **surjective** *or* **onto** *if every element of T appears in an ordered pair. Surjective functions are called* **surjections**.

We use the symbol $\mathbb{R}$ for the real numbers. We also assume familiarity with interval notation for contiguous subsets of the reals. For real numbers $a \leq b$

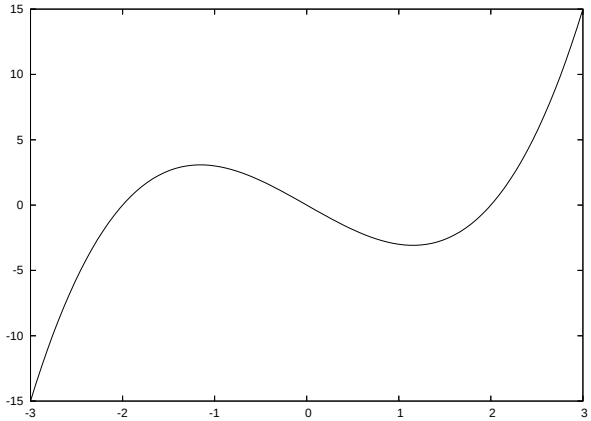
$$\begin{array}{lll} (a, b) & \text{is} & \{x : a < x < b\} \\ (a, b] & \text{is} & \{x : a < x \leq b\} \\ [a, b) & \text{is} & \{x : a \leq x < b\} \\ [a, b] & \text{is} & \{x : a \leq x \leq b\} \end{array}$$

Example 2.20 *The function $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(n) = 5 - n$ is a surjection. If we set $m = 5 - n$ then $n = 5 - m$. This means that if we want to find some n so that $f(n)$ is, for example, 8, then $5 - 8 = -3$ and we see that $f(-3) = 8$. This demonstrates that all m have some n so that $f(n) = m$, showing that all m appear as the second coordinate of an ordered pair in f.*

The function $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = \frac{x^2}{1+x^2}$ is not a surjection because $-1 < g(x) < 1$ for all $x \in \mathbb{R}$.

Definition 2.22 *A function that is both surjective and injective is said to be* **bijective**. *Bijective functions are called* **bijections**.

Example 2.21 *The function $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(n) = n$ is a bijection. All of its ordered pairs have the same first and second coordinate. This function is called the* identity function.



The function $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = x^3 - 4x$ is not a bijection. It is not too hard to show that it is a surjection, but it fails to be an injection. The portion of the graph shown above demonstrates that $g(x)$ takes on the same value more than once. This means that

some numbers appear twice as second coordinates of ordered pairs in g. We can use the graph because g is a function from the real numbers to the real numbers.

For a function $f : S \to T$ to be a bijection every element of S appears in an ordered pair as the first member of an ordered pair and every element of T appears in an ordered pair as the second member of an ordered pair. Another way to view a bijection is as a matching of the elements of S and T so that every element of S is paired with an element of T. For finite sets this is clearly only possible if the sets are the same size and, in fact, this is the formal definition of "same size" for sets.

Definition 2.23 *Two sets S and T are defined to be* **the same size** *or to have* **equal cardinality** *if there is a bijection $f : S \to T$.*

Example 2.22 *The sets $A = \{a, b, c\}$ and $Z = \{1, 2, 3\}$ are the same size. This is obvious because they have the same number of elements, $|A| = |Z| = 3$ but we can construct an explicit bijection*

$$f = \{(a, 3), (b, 1), (c, 2)\}$$

with each member of A appearing once as a first coordinate and each member of B appearing once as a second coordinate. This bijection is a witness *that A and B are the same size.*

Let E be the set of even integers. Then the function

$$g : \mathbb{Z} \to E$$

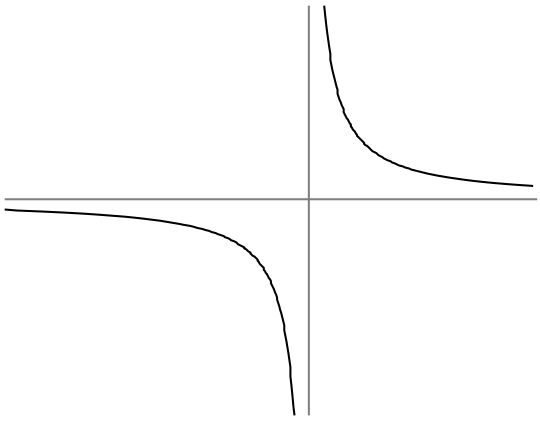
in which $g(n) = 2n$ is a bijection. Notice that each integers can be put into g and that each even integer has exactly one integer that can be doubled to make it. The existence of g is a witness that the set of integers and the set of even integers are the same size. This may seem a bit bizarre because the set $\mathbb{Z} - E$ is the infinite set of odd integers. In fact one hallmark of an infinite set is that it can be the same size as a proper subset. This also means we now have an equality set for sizes of infinite sets. We will do a good deal more with this in Chapter 3.

Bijections have another nice property: they can be unambiguously reversed.

Definition 2.24 *The* **inverse** *of a function $f : S \to T$ is a function $g : T \to S$ so that for all $x \in S$, $g(f(x)) = x$ and for all $y \in T$, $f(g(y)) = y$.*

If a function f has an inverse we use the notation f^{-1} for that inverse. Since an exponent of -1 also means reciprocal in some circumstances this can be a bit confusing. The notational confusion is resolved by considering context. So long as we keep firmly in mind that functions are sets of ordered pairs it is easy to prove the proposition/definition that follows after the next example.

Example 2.23 *If E is the set of even integers then the bijection $f(n) = 2n$ from $\mathbb{Z}$ to E has the inverse $f^{-1} : E \to \mathbb{Z}$ given by $g(2n) = n$. Notice that defining the rule for g as depending on the argument $2n$ seamlessly incorporates the fact that the domain of g is the even integers.*



If $g(x) = \frac{x}{x-1}$, shown above with its asymptotes $x = 1$ and $y = 1$ then f is a function from the set $H = \mathbb{R} - \{1\}$ to itself. The function was chosen to have asymptotes at equal x and y values; this is a bit unusual. The function g is a bijection. Notice that the graph intersects any horizontal or vertical line in at most one point. Every value except $x = 1$ may be put into g meaning that g is a function on H. Since the vertical asymptote goes off to ∞ in both directions, all values in H come out of g. This demonstrates g is a bijection. This means that it has an inverse which we now compute using a standard

technique from calculus classes.

$$\begin{aligned} y &= \frac{x}{x-1} \\ y(x-1) &= x \\ xy - y &= x \\ xy - x &= y \\ x(y-1) &= y \\ x &= \frac{y}{y-1} \end{aligned}$$

which tells us that $g^{-1}(x) = \frac{x}{x-1}$ *so* $g = g^{-1}$*: the function is its own inverse.*

Proposition 2.5 *A function has an inverse if and only if it is a bijection.*

Proof:

Suppose that $f : S \to T$ is a bijection. Then if $g : T \to S$ has ordered pairs that are the exact reverse of those given by f it is obvious that for all $x \in S$, $g(f(x)) = x$, likewise that for all $y \in T$, $f(g(y)) = y$. We have that bijections posses inverses. It remains to show that non-bijections do not have inverses.

If $f : S \to T$ is not a bijection then either it is not a surjection or it is not an injection. If f is not a surjection then there is some $t \in T$ that appears in no ordered pair of f. This means that no matter what $g(t)$ is, $f(g(t)) \neq t$ and we fail to have an inverse. If, on the other hand, $f : S \to T$ is a surjection but fails to be an injection then for some distinct $a, b \in S$ we have that $f(a) = t = f(b)$. For $g : T \to S$ to be an inverse of f we would need $g(t) = a$ and $g(t) = b$, forcing t to appear as the first coordinate of two ordered pairs in g and so rendering g a non-function. We thus have that non-bijections do not have inverses. □

The type of inverse we are discussing above is a *two-sided inverse*. The functions f and f^{-1} are mutually inverses of one another. It is possible to find a function that is a one-way inverse of a function so that $f(g(x)) = x$ but $g(f(x))$ is not even defined. These are called *one-sided inverses*.

Note on mathematical grammar: Recall that when two notions, such as "bijection" and "has an inverse" are equivalent we use the phrase "if and only if" (abbreviated iff) to phrase a proposition declaring that the notions are equivalent. A proposition that A iff B is proven by first assuming A and deducing B and then separately assuming B and deducing A. The formal symbol for A iff B is $A \Leftrightarrow B$. Likewise we have symbols for the ability to deduce B given A, $A \Rightarrow B$ and vice-versa $B \Rightarrow A$. These symbols are spoken "A implies B" and "B implies A" respectively.

Proposition 2.6 *Suppose that* X*,* Y*, and* Z *are sets. If* $f : X \to Y$ *and* $g : Y \to Z$ *are bijections then so is* $g \circ f : X \to Z$*.*

Proof: *this proof is left as an exercise.*

Definition 2.25 *Suppose that* $f : A \to B$ *is a function. The* **image of** A **in** B *is the subset of* B *made of elements that appear as the second element of ordered pairs in* f*. Colloquially the image of* f *is the set of elements of* B *hit by* f*. We use the notation* $Im(f)$ *for images. In other words* $Im(f) = \{f(a) : a \in A\}$*.*

Example 2.24 *If* $f : \mathbb{N} \to \mathbb{N}$ *is given by the rule* $f(n) = 3n$ *then the set* $T = \{0, 3, 6, \ldots\}$ *of natural numbers that are multiples of three is the image of* f*. Notation:* $Im(f) = T$*.*

If $g : \mathbb{R} \to \mathbb{R}$ *given by* $g(x) = x^2$ *then*

$$Im(g) = \{y : y \geq 0, y \in \mathbb{R}\}$$

There is a name for the set of all ordered pairs drawn from two sets.

Definition 2.26 *If* A *and* B *are sets then the set of all ordered pairs with the first element from* A *and the second from* B *is called the* **Cartesian Product** *of* A *and* B*.*

The notation for the Cartesian product of A and B is $A \times B$. using curly brace notation:

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

Example 2.25 *If* $A = \{1, 2\}$ *and* $B = \{x, y\}$ *then*

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

The **Cartesian plane** *is an example of a Cartesian product of the real numbers with themselves:* $\mathbb{R} \times \mathbb{R}$*.*

2.3.1 Permutations

In this section we will look at a very useful sort of function, bijections of finite sets.

Definition 2.27 *A* **permutation** *is a bijection of a finite set with itself. Likewise a bijection of a finite set X with itself is called a* **permutation of X**.

Example 2.26 *Let $A = \{a, b, c\}$ then the possible permutations of A consist of the following six functions:*

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

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

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

Notice that the number of permutations of three objects does not depend on the identity of those objects. In fact there are always six permutations of any set of three objects. We now define a handy function that uses a rather odd notation. The method of showing permutations in Example 2.26, explicit listing of ordered pairs, is a bit cumbersome.

Definition 2.28 *Assume that we have agreed on an order, e.g. a,b,c, for the members of a set $X = \{a, b, c\}$. Then* **one-line** *notation for a permutation f consists of listing the first coordinate of the ordered pairs in the agreed on order. The table in Example 2.26 would become:*

abc acb
bac bca
cab cba

in one line notation. Notice the saving of space.

Definition 2.29 *The* **factorial** *of a natural number n is the product*

$$n(n-1)(n-2)\cdots 3\cdot 2\cdot 1 = \prod_{i=1}^{n} i$$

with the convention that the factorial of 0 is 1. We denote the factorial of n as $n!$, spoken "n factorial".

Example 2.27 *Here are the first few factorials:*

n	0	1	2	3	4	5	6	7
$n!$	1	1	2	6	24	120	720	5040

Proposition 2.7 *The number of permutations of a finite set with n elements is $n!$.*

Proof: *this proof is left as an exercise.*

Notice that one implication of Proposition 2.6 is that the composition of two permutations is a permutation. This means that the set of permutations of a set is *closed* under functional composition.

Definition 2.30 *A* **fixed point** *of a function $f : S \to S$ is any $x \in S$ such that $f(x) = x$. We say that* **f fixes x**.

Problems

Problem 2.42 *Suppose for finite sets A and B that $f : A \to B$ is an injective function. Prove that*

$$|B| \geq |A|$$

Problem 2.43 *Suppose that for finite sets A and B that $f : A \to B$ is a surjective function. Prove that $|A| \geq |B|$.*

Problem 2.44 *Using functions from the integers to the integers give an example of*

(i) A function that is an injection but not a surjection.

(ii) A function that is a surjection but not an injection.

(iii) A function that is neither an injection nor a surjection.

(iv) A bijection that is not the identity function.

Problem 2.45 *For each of the following functions from the real numbers to the real numbers say if the function is surjective or injective. It may be neither.*

(i) $f(x) = x^2$ (ii) $g(x) = x^3$

(iii) $h(x) = \begin{cases} \sqrt{x} & x \geq 0 \\ -\sqrt{-x} & x < 0 \end{cases}$

Interlude

The Collatz Conjecture

One of the most interesting features of mathematics is that it is possible to phrase problems in a few lines that turn out to be incredibly hard. The Collatz conjecture was first posed in 1937 by Lothar Collatz. Define the function f from the natural numbers to the natural numbers with the rule

$$f(n) = \begin{cases} 3n+1 & \text{n odd} \\ \frac{n}{2} & \text{n even} \end{cases}$$

Collatz' conjecture is that if you apply f repeatedly to a positive integer then the resulting sequence of numbers eventually arrives at one. If we start with 17, for example, the result of repeatedly applying f is:

$$f(17) = 52, f(52) = 26, f(26) = 13, f(13) = 40, f(40) = 20, f(20) = 10,$$
$$f(10) = 5, f(5) = 16, f(16) = 8, f(8) = 4, f(4) = 2, f(2) = 1$$

The sequences of numbers generated by repeatedly applying f to a natural number are called *hailstone sequences* with the collapse of the value when a large power of 2 appears being analogous to the impact of a hailstone. If we start with the number 27 then 111 steps are required to reach one and the largest intermediate number is 9232. This quite irregular behavior of the sequence is not at all apparent in the original phrasing of the problem.

The Collatz conjecture has been checked for numbers up to 5×2^{61} (about 5.764×10^{18}) by using a variety of computational tricks. It has not, however, been proven or disproven. The very simple statement of the problem causes mathematicians to underestimate the difficulty of the problem. At one point a mathematician suggested that the problem might have been developed by the Russians as a way to slow American mathematical research. This was after several of his colleagues spent months working on the problem without obtaining results.

A simple (but incorrect) argument suggests that hailstone sequences ought to grow indefinitely. Half of all numbers are odd, half are even. The function f slightly more than triples odd numbers and divides even numbers in half. Thus, on average, f increases the value of numbers. The problem is this: half of all even numbers are multiples of four and so are divided in half twice. One-quarter of all even numbers are multiples of eight and so get divided in half three times, and so on. The net effect of factors that are powers of two is to defeat the simple argument that f grows "on average".

Problem 2.46 *True or false (and explain): The function* $f(x) = \frac{x-1}{x+1}$ *is a bijection from the real numbers to the real numbers.*

Problem 2.47 *Find a function that is an injection of the integers into the even integers that does not appear in any of the examples in this chapter.*

Problem 2.48 *Suppose that* $B \subset A$ *and that there exists a bijection* $f : A \to B$*. What may be reasonably deduced about the set* A*?*

Problem 2.49 *Suppose that* A *and* B *are finite sets. Prove that* $|A \times B| = |A| \cdot |B|$*.*

Problem 2.50 *Suppose that we define* $h : \mathbb{N} \to \mathbb{N}$ *as follows. If* n *is even then* $h(n) = n/2$ *but if* n *is odd then* $h(n) = 3n+1$*. Determine if* h *is a (i) surjection or (ii) injection.*

Problem 2.51 *Prove proposition 2.6.*

Problem 2.52 *Prove or disprove: the composition of injections is an injection.*

Problem 2.53 *Prove or disprove: the composition of surjections is a surjection.*

Problem 2.54 *Prove proposition 2.7.*

Problem 2.55 *List all permutations of*

$$X = \{1, 2, 3, 4\}$$

using one-line notation.

Problem 2.56 *Suppose that* X *is a set and that* f*,* g*, and* h *are permutations of* X*. Prove that the equation* $f \circ g = h$ *has a solution* g *for any given permutations* f *and* h*.*

Problem 2.57 *Examine the permutation* f *of* $Q = \{a, b, c, d, e\}$ *which is* **bcaed** *in one line notation. If we create the series* $f, f \circ f, f \circ (f \circ f), \ldots$ *does the identity function,* **abcde***, ever appear in the series? If so, what is its first appearance? If not, why not?*

Problem 2.58 *If* f *is a permutation of a finite set, prove that the sequence* $f, f \circ f, f \circ (f \circ f), \ldots$ *must contain repeated elements.*

Problem 2.59 *Suppose that* X *and* Y *are finite sets and that* $|X| = |Y| = n$*. Prove that there are* $n!$ *bijections of* X *with* Y*.*

Problem 2.60 *Suppose that* X *and* Y *are sets with* $|X| = n$*,* $|Y| = m$*. Count the number of functions from* X *to* Y*.*

Problem 2.61 *Suppose that* X *and* Y *are sets with* $|X| = n$*,* $|Y| = m$ *for* $m > n$*. Count the number of injections of* X *into* Y*.*

Problem 2.62 *For a finite set* S *with a subset* T *prove that the permutations of* S *that have all members of* T *as fixed points form a set that is closed under functional composition.*

Problem 2.63 *Compute the number of permutations of a set* S *with* n *members that fix at least* $m < n$ *points.*

Problem 2.64 *Using any technique at all, estimate the fraction of permutations of an* n*-element set that have no fixed points. This problem is intended as an exploration.*

Problem 2.65 *Let* X *be a finite set with* $|X| = n$*. Let* $C = X \times X$*. How many subsets of* C *have the property that every element of* X *appears once as a first coordinate of some ordered pair and once as a second coordinate of some ordered pair?*

Problem 2.66 *An alternate version of Sigma (*$\sum$*) and Pi (*$\prod$*) notation works by using a set as an index. So if* $S = \{1, 3, 5, 7\}$ *then*

$$\sum_{s \in S} s = 16 \quad \text{and} \quad \prod_{s \in S} s = 105$$

Given all the material so far, give and defend reasonable values for the sum and product of an empty set.

Problem 2.67 *Suppose that* $f_\alpha : [0, 1] \to [0, 1]$ *for* $-1 < \alpha < \infty$ *is given by*

$$f_\alpha(x) = \frac{(\alpha + 1)x}{\alpha x + 1},$$

prove that f_α *is a bijection.*

Problem 2.68 *Find, to five decimals accuracy:*

$$Ln(200!)$$

Explain how you obtained the answer.

2.4 $\infty + 1$

We conclude the chapter with a brief section that demonstrates a strange thing that can be accomplished with set notation. We choose to represent the natural numbers $0, 1, 2, \ldots$ by sets that contain the number of elements counted by the corresponding natural number. We also choose to do so as simply as possible, using only curly braces and commas. Given this the numbers and their corresponding sets are:

$0 : \{\}$
$1 : \{\{\}\} = \{0\}$
$2 : \{\{\},\{\{\}\}\}=\{0,1\}$
$3 : \{\{\},\{\{\}\},\{\{\},\{\{\}\}\}\}=\{0,1,2\}$
$4 : \{\{\},\{\{\}\},\{\{\},\{\{\}\}\},\{\{\},\{\{\}\},\{\{\},\{\{\}\}\}\}\}$
$=\{0,1,2,3\}$

The trick for the above representation is this. Zero is represented by the empty set. One is represented by the set of the only thing we have constructed - zero, represented as the empty set. Similarly the representation of two is the set of the representation of zero and one (the empty set and the set of the empty set). This representation is incredibly inefficient but it uses a very small number of symbols. This representation also has a useful property. As always, we will start with a definition.

Definition 2.31 *The* **minimal set representation** *of the natural numbers is constructed as follows:*

(i) Let 0 be represented by the empty set.

(ii) For $n > 0$ let n be represented by the set $\{0, 1, \ldots, n-1\}$.

The shorthand $\{0, 1\}$ for $\{\{\}, \{\{\}\}\}$ is called the *simplified notation* for the minimal set representation. We now give the useful property of the minimal set representation.

Proposition 2.8 $n + 1 = n \cup \{n\}$

Proof:
This follows directly from Definition 2.31 by considering the set difference of the representations of n and $n - 1$. □
The definition says that any set of the representations of consecutive natural numbers, starting at zero, is the representation of the next natural number. This permits us to conclude that the set of all natural numbers

$$\{0, 1, 2, \ldots\}$$

fits the definition of a natural number. Which natural number is it? It is easy to see, in the minimal set representation, that for natural numbers m and n, $m < n$ implies that the representation of m is a subset of the representation of n. Every finite natural number is a subset of the set of all natural numbers and so we conclude that $\{0, 1, 2, \ldots\}$ is an infinite natural number. The set notation thus permits us to construct an infinite number.

The set consisting of the representations of all finite natural numbers is an infinite natural number. The number has been given the name ω, the lower-case omega. In addition to being a letter omega traditionally also means "the last". The number ω comes after all the finite natural numbers. If we now apply Proposition 2.8 we see that

$$\omega \cup \{\omega\} = \omega + 1$$

This means that we can add one to an infinite number. Is the resulting number $\omega + 1$ a different number from ω? It turns out the answer is "'yes", because the representations of these numbers are different as sets. The representation of ω contains no infinite sets while the representation of $\omega + 1$ contains one.

Problems

Problem 2.69 *Find the representation for 5 using the curly-brace-and-comma notation.*

Problem 2.70 *Give the minimal set representation of $\omega + 2$ using the simplified notation.*

Problem 2.71 *Suppose that $n > m$ are natural numbers and that S is the minimal set representation of n while T is the minimal set representation of m. Is the representation of $n - m$ a member of the set difference $S - T$?*

Problem 2.72 *Give a formula, as a function of n, for the number of times that the symbol { appears in the representation of n.*

Problem 2.73 *Prove or disprove: there are an infinite number of distinct infinite numbers.*

Interlude

Russell's Paradox

Bertrand Arthur William Russell, 3rd Earl Russell, OM, FRS (18 May 1872-2 February 1970), commonly known as simply Bertrand Russell, was a British philosopher, logician, mathematician, historian, religious skeptic, social reformer, socialist and pacifist. Although he spent the majority of his life in England, he was born in Wales, where he also died.

Let Q be the set of all sets that do not contain themselves as a member. Consider the question:"Does Q contain itself?" If the answer to this question is no then Q, by definition must contain itself. If, however, Q contains itself then it is by definition unable to contain itself. This rather annoying contradiction, constructed by Russell, had a rather amusing side effect.

Friedrich Frege had just finished the second of a three volume set of works called the *Basic Laws of Arithmetic* that was supposed to remove all intuition from mathematics and place it on a purely logical basis. Russell wrote Frege, explaining his paradox. Frege added an appendix to his second volume that attempted to avoid Russell's paradox. The third volume was never published.

It is possible to resolve Russell's paradox by being much more careful about what objects may be defined to be sets; the *category* of all sets that do not contain themselves gives rise to no contradiction (it does give rise to an entire field of mathematics, category theory). The key to resolving the paradox from a set theoretic perspective is that one cannot assume that, for every property, there is a set of all things satisfying that property. This is a reason why it is important that a set is properly defined. Another consequence of Russell's paradox is a warning that self-referential statements are both potentially interesting and fairly dangerous, at least on the intellectual plane.

The original phrasing of Russell's paradox was in terms of normal and abnormal sets. A set is *normal* if it fails to contain itself and abnormal otherwise. Consider the set of all normal sets. If this set is abnormal, it contains itself but by definition the set contains only normal sets and hence it is itself normal. The normality of this set forces the set to contain itself, which makes it abnormal. This is simply a rephrasing of the original contradiction.

Puzzle: what does the circuit below have to do with Russell's paradox and what use is it?