Sets, Infinity, and Mappings

Michael J. Neely
University of Southern California
http://www-bcf.usc.edu/~mjneely

Abstract

These notes discuss sets, mappings between sets, and the notion of countably infinite. This material supplements the lecture on countably and uncountably infinite sets in the EE 503 probability class. The material is useful because probability theory is defined over abstract sets, probabilities are defined as measures on subsets, and random variables are defined by mappings from an abstract set to a real number. Further, it is important to understand the difference between "countably infinite" and "uncountably infinite." One reason is that probability techniques for uncountably infinite sets are different from those for countably infinite sets. The final sections of the notes discuss related material that is not part of the course but may be of interest. This includes Cantor's theorem for power sets, and famous paradoxes.

I. Infinity in a nutshell

A. Cardinality

A set is *finite* if it has a finite number of elements. Specifically, a finite set has a number of elements equal to a non-negative integer (the set with no elements is called the *empty set*). For example, the following set of numbers is finite because it has only three elements:

$$\{3.4, 2.7, 9\}$$

Another example is the following set of three colors:

$$\{\text{red}, \text{blue}, \text{green}\}$$

The sets $\{3.4, 2.7, 9\}$ and $\{\text{red}, \text{blue}, \text{green}\}$ are composed of very different things. However, they are both finite. Further, they both have the same "size" because they can be put into a *one-to-one correspondence* where every element of the first set is matched to a unique element of the second, and all elements of the second set are included in this assignment of matches:

$$\begin{aligned} 3.4 &\leftrightarrow \text{red} \\ 2.7 &\leftrightarrow \text{blue} \\ 9 &\leftrightarrow \text{green} \end{aligned}$$

The above is just one of several different ways to construct a one-to-one correspondence between the sets $\{3.4, 2.7, 9\}$ and $\{\text{red}, \text{blue}, \text{green}\}$. For example, one could swap the color assignments of 3.4 and 2.7 to obtain the following alternative one-to-one correspondence:

$$\begin{aligned} 3.4 &\leftrightarrow \text{blue} \\ 2.7 &\leftrightarrow \text{red} \\ 9 &\leftrightarrow \text{green} \end{aligned}$$

It is not possible to create a one-to-one correspondence between the 2-element set $\{6, 8.9\}$ and the 3-element set {red, blue, green} because any attempt would necessarily leave out a color.

Definition 1: Two sets are said to have the *same cardinality* if there exists a one-to-one correspondence between them.

Therefore, the sets $\{3.4, 2.7, 9\}$ and $\{\text{red}, \text{blue}, \text{green}\}$ have the same cardinality, but the sets $\{6, 8.9\}$ and {red, blue, green} do not. The word "cardinality" in the world of sets can be interpreted as "size." Two finite sets have the same cardinality if and only if they have the same number of elements.

B. Infinite and countably infinite sets

A set is *infinite* if it is not finite. That is, an infinite set is one that has an infinite number of elements. An example of an infinite set is the set of all positive integers:

$$\{1, 2, 3, \ldots\}$$

Another example is the set of all even positive integers:

$$\{2, 4, 6, \ldots\}$$

The set of positive integers includes all even positive integers, but also includes more (in particular, it includes all odd positive integers). However, according to Definition 1, the two sets have the same cardinality because there exists a one-to-one correspondence between them:

$$\begin{aligned} 1 &\leftrightarrow 2 \\ 2 &\leftrightarrow 4 \\ 3 &\leftrightarrow 6 \\ 4 &\leftrightarrow 8 \\ &\ldots \end{aligned}$$

so that each positive integer n is assigned to the even positive integer $2n$.

Definition 2: An infinite set is said to be *countably infinite* if it has the same cardinality as the set of positive integers.

According to the above definition, the set of even integers is a countably infinite set. Now consider any infinite set $\mathcal{X}$ whose elements can be written as an infinite list $\{x_1, x_2, x_3, \ldots\}$. It is assumed here that the list contains each element of $\mathcal{X}$ once and only once. Then $\mathcal{X}$ is a countably infinite set, as shown by the following one-to-one correspondence:

$$\begin{aligned} 1 &\leftrightarrow x_1 \\ 2 &\leftrightarrow x_2 \\ 3 &\leftrightarrow x_3 \\ 4 &\leftrightarrow x_4 \\ &\ldots \end{aligned}$$

so that each positive integer n is assigned to the element x_n.

It follows that a set is countably infinite if and only if its elements can be written out as an infinite list, where each element is on the list once and only once. For this reason, countably infinite sets are often called *listable sets*. The nonnegative integers are countably infinite. The nonnegative even integers are countably infinite. The set of *all* integers, denoted $\mathbb{Z}$, is also countably infinite because that set can be listed as follows:

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, ...\}$$

where it is easy to see the rule generating the nth integer on the list. The set of all rational numbers in the interval $[0, 1)$ is also countably infinite. This is the set of all numbers of the form p/q where p, q are integers that satisfy $0 \leq p < q$. This set can be listed with the following 2-step procedure:

1) Form the initial sequence:

$$\{0, 1/2, 1/3, 2/3, 1/4, 2/4, 3/4, 1/5, 2/5, 3/5, 4/5, 1/6, 2/6, 3/6, 4/6, 5/6, ...\}$$

   The rule for this sequence is to start with 0, then sequentially list all rationals $p/q \in (0, 1)$ with $q = 2$, then $q = 3$, then $q = 4$, and so on.

2) Sequentially trim this list to remove redundancies:

$$\{0, 1/2, 1/3, 2/3, 1/4, \cancel{2/4}, 3/4, 1/5, 2/5, 3/5, 4/5, 1/6, \cancel{2/6}, \cancel{3/6}, \cancel{4/6}, 5/6, ...\}$$

Specifically, we strike out the nth number on the initial sequence if it is equal to one of the previous $n-1$ numbers on that sequence. The resulting trimmed sequence includes only *distinct* numbers. So we have our list $\{x_1, x_2, x_3, ...\}$ with

$$\begin{aligned} x_1 &= 0 \\ x_2 &= 1/2 \\ x_3 &= 1/3 \\ x_4 &= 2/3 \\ x_5 &= 1/4 \\ x_6 &= 3/4 \\ x_7 &= 1/5 \end{aligned}$$

and so on (for example, $x_{12} = 5/6$ and $x_{13} = 1/7$).

With some extra work it can also be shown that the set of *all* rational numbers, denoted $\mathbb{Q}$, is countably infinite. One may wonder if *all* infinite sets are countably infinite. Can we take any infinite set and write all of its elements out in an infinite list? The surprising answer is "no." The next two subsections show that the set of all real numbers between 0 and 1 is a set that is so large that its elements cannot be listed.

C. Decimal expansion of a real number

Let $[0, 1)$ denote the set of all real numbers that are greater than or equal to 0 and strictly less than 1. Every real number in $[0, 1)$ can be written using a decimal expansion. For example:

$$\begin{aligned} 0.348 &= 3 \times 10^{-1} + 4 \times 10^{-2} + 8 \times 10^{-3} \\ 0.4294 &= 4 \times 10^{-1} + 2 \times 10^{-2} + 9 \times 10^{-3} + 4 \times 10^{-4} \\ 0.703 &= 7 \times 10^{-1} + 0 \times 10^{-2} + 3 \times 10^{-3} \end{aligned}$$

A number might have a decimal expansion with an infinite number of digits:

$$\begin{aligned} 0.3333\ldots &= 3 \times 10^{-1} + 3 \times 10^{-2} + 3 \times 10^{-3} + 3 \times 10^{-4} + \cdots \\ 0.212121\ldots &= 2 \times 10^{-1} + 1 \times 10^{-2} + 2 \times 10^{-3} + 1 \times 10^{-4} + 2 \times 10^{-5} + 1 \times 10^{-6} \cdots \end{aligned}$$

The decimal expansion of a number is unique, *except for the possibility of having an infinite tail of 9s*. The following examples show that, if we are allowed to use an infinite tail of 9s, we can often write the same number in two different ways:

$$\begin{aligned} 0.999999\ldots &= 1 \\ 0.1399999\ldots &= 0.14 \end{aligned}$$

This fact about infinite tails of 9s can be formally proven, but that is a tangential detail that is omitted for brevity.[1] In the special case when a number can be written with two different decimal expansions, one of those expansions must have an infinite tail of 9s while the other does not. In that case, we can just choose the expansion that does *not* have the infinite tail of 9s. This is formalized by the following fact.

Fact 1: Every real number x in the set $[0, 1)$ can be uniquely written as a decimal expansion:

$$x = 0.a_1a_2a_3\cdots = a_1 \times 10^{-1} + a_2 \times 10^{-2} + a_3 \times 10^{-3} + \cdots$$

where the digits $a_1, a_2, a_3, \ldots$ satisfy the following:

- Each digit a_i is an integer in the 10-element set $\{0, 1, 2, \ldots, 9\}$.
- The sequence of digits $\{a_1, a_2, a_3, \ldots\}$ does not have an infinite tail of 9s.

[1] A proof uses the fact that, for all integers k, one has the identity $\sum_{i=k}^{\infty}[9 \times 10^{-i}] = 10^{-k+1}$.

D. Cantor's diagonal argument

Definition 3: A set is *uncountably infinite* if it is infinite but not countably infinite.

Intuitively, an uncountably infinite set is an infinite set *that is too large to list*. This subsection proves the existence of an uncountably infinite set. In particular, it proves that the set of all real numbers in the interval $[0, 1)$ is uncountably infinite. The proof starts by assuming that $[0, 1)$ is countably infinite, and then reaches a contradiction.[2]

Suppose the set $[0, 1)$ is countably infinite, so that its elements can be written as an infinite list (we will reach a contradiction). Then we can list all elements of $[0, 1)$ as $\{x_1, x_2, x_3, \ldots\}$ where:

$$\begin{aligned} x_1 &= 0.a_{11}a_{12}a_{13}a_{14}a_{15}\cdots \\ x_2 &= 0.a_{21}a_{22}a_{23}a_{24}a_{25}\cdots \\ x_3 &= 0.a_{31}a_{32}a_{33}a_{34}a_{35}\cdots \\ &\cdots \\ x_i &= 0.a_{i1}a_{i2}a_{i3}a_{i4}a_{i5}\cdots \\ &\cdots \end{aligned}$$

where each number x_i on the list is written with its unique decimal expansion. Now draw a box around each "diagonal digit" in this list of decimal expansions:

$$\begin{aligned} x_1 &= 0.\boxed{a_{11}}a_{12}a_{13}a_{14}a_{15}\cdots \\ x_2 &= 0.a_{21}\boxed{a_{22}}a_{23}a_{24}a_{25}\cdots \\ x_3 &= 0.a_{31}a_{32}\boxed{a_{33}}a_{34}a_{35}\cdots \end{aligned}$$

The idea is to construct a new real number x^* that is not on the list by designing its decimal expansion $0.b_1b_2b_3\cdots$ such that each digit b_i differs from the boxed digit a_{ii}. Specifically, define the real number x^* as follows:

$$x^* = 0.b_1b_2b_3b_4b_5\cdots$$

where each digit b_i is defined:

$$b_i = \begin{cases} 8 & \text{if } a_{ii} = 7 \\ 7 & \text{if } a_{ii} \neq 7 \end{cases}$$

Since the decimal expansion of x^* has only 7s and 8s, it has no infinite tail of 9s. Thus, $0.b_1b_2b_3\cdots$ is the unique decimal expansion of x^*. Also, by this construction we have:

$$\begin{aligned} b_1 &\neq a_{11} \\ b_2 &\neq a_{22} \\ b_3 &\neq a_{33} \\ &\cdots \end{aligned}$$

so that, for every digit $i \in \{1, 2, 3, ...\}$, we have $b_i \neq a_{ii}$. *It follows that x^* is a real number that is not on the list, since its unique decimal expansion $0.b_1b_2b_3b_4\cdots$ differs (in at least one digit) from the unique decimal expansion of every number on the list.*

For example, x^* cannot be the same as x_3 because:

$$\begin{aligned} x^* &= 0. \quad b_1 \quad b_2 \quad \boxed{b_3} \quad b_4 \quad b_5 \quad \cdots \\ x_3 &= 0. \quad a_{31} \quad a_{32} \quad \boxed{a_{33}} \quad a_{34} \quad a_{35} \quad \cdots \end{aligned}$$

and $b_3 \neq a_{33}$. Likewise, x^* cannot be the same as x_4 because:

$$\begin{aligned} x^* &= 0. \quad b_1 \quad b_2 \quad b_3 \quad \boxed{b_4} \quad b_5 \quad \cdots \\ x_3 &= 0. \quad a_{41} \quad a_{42} \quad a_{43} \quad \boxed{a_{44}} \quad a_{45} \quad \cdots \end{aligned}$$

[2]Formally, one should first note that $[0, 1)$ is indeed an infinite set. For example, it includes all numbers in the infinite list $\{1/2, 1/3, 1/4, 1/5, \ldots\}$.

and $b_4 \neq a_{44}$.

However, x^* is indeed a real number in the interval $[0, 1)$ because its decimal expansion has only 7s and 8s, so:

$$0 \leq 0.7777777\ldots \leq x^* \leq 0.888888\ldots < 1$$

Therefore, x^* is a real number in the interval $[0, 1)$ that is not on the list. This contradicts the assumption that the list contains all real numbers in $[0, 1)$.

E. Discussion

The proof in the previous subsection assumed existence of an infinite list $\{x_1, x_2, x_3, \ldots\}$ that contains all real numbers in the interval $[0, 1)$, and reached a contradiction by finding a particular real number x^* in $[0, 1)$ that was *not* on the list. One might object as follows: Why don't we simply take the old list and add to it the single real number x^*? This would form a new list $\{x^*, x_1, x_2, x_3, \ldots\}$, where x^* is now the first on the list, the old x_1 is now the second on the list, and so on.

This does not work because the argument can be repeated to find yet another number y^* in $[0, 1)$ that is not on the infinite list $\{x^*, x_1, x_2, x_3, \ldots\}$. Adding y^* to the list to form yet another list $\{y^*, x^*, x_1, x_2, x_3, \ldots\}$ does not help, for the same reason. Overall, the proof required us to only find one such number that was not on the list. However, it turns out that for any list containing numbers in $[0, 1)$, the set of all real numbers in $[0, 1)$ that are *not* on the list is, in fact, uncountably infinite.

F. Diagonalization exercises

Exercise 1: Let $\{x_1, x_2, x_3, ...\}$ be a sequence of real numbers in the interval $[0, 1)$. For each positive integer i, the unique decimal expansion of x_i is given by $x_i = 0.a_{i1}a_{i2}a_{i3}...$, so the expansion does not have an infinite tail of 9s, and $x_i = \sum_{j=1}^{\infty} a_{ij}10^{-j}$. Construct a real number $y \in [0, 1)$ that is not on the list $\{x_1, x_2, x_3, ...\}$.

Exercise 2: Let $\{x_1, x_2, x_3, \ldots\}$ be a listing of all rational numbers in the set $[0, 1)$. For each positive integer i, define the decimal expansion of x_i as:

$$x_i = 0.a_{i1}a_{i2}a_{i3}\cdots = \sum_{k=1}^{\infty} a_{ik}10^{-k}$$

where the expansion does not have an infinite tail of 9s. We want to prove that the following number z is irrational:

$$z = \sum_{k=1}^{\infty} a_{kk}10^{-k}$$

a) For each positive integer k, define:

$$b_k = \begin{cases} 3 & \text{if } a_{kk} = 5 \\ 5 & \text{if } a_{kk} \neq 5 \end{cases}$$

Define $y = \sum_{k=1}^{\infty} b_k 10^{-k}$. Prove that y is irrational.

b) Use part (a) to prove that z is irrational. *Hint: A number is rational if and only if its decimal expansion has an eventually repeating pattern. Suppose* $\{a_{11}, a_{22}, a_{33}, \ldots\}$ *has an eventually repeating pattern.*

G. Infinite levels of infinity

If $\mathcal{A}$ and $\mathcal{B}$ are finite sets and $\mathcal{A}$ has fewer elements than $\mathcal{B}$, we say that set $\mathcal{A}$ has *strictly smaller cardinality* than set $\mathcal{B}$. It is possible to precisely define the notion of "strictly smaller cardinality" for arbitrary (possibly infinite) sets. The statement "set $\mathcal{A}$ has strictly smaller cardinality than set $\mathcal{B}$" is often written with the simpler shorthand:

$$|\mathcal{A}| < |\mathcal{B}|$$

where the notation $|\mathcal{A}|$ is read "the cardinality of $\mathcal{A}$."

For any set $\mathcal{A}$, one can define a new set $P(\mathcal{A})$, called the *power set of* $\mathcal{A}$, as the set of all subsets of $\mathcal{A}$. For example, the 3-element set $\mathcal{A} = \{1, 2, 3\}$ has an 8-element power set:

$$P(\mathcal{A}) = \{\{\}, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

where "$\{\}$" represents the "empty set" (alternatively, the empty set is often denoted by the symbol ϕ). Another simple argument of Cantor shows that any non-empty set $\mathcal{A}$ has strictly smaller cardinality than its power set (see Theorem 2 in a later section). That is:

$$|\mathcal{A}| < |P(\mathcal{A})| \tag{1}$$

This can be used to show that there are an infinite number of levels of infinity. Indeed, let $\mathbb{N}$ denote the set of all positive integers (also called *natural numbers*). Then:

$$|\mathbb{N}| < |P(\mathbb{N})| < |P(P(\mathbb{N}))| < |P(P(P(\mathbb{N})))| < \cdots \tag{2}$$

Thus, starting with the infinite set $\mathbb{N}$, the power set operation iteratively produces new infinite sets with strictly larger cardinality than all of their predecessors.

All sets with cardinality larger than $|\mathbb{N}|$ are uncountably infinite. It can be shown that the set of all real numbers, denoted $\mathbb{R}$, has the same cardinality as $P(\mathbb{N})$. The set $\mathbb{R}$ also has the same cardinality as the interval $[0, 1)$, which has the same cardinality as the interval $[0, 1]$. Indeed, it can be shown that adding a single element to an infinite set does not change the cardinality of that set. Further, for any positive integer n, it can be shown that $\mathbb{R}$ has the same cardinality as the set $\mathbb{R}^n$ that consists of all n-tuples of real numbers $(x_1, ..., x_n)$. That is

$$|P(\mathbb{N})| = |\mathbb{R}| = |[0, 1)| = |[0, 1]| = |\mathbb{R}^n| \quad \forall n \in \{1, 2, 3, ...\}$$

The set of all infinite real-valued sequences $(x_1, x_2, x_3, ...)$, sometimes denoted $\mathbb{R}^{\mathbb{N}}$, can also be shown to have the same cardinality as $\mathbb{R}$. In particular, while (2) shows there are an infinite number of levels of uncountably infinite sets, the uncountably infinite sets of most practical interest (that is, the uncountably infinite sets we shall work with in the EE 503 class) have the same cardinality as $\mathbb{R}$. An example set that has cardinality larger than $|\mathbb{R}|$ is the set of all real-valued functions $f : \mathbb{R} \to \mathbb{R}$. Indeed, this set can be shown to have cardinality $|P(\mathbb{R})|$, which is the same as $|P(P(\mathbb{N}))|$. However, it can be shown that the set of all *continuous* functions $f : \mathbb{R} \to \mathbb{R}$ has cardinality $|\mathbb{R}|$. In particular, there are "way more" discontinuous functions than there are continuous functions.

H. A deep paradox

The chain of sets given in (2) shows there are infinitely many levels of infinity. That is, there are infinitely many different infinite cardinalities. Of course, this iterative method can only demonstrate existence of a *countably infinite* number of levels of infinity. This leads one to wonder if the set of all distinct levels of infinity is in fact a countably infinite set.

There is a simple proof that the answer is "no." It is a resounding "no." It is the loudest possible "no" that one can imagine! *The set of all distinct levels of infinity is so large that it cannot even be called a set!* The standard proof of this fact starts by assuming we can represent all distinct infinite cardinalities as a set, call it set $\mathcal{C}$. It then uses a variation on the power set principle to construct a new infinite set with cardinality that is not in $\mathcal{C}$.

The specific proof is as follows: Suppose $\mathcal{C}$ is the collection of all distinct infinite cardinalities. Suppose that for each distinct cardinality c in $\mathcal{C}$, we have a representative set $\mathcal{X}_c$ that has this cardinality. Hence, we have a collection of sets, one for each possible cardinality. It can be shown that the cardinality of any particular set in the collection is less than or equal to the cardinality of the *union of all sets* in the collection, and so the power set of this union has *strictly greater* cardinality than any set in the collection.[3] That is, the following holds for every cardinality d in the set $\mathcal{C}$:

$$|\mathcal{X}_d| \leq |\cup_{c\in\mathcal{C}} \mathcal{X}_c| < |P(\cup_{c\in\mathcal{C}}\mathcal{X}_c)| \tag{3}$$

This result is arguably one of the deepest "paradoxes" of mathematics.

[3]Formally, the following intuitive fact can be proven: if $\mathcal{A}$ and $\mathcal{B}$ are sets that satisfy $\mathcal{A} \subseteq \mathcal{B}$ (so that all elements of $\mathcal{A}$ are also in $\mathcal{B}$), then $|\mathcal{A}| \leq |\mathcal{B}|$. The first inequality in (3) follows immediately from this fact together with the observation that $\mathcal{X}_d \subseteq \cup_{c\in\mathcal{C}}\mathcal{X}_c$ whenever d is in the set $\mathcal{C}$. The second inequality in (3) follows from (1). The notation "$\cup_{c\in\mathcal{C}}\mathcal{X}_c$" means "the union of all sets $\mathcal{X}_c$ such that c is in the set $\mathcal{C}$" and represents the set of all elements that are in one or more of the sets in the collection.

II. Notation and symbolism

This section discusses standard notation and symbolism for sets.

A. What is it?

A set $\mathcal{A}$ is a collection of *objects*, also called *elements*, *members*, or *points*. The term *objects* emphasizes that sets can be formed from different types of things (i.e., sets of numbers, shapes, colors, functions, vectors, and so on). For example, the following sets $\mathcal{A}$ and $\mathcal{B}$ consist of objects related to writing:

$$\mathcal{A} = \{\text{pencil}, \text{pen}, \text{computer}\} \tag{4}$$
$$\mathcal{B} = \{\text{pencil}, \text{pen}\} \tag{5}$$

The term *elements* emphasizes that the objects of a set can be viewed as its smallest "atomic" pieces. The union of all pieces forms the whole set. The set $\mathcal{A}$ defined in (4) is a 3-element set (it consists of 3 objects). The set $\mathcal{B}$ defined in (5) is a 2-element set.

The term *members* is useful when sets are defined from larger sets by including only those elements that have certain distinguishing properties. For example, let $\mathbb{R}$ denote the set of all real numbers. Out of this set, one can define a new set $\mathbb{N}$ whose members consist only of positive integers (called *natural numbers*):

$$\mathbb{N} = \{1, 2, 3, \ldots\}$$

Out of the set $\mathbb{N}$, one can define a new set whose membership is restricted to those elements of $\mathbb{N}$ that have values less than or equal to 4:

$$\mathcal{C} = \{x \in \mathbb{N} : x \leq 4\} = \{1, 2, 3, 4\}$$

The notation "$\{x \in \mathbb{N} : x \leq 4\}$," used above, is read "The set of all x in the set $\mathbb{N}$ such that x is less than or equal to 4." In particular, the colon ":" represents the phrase "such that." With these definitions for $\mathbb{N}$ and $\mathcal{C}$, the real number 3.4 is not a member of $\mathbb{N}$ because it is not a positive integer. The positive integer 8 is not a member of $\mathcal{C}$ because its value is larger than 4.

The term *points* is useful when attempting to visualize properties of abstract sets by drawing pictures of circles on a piece of paper and pretending that all objects of the set are physical points inside the circle. This visualization is useful for understanding intersections between one or more sets, even when the sets in question are not 2-dimensional points.

B. Notation

The following notation is useful:

- $x \in \mathcal{A}$: This means "x is an element of $\mathcal{A}$."
- $x \notin \mathcal{A}$: This means "x is not an element of $\mathcal{A}$."
- $\mathcal{B} \subseteq \mathcal{A}$: This means "$\mathcal{B}$ is a subset of $\mathcal{A}$." Specifically, this means that all elements of $\mathcal{B}$ are also in $\mathcal{A}$.
- $\mathcal{A} = \mathcal{B}$: This means that sets $\mathcal{A}$ and $\mathcal{B}$ are the same.

A statement of the form "$\mathcal{B} \subseteq \mathcal{A}$" is called an *inclusion* (such statements are analogous to *inequality statements* of the form "$x \leq y$" for real numbers x and y). For example, if sets $\mathcal{A}$ and $\mathcal{B}$ are defined by (4)-(5), then the inclusion $\mathcal{B} \subseteq \mathcal{A}$ is true because both "pencil" and "pen" are in $\mathcal{A}$. A set is always a subset of itself. That is, the inclusion $\mathcal{A} \subseteq \mathcal{A}$ holds for all sets $\mathcal{A}$. A set $\mathcal{B}$ is defined to be a *proper subset of* $\mathcal{A}$, written $\mathcal{B} \subsetneq \mathcal{A}$, if $\mathcal{B} \subseteq \mathcal{A}$ but $\mathcal{B}$ is not the same as $\mathcal{A}$. This happens when all elements of $\mathcal{B}$ are also in $\mathcal{A}$, but $\mathcal{A}$ has at least one element that is not in $\mathcal{B}$. With the example sets $\mathcal{A}$ and $\mathcal{B}$ from (4)-(5), it is clear that $\mathcal{B}$ is a proper subset of $\mathcal{A}$.

The following sets of numbers and tuples of numbers are standard:

$$\begin{aligned} \mathbb{R} &= \text{The set of real numbers} \\ \mathbb{N} &= \text{The set of natural numbers} = \{1, 2, 3, \ldots\} \\ \mathbb{R}^N &= \{(x_1, \ldots, x_N) : x_i \in \mathbb{R} \text{ for all } i \in \{1, \ldots, N\}\} \end{aligned}$$

In the definition of $\mathbb{R}^N$ above, it is assumed that N has already been defined as a positive integer. Thus, $\mathbb{R}^N$ is the set of all N-tuples of real numbers, also called *N-dimensional Euclidean space*. The following notation for *intervals* of the real number line $\mathbb{R}$ is also used: Let a and b be two real numbers that satisfy $a < b$. Then:

$$\begin{aligned}[a,b] &= \{x \in \mathbb{R} : a \leq x \leq b\} \\ (a,b) &= \{x \in \mathbb{R} : a < x < b\} \\ (a,b] &= \{x \in \mathbb{R} : a < x \leq b\}\end{aligned}$$

Exercise 3: Write definitions of intervals $[a,b)$, $(-\infty,b]$, and (a,∞) using the set structure as above. That is, write: $[a,b) = \{x \in \mathbb{R} : \cdots\}$ (where you fill in the rest), and so on.

C. Unions, intersections, and complements

For two sets $\mathcal{A}$ and $\mathcal{B}$, define:

- $\mathcal{A} \cup \mathcal{B}$: This is the *union* of $\mathcal{A}$ and $\mathcal{B}$, consisting of the set of all objects that are in either $\mathcal{A}$ or $\mathcal{B}$.
- $\mathcal{A} \cap \mathcal{B}$. This is the *intersection* of $\mathcal{A}$ and $\mathcal{B}$, consisting of all objects in *both* $\mathcal{A}$ and $\mathcal{B}$.

In order to define intersections in cases when $\mathcal{A}$ and $\mathcal{B}$ have no elements in common, it is useful to define the *empty set* ϕ as the set with no elements. Two sets with no elements in common are said to be *disjoint*. Thus, $\mathcal{A}$ and $\mathcal{B}$ are disjoint if and only if $\mathcal{A} \cap \mathcal{B} = \phi$. By convention, the empty set is viewed as a subset of every set, so that $\phi \subseteq \mathcal{A}$ for all sets $\mathcal{A}$. The empty set is often represented with the alternative notation $\{\}$.

For an infinite sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$, the infinite unions and intersections are defined:

- $\cup_{n=1}^{\infty} \mathcal{A}_n$: This is the set of all objects that are in at least one set $\mathcal{A}_n$ for some positive integer n.
- $\cap_{n=1}^{\infty} \mathcal{A}_n$: This is the set of all objects that are in *all* sets $\mathcal{A}_n$ (for all $n \in \{1,2,3,\ldots\}$).

In some cases it is useful to define sets $\mathcal{A}_x$ using an index x that takes values in some general *index set* $\mathcal{I}$ (where $\mathcal{I}$ is an arbitrary nonempty set, possibly different from $\mathbb{N}$). Unions and intersections are then defined:

- $\cup_{x \in \mathcal{I}} \mathcal{A}_x$: This is the set of all objects that are in at least one set $\mathcal{A}_x$ for some $x \in \mathcal{I}$.
- $\cap_{x \in \mathcal{I}} \mathcal{A}_x$: This is the set of all objects that are in *all* sets $\mathcal{A}_x$ (for all $x \in \mathcal{I}$).

For many problems it is understood that the sets of interest are subsets of some larger set $\mathcal{S}$. In that case, if $\mathcal{A}$ is a subset of $\mathcal{S}$, its *complement*, denoted $\mathcal{A}^c$, is the set of all objects in $\mathcal{S}$ that are not in $\mathcal{A}$. Relationships between unions, intersections, and complements are revealed by *DeMorgan's laws*.

DeMorgan's laws (version 1): For any two sets $\mathcal{A}$ and $\mathcal{B}$ that are subsets of $\mathcal{S}$, we have:

- $(\mathcal{A} \cup \mathcal{B})^c = \mathcal{A}^c \cap \mathcal{B}^c$.
- $(\mathcal{A} \cap \mathcal{B})^c = \mathcal{A}^c \cup \mathcal{B}^c$.

DeMorgan's laws (version 2): Let $\mathcal{I}$ be an arbitrary index set. For each $x \in \mathcal{I}$, let $\mathcal{A}_x$ be a subset of $\mathcal{S}$. Then:

- $(\cup_{x \in \mathcal{I}} \mathcal{A}_x)^c = \cap_{x \in \mathcal{I}} \mathcal{A}_x^c$.
- $(\cap_{x \in \mathcal{I}} \mathcal{A}_i)^c = \cup_{x \in \mathcal{I}} \mathcal{A}_x^c$.

To prove two sets $\mathcal{A}$ and $\mathcal{B}$ are the same (so that $\mathcal{A} = \mathcal{B}$), we typically first prove $\mathcal{A} \subseteq \mathcal{B}$, and then prove $\mathcal{B} \subseteq \mathcal{A}$. This method is used in the following proof of DeMorgan's laws.

Proof: (DeMorgan's laws) For brevity we only prove that $(\cup_{x \in \mathcal{I}} \mathcal{A}_x)^c = \cap_{x \in \mathcal{I}} \mathcal{A}_x^c$.

Forward inclusion: Suppose $a \in (\cup_{x \in \mathcal{I}} \mathcal{A}_x)^c$. Then a is not in $\mathcal{A}_x$ for any $x \in \mathcal{I}$. That is, $a \in \mathcal{A}_x^c$ for all $x \in \mathcal{I}$. Hence, $a \in \cap_{x \in \mathcal{I}} \mathcal{A}_x^c$. Thus, $(\cup_{x \in \mathcal{I}} \mathcal{A}_x)^c \subseteq \cap_{x \in \mathcal{I}} \mathcal{A}_x^c$.

Reverse inclusion: Now suppose $a \in \cap_{x \in \mathcal{I}} \mathcal{A}_x^c$. Then a is not in $\mathcal{A}_x$ for any $x \in \mathcal{I}$. Thus, $a \notin \cup_{x \in \mathcal{I}} \mathcal{A}_x$. Hence, $a \in (\cup_{x \in \mathcal{I}} \mathcal{A}_x)^c$, and so $\cap_{x \in \mathcal{I}} \mathcal{A}_x^c \subseteq (\cup_{x \in \mathcal{I}} \mathcal{A}_x)^c$. □

D. Defining sets

A set can be defined by either listing its elements, or specifying a *defining property* that is satisfied by all of its members (and none of its non-members), such as:

$$\begin{aligned}
\mathcal{A} &= \{\text{three letter words in the above sentence}\} \\
&= \{\text{set, can, its, all, and}\} \\
\mathcal{B} &= \{\text{even integers}\} \\
&= \{2, 4, 6, 8, \ldots\} \\
\mathcal{C} &= \{\text{integers that are perfect squares}\} \\
&= \{1, 4, 9, 16, 25, \ldots\} \\
&= \{n \in \mathbb{N} : n = a^2 \text{ for some } a \in \mathbb{N}\}
\end{aligned}$$

One should use a defining property in cases when it is difficult or impossible to list all the elements.

III. Mappings between sets

Let $\mathcal{A}$ and $\mathcal{B}$ be two abstract sets. A function $f(a)$ from set $\mathcal{A}$ to set $\mathcal{B}$ is represented by the notation:

$$f : \mathcal{A} \to \mathcal{B}$$

The function f assigns every element $a \in \mathcal{A}$ a "value" $f(a) \in \mathcal{B}$. A function from a set $\mathcal{A}$ to a set $\mathcal{B}$ is also called a *mapping from $\mathcal{A}$ into $\mathcal{B}$*.

For example, define sets $\mathcal{A}$ and $\mathcal{B}$ as follows:

$$\begin{aligned}
\mathcal{A} &= \{\text{pencil, pen, computer}\} \\
\mathcal{B} &= \{(0,0), (0,1)\}
\end{aligned}$$

For these sets, one can define the following function $f : \mathcal{A} \to \mathcal{B}$:

$$f(\text{pencil}) = (0,1) \tag{6}$$
$$f(\text{pen}) = (0,1) \tag{7}$$
$$f(\text{computer}) = (0,0) \tag{8}$$

This function f has the property that every element of $\mathcal{B}$ is mapped to from at least one element of $\mathcal{A}$. However, the element "$(0,1)$" in $\mathcal{B}$ is mapped to from two distinct elements of $\mathcal{A}$.

As an example function defined over an infinite set, consider $g : \mathbb{N} \to \mathbb{R}$ defined by $g(x) = \sqrt{x}$. This function maps the natural numbers into the set of real numbers. Every natural number maps to a unique real number, but not all real numbers get mapped to.

Definition 4: A function $f : \mathcal{A} \to \mathcal{B}$ is an *injection* (also called "one-to-one") if $f(x) = f(y)$ only when $x = y$. That is, all elements of $\mathcal{A}$ are mapped to *unique* elements of $\mathcal{B}$.

Definition 5: A function $f : \mathcal{A} \to \mathcal{B}$ is a *surjection* (also called "onto") if for all $b \in \mathcal{B}$, there is a (possibly non-unique) $a \in \mathcal{A}$ such that $f(a) = b$. That is, all elements of $\mathcal{B}$ get mapped to.

Definition 6: A function $f : \mathcal{A} \to \mathcal{B}$ is a *bijection* if it is both an injection and a surjection.

A function is called *injective* if it is an injection, *surjective* if it is a surjection, and *bijective* if it is a bijection. A bijection between sets $\mathcal{A}$ and $\mathcal{B}$ is also called a *one-to-one correspondence between $\mathcal{A}$ and $\mathcal{B}$*. For example, the function $f : \mathcal{A} \to \mathcal{B}$ defined in (6)-(8) is surjective but not injective. The function $g : \mathbb{N} \to \mathbb{R}$ defined by $g(x) = \sqrt{x}$ is injective but not surjective. Define $\mathcal{C}$ as the set of even positive integers, so that:

$$\mathcal{C} = \{2, 4, 6, 8, \ldots\}$$

The function $h : \mathbb{N} \to \mathcal{C}$ defined by $f(n) = 2n$ is a bijection from $\mathbb{N}$ to $\mathcal{C}$.

A. Proving injectivity and surjectivity

To prove that a function $f : \mathcal{A} \to \mathcal{B}$ is an injection, we start the proof as follows:

"Suppose x and y are in $\mathcal{A}$ and satisfy $f(x) = f(y)$. We want to show that $x = y$."

To prove that a function $f : \mathcal{A} \to \mathcal{B}$ is a surjection, we start the proof as follows:

"Suppose $y \in \mathcal{B}$. We want to show there is an $x \in \mathcal{A}$ such that $f(x) = y$."

B. Simple examples

Example 1: Define $\mathcal{A} = \{red, green, blue\}$ and $\mathcal{B} = \{1, 2, 3, 4, 5\}$. Define $f : \mathcal{A} \to \mathcal{B}$ and $g : \mathcal{A} \to \mathcal{B}$ by:

- $f(red) = 1, f(green) = 4, f(blue) = 2$
- $g(red) = 2, g(green) = 5, g(blue) = 2$

Then function f is an injection, but not a surjection. The function g is neither injective nor surjective. Note that it is impossible to define a surjective function $h : \mathcal{A} \to \mathcal{B}$ because set $\mathcal{A}$ only has 3 elements, while set $\mathcal{B}$ has 5.

Example 2: Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 2x+5$. Prove that this function is both injective and surjective (and hence it is a bijection from $\mathbb{R}$ to $\mathbb{R}$).

Proof: (Injective) Suppose a and b are elements of $\mathbb{R}$ such that $f(a) = f(b)$. We want to show that $a = b$. Since $f(a) = f(b)$, we have:

$$2a + 5 = 2b + 5$$

Simplifying the above equation gives $a = b$. □

Proof: (Surjective) Suppose $y \in \mathbb{R}$. We want to show there is an $x \in \mathbb{R}$ such that $f(x) = y$. That is, we want to show there exists an $x \in \mathbb{R}$ that satisfies $2x + 5 = y$. It is clear that the value $x = (y - 5)/2$ is a real number that satisfies $2x + 5 = y$, and so we are done. □

Example 3: Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. Prove that f is neither injective nor surjective.

Proof: (Not injective) $f(1) = f(-1)$, but $1 \neq -1$. Thus, f is not injective. □

Proof: (Not surjective) Since $f(x) = x^2$, we have $f(x) \geq 0$ for all $x \in \mathbb{R}$. So $-1 \in \mathbb{R}$, but there is no $x \in \mathbb{R}$ such that $f(x) = -1$. So f is not surjective. □

Example 4: Consider the function $f : [0, \infty) \to [0, \infty)$ defined by $f(x) = x^2$. Show that f is a bijection.
Proof: (Injective) Suppose a and b are real numbers in the interval $[0, \infty)$ that satisfy $f(a) = f(b)$. We want to show that $a = b$. Since $f(a) = f(b)$, we have $a^2 = b^2$. Since a and b are both nonnegative, it must be that $a = b$. □

Proof: (Surjective) Let $y \in [0, \infty)$. We want to show there is an $x \in [0, \infty)$ such that $f(x) = y$. That is, we want to find a nonnegative real number x that satisfies $x^2 = y$. Define $x = \sqrt{y}$. Since y is nonnegative, the value x is well defined as a real and nonnegative number, and clearly $x^2 = y$. □

Example 5: Find a bijection $f : (0, 1) \to (2, 8)$.

Solution: The simplest example is a linear function. Define $f(x) = 2 + 6x$. It is easy to show (using the same methods as the above example proofs) that $f(x)$ is a bijection from $(0, 1)$ to $(2, 8)$.

Example 6: Is it possible to find an injection $f : (-\pi/2, \pi/2) \to \mathbb{R}$?

Solution: Yes, consider the function $f(x) = \tan(x)$. A formal proof that this is a bijection can be constructed via results of exercises 5 and 7 below. The function $\tan(x)$ also has an *inverse* $\arctan(x)$.

C. Inverses

Let $\mathcal{A}$ and $\mathcal{B}$ be two abstract sets. If $f : \mathcal{A} \to \mathcal{B}$ is bijective, then we can define its *inverse function* $f^{-1} : \mathcal{B} \to \mathcal{A}$, where for each $b \in \mathcal{B}$, the value $f^{-1}(b)$ is defined as the (unique) $a \in \mathcal{A}$ such that $f(a) = b$. With this definition, it follows that:

$$f(f^{-1}(b)) = b \quad \text{for all } b \in \mathcal{B} \tag{9}$$

$$f^{-1}(f(a)) = a \quad \text{for all } a \in \mathcal{A} \tag{10}$$

Lemma 1: If $f : \mathcal{A} \to \mathcal{B}$ is bijective, then $f^{-1} : \mathcal{B} \to \mathcal{A}$ is also bijective.

Proof: (Injective) Suppose b_1, b_2 are elements of $\mathcal{B}$ that satisfy:

$$f^{-1}(b_1) = f^{-1}(b_2)$$

We want to show that $b_1 = b_2$. Taking the function $f(\cdot)$ of both sides of the above equality gives:

$$f(f^{-1}(b_1)) = f(f^{-1}(b_2))$$

It follows by (9) that $b_1 = b_2$. □

Proof: (Surjective) Suppose $a \in \mathcal{A}$. We want to show there is a $b \in \mathcal{B}$ such that $f^{-1}(b) = a$. Define $b = f(a)$. Then $b \in \mathcal{B}$, and $f^{-1}(b) = f^{-1}(f(a)) = a$. □

D. Exercises

Exercise 4: Give an example of sets $\mathcal{A}$ and $\mathcal{B}$, where $\mathcal{B}$ has 5 elements, and a function $f : \mathcal{A} \to \mathcal{B}$ that is surjective but not injective.

Exercise 5: Let $\mathcal{A} \subseteq \mathbb{R}$. Suppose $f : \mathcal{A} \to \mathbb{R}$ is a strictly increasing function, so that if a and b are elements of $\mathcal{A}$ that satisfy $a < b$, then $f(a) < f(b)$. Prove that f is injective.

Exercise 6: Is the result of the previous exercise true if the "strictly increasing" assumption is changed to "nondecreasing"?

Exercise 7: Let $\mathcal{A}$ be a (possibly infinite) interval of $\mathbb{R}$, and let $f : \mathcal{A} \to \mathbb{R}$ be a continuous function that has arbitrarily large and small values over $\mathcal{A}$. That is, for every real number M, there are points a and b in $\mathcal{A}$ such that $f(a) \leq M \leq f(b)$. Use the intermediate value theorem to prove that f is surjective.

Exercise 8: Give an example of a function $f : \mathbb{R} \to \mathbb{R}$ that is strictly increasing but not surjective.

Exercise 9: Find a bijection $f : (-1, 1) \to (4, 10)$. (This can be done with a continuous function).

Exercise 10: Find a bijection $f : (0, 1) \to \mathbb{R}$. (This can be done with a continuous function).

Exercise 11: Find a bijection $f : (0, 1) \to [0, \infty)$. Can this be done with a continuous function?

Exercise 12: Find a bijection $f : [0, 1) \to \mathbb{R}$. Can this be done with a continuous function?

Lemma 2: Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be three abstract sets.

a) If $f : \mathcal{A} \to \mathcal{B}$ and $g : \mathcal{B} \to \mathcal{C}$ are two injective functions, then the function $h : \mathcal{A} \to \mathcal{C}$ defined by $h(a) = g(f(a))$ is also injective.

b) If $f : \mathcal{A} \to \mathcal{B}$ and $g : \mathcal{B} \to \mathcal{C}$ are two surjective functions, then the function $h : \mathcal{A} \to \mathcal{C}$ defined by $h(a) = g(f(a))$ is also surjective.

c) If $f : \mathcal{A} \to \mathcal{B}$ is a bijection, and if $g : \mathcal{B} \to \mathcal{C}$ is a bijection, then the function $h : \mathcal{A} \to \mathcal{C}$ defined by $h(a) = g(f(a))$ is a bijection.

Proof: Exercise. □

IV. Cardinality

A. Finite sets and infinite sets

Definition 7: A set $\mathcal{A}$ is *finite* if its number of elements is a nonnegative integer.

Definition 8: A set is *infinite* if it is not finite.

A set is finite if and only if it is either empty or can be put into one-to-one correspondence with the set $\{1, 2, \ldots, n\}$ for some positive integer n. A subset of a finite set must also be finite. If a set $\mathcal{A}$ contains an infinite subset, then $\mathcal{A}$ is also infinite.

Let $\mathcal{A}$ be a finite set. Define the *cardinality of* $\mathcal{A}$, denoted $|\mathcal{A}|$, as the number of elements of $\mathcal{A}$. Thus, the cardinality of the empty set is 0, so that $|\phi| = 0$. If a set $\mathcal{A}$ has n elements, where n is a positive integer, then $|\mathcal{A}| = n$. Any two finite sets with the same cardinality can be put into one-to-one correspondence. This property motivates the following definition for general (possibly infinite) sets.

Definition 9: Two nonempty sets $\mathcal{A}$ and $\mathcal{B}$ are said to have the *same cardinality*, written $|\mathcal{A}| = |\mathcal{B}|$, if they can be put into one-to-one correspondence. That is, $|\mathcal{A}| = |\mathcal{B}|$ if and only if there is a bijection $f : \mathcal{A} \to \mathcal{B}$.

It is important to emphasize that, for an infinite set $\mathcal{A}$, the object $|\mathcal{A}|$ is *not* a real number, and is *not* ∞. Indeed, we shall soon find that two distinct infinite sets can have different cardinalities. When $\mathcal{A}$ is infinite, it is not necessary to think of $|\mathcal{A}|$ as an object at all. Rather, one can formally view the statement "$|\mathcal{A}| = |\mathcal{B}|$" as shorthand for the statement "sets $\mathcal{A}$ and $\mathcal{B}$ can be put into one-to-one correspondence."

B. Transitivity of cardinality

Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be three abstract sets.

Lemma 3: If $|\mathcal{A}| = |\mathcal{B}|$ and $|\mathcal{B}| = |\mathcal{C}|$, then $|\mathcal{A}| = |\mathcal{C}|$.

Proof: If $|\mathcal{A}| = |\mathcal{B}|$, then there is a bijection $f : \mathcal{A} \to \mathcal{B}$. If $|\mathcal{B}| = |\mathcal{C}|$, then there is a bijection $g : \mathcal{B} \to \mathcal{C}$. Lemma 2c ensures that the function $h(a) = g(f(a))$ defines a bijection from $\mathcal{A}$ to $\mathcal{C}$. □

C. Examples of infinite sets with the same cardinality

Example 7: Let $\mathcal{A} = \{1, 2, 3, \ldots\}$ and $\mathcal{B} = \{2, 4, 6, \ldots\}$. Then $f : \mathcal{A} \to \mathcal{B}$ defined by $f(a) = 2a$ is a bijection.

Example 8: Intervals $(0, 1)$ and $(2, 8)$ have the same cardinality, which follows from Example 5.

Example 9: Intervals $(-\pi/2, \pi/2)$ and $\mathbb{R}$ have the same cardinality, which follows from Example 6.

Example 10: The sets $[0, 1)$ and $\mathbb{R}$ have the same cardinality by the result of Exercise 12.

Example 11: Let a, b be real numbers that satisfy $a < b$. Then (a, b) and $\mathbb{R}$ have the same cardinality.

Proof: Just define a bijection $f : (a, b) \to \mathbb{R}$, such as $f(x) = \tan(cx + d)$ for suitable values of c and d. □

V. Countably infinite

Recall that $\mathbb{N} = \{1, 2, 3, \ldots\}$ is the set of natural numbers. It is clear that $\mathbb{N}$ is an infinite set. Thus, any set with the same cardinality as $\mathbb{N}$ is also infinite.

Definition 10: A set $\mathcal{A}$ is *countably infinite* if it can be put into one-to-one correspondence with the natural numbers $\mathbb{N}$. That is, if $|\mathcal{A}| = |\mathbb{N}|$.

Thus, if a set $\mathcal{A}$ is countably infinite, then there is a bijective function $f : \mathbb{N} \to \mathcal{A}$ such that:

$$\begin{aligned} f(1) &= a_1 \\ f(2) &= a_2 \\ f(3) &= a_3 \\ \cdots &\quad \cdots \\ f(n) &= a_n \\ \cdots &\quad \cdots \end{aligned}$$

where a_i are distinct values of the set $\mathcal{A}$, and $\cup_{i=1}^{\infty}\{a_i\} = \mathcal{A}$. This allows the elements of $\mathcal{A}$ to be *listed*, so that:

$$\mathcal{A} = \{a_1, a_2, a_3, \ldots\}$$

where a_1 is the first element on the list, a_2 is the second element on the list, and so on. A list formed in this way is infinite, includes no repetitions, includes all elements of $\mathcal{A}$, and has no elements that are *not* in $\mathcal{A}$. In particular, a set $\mathcal{A}$ is countably infinite if and only if it can be written as an infinite list $\mathcal{A} = \{a_1, a_2, a_3, \ldots\}$ that includes each element of $\mathcal{A}$ once and only once.

Examples of countably infinite sets:

1) $\{2, 4, 6, 8, \ldots\}$
2) $\{integers\} = \{0, 1, -1, 2, -2, 3, -3, \ldots\}$
3) $\mathbb{N}^2$ (also written $\mathbb{N} \times \mathbb{N}$), defined by:

$$\mathbb{N}^2 = \{(a, b) : a \in \mathbb{N}, b \in \mathbb{N}\}$$

   (just draw the 2-dimensional grid of such ordered pairs and find a clever way to list them)
4) $\mathbb{N}^N$ for a given positive integer N.
5) $\{\text{rational numbers}\} = \{x \in \mathbb{R} : x = p/q \text{ for some integers } p, q \text{ such that } q \neq 0\}$.
6) The set of all rational numbers in the interval $[n, n + 1)$ for a given integer n.

The last example considers rational numbers in some restricted interval. That this is a countably infinite set can be proven as a special case of the following result. The proof uses a useful technique of *trimming a list* to remove redundancy.

Lemma 4: Every infinite subset of a countably infinite set is countably infinite.

Proof: Let $\mathcal{A}$ be a countably infinite set, and let $\mathcal{B}$ be an infinite subset of $\mathcal{A}$. We want to show that $\mathcal{B}$ is countably infinite. List the elements of $\mathcal{A}$, so that $\mathcal{A} = \{a_1, a_2, a_3, \ldots\}$. Now trim this list by removing those elements that are not in $\mathcal{B}$. The trimmed list contains only elements of $\mathcal{B}$, and contains *all* elements of $\mathcal{B}$. Since the original list had no redundancy, the trimmed list has no redundancy. Finally, this list must be infinite (since $\mathcal{B}$ is infinite). Thus, the set $\mathcal{B}$ is countably infinite. □

A. Unions of countably infinite sets

Lemma 5: If $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$ is an infinite list of finite sets, then $\cup_{i=1}^{\infty}\mathcal{A}_i$ is finite or countably infinite.

Lemma 6: If $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots$ is an infinite list of countably infinite sets, then $\cup_{i=1}^{\infty}\mathcal{A}_i$ is countably infinite.

Proof: (Lemma 5) Without loss of generality, assume each set $\mathcal{A}_k$ is nonempty and write this set as a finite list with $|\mathcal{A}_k|$ elements (where $1 \leq |\mathcal{A}_k| < \infty$). Every positive integer n can be written uniquely as $n = \sum_{k=1}^{b-1}|\mathcal{A}_k| + r$ for some integer $b \geq 1$ and some integer $r \in \{1, \ldots, |\mathcal{A}_b|\}$. Now define the *concatenated list* that first lists all elements of $\mathcal{A}_1$, then all elements of $\mathcal{A}_2$, and so on. Element n on this concatenated list is equal to the rth element of the list for $\mathcal{A}_b$, where r and b are determined from the unique decomposition $n = \sum_{k=1}^{b-1}|\mathcal{A}_k| + r$. From this concatenated list, remove all repetitions. This results is a new list (with no redundancy) that contains all elements of $\cup_{i=1}^{\infty}\mathcal{A}_i$ (and no other elements). This new list is either finite or countably infinite. □

Proof: (Lemma 6) Define $\mathbb{Q}^+$ as the set of all nonnegative rational numbers. We know this set is countably infinite. For each positive integer k, define $\mathbb{Q}_k$ as the set of all rational numbers in the interval $[k-1, k)$. Each set $\mathbb{Q}_k$ is countably infinite, and:

$$\mathbb{Q}^+ = \cup_{k=1}^{\infty}\mathbb{Q}_k$$

That is, $\mathbb{Q}^+$ is itself a countably infinite union of countably infinite sets. For each positive integer k, the set $\mathbb{Q}_k$ can be put into one-to-one correspondence with the set $\mathcal{A}_k$ via an invertible function $g_k : \mathbb{Q}_k \to \mathcal{A}_k$. This defines a surjection $f : \mathbb{Q}^+ \to \cup_{k=1}^{\infty}\mathcal{A}_k$. Specifically, for each $q \in \mathbb{Q}^+$, define $f(q)$ as follows: First determine the unique integer $k \in \{1, 2, 3, \ldots\}$ for which $q \in [k-1, k)$. Call this integer k_q to emphasize dependence on q. Next, define $f(q) = g_{k_q}(q)$. This defines the function $f : \mathbb{Q}^+ \to \cup_{k=1}^{\infty}\mathcal{A}_k$. To show it is a surjection, note that if $a \in \cup_{k=1}^{\infty}\mathcal{A}_k$ then $a \in \mathcal{A}_k$ for some integer $k \geq 1$, and so $g_k^{-1}(a) \in \mathbb{Q}_k \subseteq \mathbb{Q}^+$. Thus, $f(g_k^{-1}(a)) = g_k(g_k^{-1}(a)) = a$.

Now list the elements of $\mathbb{Q}^+$, so that $\mathbb{Q}^+ = \{x_1, x_2, x_3, \ldots\}$. From this, form a new list $\{f(x_1), f(x_2), f(x_3), \ldots\}$. This new list contains all elements of $\cup_{i=1}^{\infty}\mathcal{A}_i$, and only elements of $\cup_{i=1}^{\infty}\mathcal{A}_i$. If there is any redundancy, trim the list to remove the redundancy. □

B. Every infinite set has a countably infinite subset

It can be shown that if $\mathcal{A}$ has the same cardinality as a set that is known to be countably infinite, then $\mathcal{A}$ is countably infinite. Likewise, if $\mathcal{A}$ has the same cardinality as a set that is known to be uncountably infinite, then $\mathcal{A}$ is uncountably infinite. A countably infinite set cannot have the same cardinality as an uncountably infinite set. If $\mathcal{A}$ contains a subset that is uncountably infinite, then $\mathcal{A}$ is uncountably infinite.

Lemma 7: Every infinite set has a countably infinite subset.

Proof: Let $\mathcal{A}$ be an infinite set. Define a countably infinite list of elements of $\mathcal{A}$ as follows: Let x_1 be any element of $\mathcal{A}$. Let x_2 be any element of $\mathcal{A}$ that is not the same as x_1. For each step $k > 2$, let x_k be any element of $\mathcal{A}$ that is not in the set $\{x_1, \ldots, x_{k-1}\}$. Such an element can be found since the original set $\mathcal{A}$ is infinite. This procedure defines a countably infinite list of distinct elements of $\mathcal{A}$. This is a countably infinite subset of $\mathcal{A}$. □

C. Cardinality exercises

Exercise 13: Let $\mathcal{A}$ be a countably infinite list of real numbers that does not include the numbers π and e. Prove that $\mathcal{A} \cup \{\pi, e\}$ is countably infinite.

Exercise 14: Let $\mathcal{A}$ and $\mathcal{B}$ be two disjoint countably infinite sets. Prove that $\mathcal{A} \cup \mathcal{B}$ is countably infinite.

Exercise 15: Let $f : \mathbb{N}^2 \to \mathbb{N}$ be a bijection from $\mathbb{N}^2$ to $\mathbb{N}$. For all $(a, b, c) \in \mathbb{N}^3$, define $h(a, b, c) = f(f(a, b), c)$. Prove that h is a bijection from $\mathbb{N}^3$ to $\mathbb{N}$.

Exercise 16: Show that if a set $\mathcal{A}$ is infinite, then for any object x we have $|\{x\} \cup \mathcal{A}| = |\mathcal{A}|$. That is, adding a single point to an infinite set does not change its cardinality. *Hint: Use Lemma 7.*

Exercise 17: Show that if a set $\mathcal{A}$ is infinite, then for any finite or countably infinite set $\mathcal{B}$ we have $|\mathcal{A} \cup \mathcal{B}| = |\mathcal{A}|$. (This generalizes the previous exercise).

Exercise 18: If $\mathcal{A}$ has an uncountably infinite subset, then $\mathcal{A}$ is uncountably infinite.

VI. Advanced material on mappings, levels of cardinality, and power sets

Material in this section is not part of the EE 503 class and is included only for your interest.

A. Forming bijections from injections both ways

In the following, let $\mathcal{A}$ and $\mathcal{B}$ represent two abstract sets.

Lemma 8: If there exists an injective function $f : \mathcal{A} \to \mathcal{B}$, then there exists a surjective function $g : \mathcal{B} \to \mathcal{A}$.

Proof: Let a^* be an element of $\mathcal{A}$, and let $f : \mathcal{A} \to \mathcal{B}$ be an injection. Define the surjection $g : \mathcal{B} \to \mathcal{A}$ as follows: For all $b \in \mathcal{B}$ that get mapped to from f, define $g(b)$ as the unique $a \in \mathcal{A}$ that satisfies $f(a) = b$ (this a is unique because f is injective). For all other $b \in \mathcal{B}$, define $g(b) = a^*$. It is easy to see that $g(b)$ is a surjection. □

Theorem 1: (Cantor-Bernstein-Schroeder) If there exists an injective function $f : \mathcal{A} \to \mathcal{B}$ and another injective function $g : \mathcal{B} \to \mathcal{A}$, then there exists a bijection $h : \mathcal{A} \to \mathcal{B}$ (so that $\mathcal{A}$ and $\mathcal{B}$ have the same cardinality).

Proof: Omitted for brevity. □

B. Forming bijections from surjections both ways via the axiom of choice

The axiom of choice says that if we have an uncountably infinite collection of sets, then we can choose a representative element from each set.[4] This is a seemingly mild axiom that is used when making an uncountably infinite number of choices. However, it can lead to surprising and paradoxical results (see Section VII-D). Thus, it is best to invoke the axiom of choice only when needed, and to be explicit about where it is used.

Lemma 9: If there exists a surjective function $f : \mathcal{A} \to \mathcal{B}$, then, *with the axiom of choice*, there exists an injective function $g : \mathcal{B} \to \mathcal{A}$.

Proof: (Lemma 9) Let $f : \mathcal{A} \to \mathcal{B}$ be a surjection. Use the axiom of choice to construct an injection $g : \mathcal{B} \to \mathcal{A}$ as follows: For each $b \in \mathcal{B}$, define $\mathcal{A}_b = \{a \in \mathcal{A} : f(a) = b\}$. The set $\mathcal{A}_b$ is nonempty for all $b \in \mathcal{B}$ because f is a surjection. Further, it is clear that if b_1 and b_2 are distinct elements of $\mathcal{B}$, then sets $\mathcal{A}_{b_1}$ and $\mathcal{A}_{b_1}$ are disjoint. We now have to choose, for each $b \in \mathcal{B}$, an element in $\mathcal{A}_b$. Call this element $g(b)$. The only reason this is difficult is that the set $\mathcal{B}$ may be uncountably infinite, and so we use the axiom of choice to make these uncountably infinite choices. The the function $g(b)$ maps $\mathcal{B}$ into $\mathcal{A}$, and it is injective. Indeed, if b_1 and b_2 are elements of $\mathcal{B}$ that satisfy $g(b_1) = g(b_2)$, then $g(b_1) \in \mathcal{A}_{b_1}$ and $g(b_1) = g(b_2) \in \mathcal{A}_{b_2}$, so $\mathcal{A}_{b_1}$ and $\mathcal{A}_{b_2}$ are not disjoint, and so $b_1 = b_2$. □

Lemma 10: If there is a surjective function $f : \mathcal{A} \to \mathcal{B}$ and another surjective function $g : \mathcal{B} \to \mathcal{A}$, then, *with the axiom of choice*, there exists a bijection $h : \mathcal{A} \to \mathcal{B}$ (so that $\mathcal{A}$ and $\mathcal{B}$ have the same cardinality).

Proof: This follows immediately from Theorem 1 and Lemma 9. □

C. Levels of cardinality

Definition 11: We say that the *cardinality of $\mathcal{A}$ is less than or equal to the cardinality of $\mathcal{B}$*, written $|\mathcal{A}| \leq |\mathcal{B}|$, if there exists an injective function $f : \mathcal{A} \to \mathcal{B}$.

Definition 12: We say that the *cardinality of $\mathcal{A}$ is strictly less than the cardinality of $\mathcal{B}$*, written $|\mathcal{A}| < |\mathcal{B}|$, if there exists an injective function $f : \mathcal{A} \to \mathcal{B}$, but there is no injective function $g : \mathcal{B} \to \mathcal{A}$.[5]

The above definitions implicitly assume the sets $\mathcal{A}$ and $\mathcal{B}$ are nonempty (else, we cannot define a valid function between them). We can formally define the cardinality of the empty set to be strictly less than the cardinality of any nonempty set, so that for any set $\mathcal{A}$ we have: $|\phi| \leq |\mathcal{A}|$, with equality if and only if $\mathcal{A} = \phi$.

Lemma 11: If $\mathcal{A}$ and $\mathcal{B}$ are finite sets, then:

- $|\mathcal{A}| = |\mathcal{B}|$ if and only if $\mathcal{A}$ and $\mathcal{B}$ have the same number of elements.
- $|\mathcal{A}| \leq |\mathcal{B}|$ if and only if the number of elements of $\mathcal{A}$ is less than or equal to the number of elements of $\mathcal{B}$.
- $|\mathcal{A}| < |\mathcal{B}|$ if and only if the number of elements of $\mathcal{A}$ is less than the number of elements of $\mathcal{B}$.
- If $\mathcal{C}$ is an infinite set then $|\mathcal{A}| < |\mathcal{C}|$.

Proof: Exercise. □

[4]By convention, it is assumed we can always choose a representative element from a finite or countably infinite collection of sets, even without the axiom of choice.

[5]By Lemma 1, the statement $|\mathcal{A}| < |\mathcal{B}|$ is equivalent to the statement that there is an injection from $\mathcal{A}$ to $\mathcal{B}$, but no bijection.

Lemma 12: If $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are abstract sets then:[6]

- $|\mathcal{A}| = |\mathcal{B}|$ implies $|\mathcal{A}| \leq |\mathcal{B}|$.
- $|\mathcal{A}| < |\mathcal{B}|$ implies $|\mathcal{A}| \leq |\mathcal{B}|$.
- $|\mathcal{A}| \leq |\mathcal{B}|$ and $|\mathcal{B}| \leq |\mathcal{A}|$ implies $|\mathcal{A}| = |\mathcal{B}|$.
- If $|\mathcal{A}| \leq |\mathcal{B}|$ and $|\mathcal{B}| \leq |\mathcal{C}|$, then $|\mathcal{A}| \leq |\mathcal{C}|$.
- If $|\mathcal{A}| \leq |\mathcal{B}|$ and $|\mathcal{B}| < |\mathcal{C}|$, then $|\mathcal{A}| < |\mathcal{C}|$.
- If $|\mathcal{A}| < |\mathcal{B}|$ is true, then $|\mathcal{B}| \leq |\mathcal{A}|$ is false.
- If $\mathcal{A} \subseteq \mathcal{B}$ then $|\mathcal{A}| \leq |\mathcal{B}|$.

Proof: Exercise. □

Lemma 13: $|\mathbb{N}| < |\mathbb{R}|$.

Proof: Exercise. □

D. Power sets

Let $\mathcal{A}$ be a finite nonempty set. Consider the set of all binary-valued functions $f : \mathcal{A} \to \{0, 1\}$. There are exactly $2^{|\mathcal{A}|}$ such functions. For example, if $\mathcal{A} = \{1, 2\}$, then there are 4 such functions:

- Function 1: $(f(1), f(2)) = (0, 0)$
- Function 2: $(f(1), f(2)) = (0, 1)$
- Function 3: $(f(1), f(2)) = (1, 0)$
- Function 4: $(f(1), f(2)) = (1, 1)$

Now let $\mathcal{A}$ be a general nonempty set (possibly infinite). Each specific binary-valued function $f : \mathcal{A} \to \{0, 1\}$ can be viewed as defining a *subset* $\mathcal{B}_f \subseteq \mathcal{A}$ by including each element $a \in \mathcal{A}$ in the set $\mathcal{B}_f$ if and only if $f(a) = 1$. The subset $\mathcal{B}_f$ defined by the all-zero function is the empty set ϕ. It is easy to see that the set of all such functions can be put into one-to-one correspondence with the *set of all subsets of* $\mathcal{A}$ (where the set of all subsets includes the empty set ϕ).

Definition 13: Let $\mathcal{A}$ be a nonempty set. The *power set of* $\mathcal{A}$, written $2^{\mathcal{A}}$, is the set of all subsets of $\mathcal{A}$ (including the empty set).

For example, if $\mathcal{A} = \{1, 2, 3\}$, then:

$$2^{\mathcal{A}} = \{\phi, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

The next fundamental result shows that the power set of any nonempty set gives another set with strictly larger cardinality.

Theorem 2: (Cantor's theorem for power sets) If $\mathcal{A}$ is a nonempty set, then $|\mathcal{A}| < |2^{\mathcal{A}}|$.

Proof: Let $\mathcal{A}$ be a nonempty set. It is easy to show that there is an injection from $\mathcal{A}$ to $2^{\mathcal{A}}$. To show that there is no injection from $2^{\mathcal{A}}$ to $\mathcal{A}$, suppose there is (we reach a contradiction). Lemma 8 implies there is a surjection $g : \mathcal{A} \to 2^{\mathcal{A}}$. For each $a \in \mathcal{A}$, $g(a)$ is a subset of $\mathcal{A}$. Thus, it makes sense to ask whether or not $a \in g(a)$. Define:

$$\mathcal{B} = \{a \in \mathcal{A} : a \notin g(a)\}$$

Thus $\mathcal{B} \in 2^{\mathcal{A}}$. The proof is now almost finished. The remaining steps are exercises:

- Argue that there is an element $c \in \mathcal{A}$ such that $g(c) = \mathcal{B}$.
- Ask the question: Is it true that $c \in \mathcal{B}$?
- Ask the question: Is it true that $c \notin \mathcal{B}$?

□

In summary, the above proof shows that if $\mathcal{A}$ is a nonempty set, then it is impossible to find a surjection from $\mathcal{A}$ to $2^{\mathcal{A}}$.

Corollary 1: Let $\mathcal{A}_c$ be a collection of sets defined for all $c \in \mathcal{C}$, where $\mathcal{C}$ is some general index set. Then for each $c \in \mathcal{C}$ we have:

$$|\mathcal{A}_c| < |2^{\cup_{d \in \mathcal{C}} \mathcal{A}_d}|$$

Proof: Exercise. *Hint: Note that for all* $c \in \mathcal{C}$ *we have* $\mathcal{A}_c \subseteq \cup_{d \in \mathcal{C}} \mathcal{A}_d$. □

[6]Given any two nonempty sets $\mathcal{A}$ and $\mathcal{B}$, is it necessarily true that either $|\mathcal{A}| \leq |\mathcal{B}|$ or $|\mathcal{B}| \leq |\mathcal{A}|$? A result of Hartogs states that this is always true if we use the axiom of choice (and is *equivalent* to the axiom of choice).

The above corollary shows that if we have any (possibly uncountably infinite) collection of sets, then we can always define a new set with a strictly larger cardinality than every set in the collection. This leads to a strange paradox: It is impossible to define a collection of representative sets $\mathcal{A}_c$ for all distinct cardinalities, since from any collection of sets we can define a new set with cardinality strictly larger than all those in the collection! This is discussed more in Section VII.

E. Related exercises

Exercise 19: Show that $|[0,1]| = |\mathbb{R}|$.

Exercise 20: Show that $|\mathbb{R}^2| = |\mathbb{R}|$. It can also be shown that $|\mathbb{R}^N| = |\mathbb{R}|$ for any given positive integer N.

Exercise 21: Show that if $|\mathcal{A}| \leq |\mathcal{B}|$, then $|2^{\mathcal{A}}| \leq |2^{\mathcal{B}}|$.

Exercise 22: Show that if $|\mathcal{A}| = |\mathcal{B}|$ then $|2^{\mathcal{A}}| = |2^{\mathcal{B}}|$.

Exercise 23: Define $\mathbb{R}^{\mathbb{N}}$ as the set of all infinite sequences $(x_1, x_2, x_3, \ldots)$, where $x_i \in \mathbb{R}$ for all $i \in \{1, 2, 3, \ldots\}$. Prove that $|\mathbb{R}| = |\mathbb{R}^{\mathbb{N}}|$.

Exercise 24: Let $\mathcal{A}$ be the set of all continuous functions $f : \mathbb{R} \to \mathbb{R}$. For example, the functions $\sin(x)$, $x^2+2.3$, and $e^x + 1/(x^2+1)$ are all in the set $\mathcal{A}$. Use the result of the previous exercise to show that $|\mathcal{A}| = |\mathbb{R}|$. *Hint: A continuous function $f : \mathbb{R} \to \mathbb{R}$ is defined by its values on the rational numbers $\mathbb{Q}$.*

Exercise 25: Prove that $|2^{\mathbb{N}}| = |\mathbb{R}|$.

VII. Paradoxes

Material in this section is not part of the EE 503 class and is included only for your interest.

A. The cardinality paradox

Theorem 2 implies that there are an infinity of levels of infinity, since:

$$|\mathbb{N}| < |2^{\mathbb{N}}| < |2^{2^{\mathbb{N}}}| < |2^{2^{2^{\mathbb{N}}}}| < \cdots$$

In particular, we can iteratively take power sets to get infinite sets with strictly growing cardinality. This procedure produces a countably infinite list of infinite sets, and hence identifies a countably infinite number of different levels of infinity. This suggests that the "set" of all levels of infinity is a countably infinite set. *This is not true*. The "set" of all levels of infinity is larger than a power set of the naturals, a power set of a power set of the naturals, and so on. In fact, the "set" of all levels of infinity is so large that it is too large to be called a set!

For a partial understanding of this paradox, let us define two sets to be in the same *equivalence class* if they have the same cardinality. Suppose we can define a set $\mathcal{C}$ of all equivalence classes. Then each $c \in \mathcal{C}$ represents a particular cardinality. For each $c \in \mathcal{C}$, let $\mathcal{A}_c$ be a representative set in equivalence class c (choosing such representatives $\mathcal{A}_c$ for all $c \in \mathcal{C}$ requires the axiom of choice). By the corollary to Cantor's power set theorem (Corollary 1), we have that for each $c \in \mathcal{C}$:

$$|\mathcal{A}_c| < |2^{\cup_{d \in c} \mathcal{A}_d}| \tag{11}$$

However, $2^{\cup_{d \in c} \mathcal{A}_d}$ is a set, so it must be in some equivalence class $c^* \in \mathcal{C}$. Thus, the representative set $\mathcal{A}_{c^*}$ for equivalence class c^* has the same cardinality as $2^{\cup_{d \in c} \mathcal{A}_d}$. That is, $|\mathcal{A}_{c^*}| = |2^{\cup_{d \in c} \mathcal{A}_d}|$. This contradicts (11). The contradiction points to the impossibility of treating all cardinalities as a set and/or the impossibility of finding representative sets for each cardinality. In a sense, we are not "allowed" to define a set of all cardinalities.

B. Russell's paradox

The definition of a power set shows that sets can be members of other sets. Are we "allowed" to define the set of all sets? If so, then this set must contain itself as a member! Are we "allowed" to define sets that contain themselves as members? Here is a famous thought experiment, called *Russell's paradox*, that shows that doing so leads to a contradiction.

Suppose we are allowed to define sets that contain themselves as members. Then we can also talk about sets that do *not* contain themselves as members. Define the following set $\mathcal{B}$:

$$\mathcal{B} = \{\text{all sets that do not contain themselves as members}\}$$

Now ask the question: Does $\mathcal{B}$ contain itself as a member? If "yes," then this leads to a contradiction. If "no," then this also leads to a contradiction.

Russell's paradox prompted the development of a more extensive set of axioms to define sets, typically called the ZF or ZFC axioms (for Zermelo-Fraenkel or Zermelo-Fraenkel with Choice). These axioms require sets to be built with more structure in order to avoid things like Russell's paradox. A discussion of these axioms is beyond the scope of these notes.

C. Gödel's theorem

Consider the following statement: "This statement is unprovable." Is this statement false? If so, then it must *not* be unprovable. Thus, it must be provable, which means it must be true. This proves it is true...which contradicts the claim that the statement is unprovable! Contradictions abound.

This thought experiment points to the need for being precise about what a statement is, whether a statement can be described as "true" or "false," what constitutes a valid proof, and whether or not everything that is provable is actually true. These questions motivate Gödel's work. This section gives a high level discussion of his work. Suppose we have a collection of axioms A and a collection of accepted logical manipulations M that allow one to derive new results from the axioms and from previously derived results. Then one can define the set P of all statements that can be derived from the axioms A via accepted logical manipulations in M. Gödel worked with axioms A for arithmetic and accepted manipulations M being standard arithmetic rules. The axioms A and rules M for arithmetic are defined to be *consistent* if every statement that can be proven from them (that is, every statement in P) is actually true.

Gödel asked two questions:

- Question 1: Is arithmetic consistent? That is, are all statements in P true?
  Gödel's answer: We don't know.
- Question 2: Assuming consistency, are all true arithmetic statements in the set P?
  Gödel's answer: No.

For the solution of question 2, Gödel constructed mappings between numbers and symbols that allowed arithmetic statements to have a higher level semantic interpretation. He then constructed a particular (lengthy) arithmetic statement with a semantic interpretation as follows:

"This statement is not in the set P"

Suppose arithmetic is consistent, so everything in P is true. Then if the above statement is false, it is *not* not in P, so it is in P, so it is true (by consistency of arithmetic), a contradiction. So the above statement (and hence the corresponding arithmetic statement) must be true. This statement, called *Gödel's sentence*, is a true statement of arithmetic that is not in the set P.

This does not mean that Gödel's statement is *true but unprovable* (since he indeed proved it). Rather, it is *true but unprovable via the axioms A and manipulations M*. Gödel's logic transcended the predefined axioms and manipulations. He also showed that the same procedure can be repeated if one attempts to augment the axioms and manipulations with additional ones (provided a certain structure is upheld). Overall, it seems that, once axioms and manipulations are fixed, there are true statements which they cannot prove.

D. The prisoner's hat problem

The following is a famous example that demonstrates unexpected results of the axiom of choice. A more famous example is the *Banach-Tarski paradox*, but this one about hats is easier to understand.

Suppose a countably infinite number of prisoners, labeled $\{1, 2, 3, \ldots\}$, will each be given a hat. Each hat has a certain *color* in a set $\mathcal{C}$ of colors. Assume the set $\mathcal{C}$ contains at least two colors, but is otherwise arbitrary (it can be countably or uncountably infinite). Let c_i be the hat color of prisoner i. Thus, the hat colors are described by an infinite list $(c_1, c_2, c_3, \ldots)$, where $c_i \in \mathcal{C}$ for all $i \in \{1, 2, 3, \ldots\}$. Each prisoner will be able to see the hat color of all other prisoners, but cannot see his own hat color. Each prisoner i is then given a chance to guess his own hat color, based only on knowledge of c_j for all $j \neq i$. If the guess is *exactly correct*, the prisoner will go free. Else, he will be executed. Before hats are given out, prisoners are allowed to agree on a guessing strategy. However, they cannot communicate after the hats are given, and they do not observe anything else (except for everyone else's hat)

once hats are given. Design a guessing strategy that ensures at most a finite number of prisoners will be executed, regardless of the hat color sequence.

This seems impossible, even if there are only two colors! For example, suppose $\mathcal{C} = \{red, blue\}$, and suppose the hats are given out according to an independent and identically distributed (i.i.d.) sequence where each color is equally likely. Then seeing all other hats seems to provide no information about your own hat! Intuitively, in this situation one expects the probability of guessing correctly to be 1/2, regardless of the guess. This leads one to conclude that, with nonzero probability, an infinite number of prisoners will be executed. *This is true if the guessing function is probabilistically measurable, but not true in general* (see appendix).

Remarkably, it is possible to solve this problem (via the axiom of choice) for any arbitrarily large cardinality for $\mathcal{C}$ (for example, the prisoners might be expected to guess a real number). Here is how: Let $\mathcal{A}$ be the set of all possible infinite sequences $(c_1, c_2, c_3, \ldots)$, where $c_i \in \mathcal{C}$ for all $i \in \{1, 2, 3, \ldots\}$. Divide this set $\mathcal{A}$ into disjoint *equivalence classes*, where two sequences are in the same equivalence class if and only if they differ in at most a finite number of terms. Let $\mathcal{X}$ denote the set of equivalence classes (this is an uncountably infinite set). For each equivalence class $x \in \mathcal{X}$, use the axiom of choice to choose a representative sequence $\hat{c}(x) = (\hat{c}_1(x), \hat{c}_2(x), \hat{c}_3(x), \ldots)$ in that equivalence class. Before hats are given out, the prisoners agree on the representative sequences $\hat{c}(x)$ to use for each equivalence class $x \in \mathcal{X}$. The guessing strategy is then as follows:

- After all hats are given out, each prisoner i observes c_j for all $j \neq i$.
- From this, all prisoners correctly determine the equivalence class x of the hat color sequence.
- Each prisoner i gives the guess "$\hat{c}_i(x)$."

The second step is possible because each prisoner can determine the equivalence class by knowing all but one term of the hat color sequence. The resulting sequence of guesses is in the same equivalence class as the true hat color sequence, and so it differs by at most a finite number of terms. Thus, at most a finite number of prisoners will be executed.

In the special case when hats are given i.i.d. with only 2 colors red and blue, this proves the following: Let $\mathcal{S}$ denote the set of all possible sequences $c = (c_1, c_2, c_3, \ldots)$. For each sequence $c \in \mathcal{S}$, let $x(c)$ denote its equivalence class. Then for any representative sequences $\hat{c}(x)$ defined for all $x \in \mathcal{X}$, there is an $i \in \{1, 2, 3, \ldots\}$ for which the following set is not probabilistically measurable:

$$\{c \in \mathcal{S} : \hat{c}_i(x(c)) = blue\}$$

The above set represents the set of all sequences that lead prisoner i to guess "blue." Thus, the probability of guessing "blue" is *not* 1/2. Rather, it unmeasurable.

APPENDIX

This appendix material is not part of the EE 503 class and is included only for your own interest.

This appendix shows that, for the prisoner hat problem with i.i.d. hats of two colors, measurable guessing leads to an infinite number of executions. The section uses techniques of probability that we will learn in the EE 503 class (such as indicator functions, the law of total expectation, and computing probabilities from limits of expanding sets). If all prisoners use guesses that are deterministic and measurable functions of their observations, then each has a 1/2 probability of being executed.[7] Of course, the execution events may or may not be independent for each prisoner. Let $N(k)$ be the random number of executions in the first k prisoners, so that:

$$N(k) = Y_1 + Y_2 + \cdots + Y_k$$

where each Y_k is an indicator variable that is 1 if prisoner k is executed, and 0 else. Clearly $0 \leq N(k) \leq k$. Also:

$$\mathbb{E}\left[N(k)\right] = \mathbb{E}\left[Y_1\right] + \cdots + \mathbb{E}\left[Y_k\right] = k/2$$

[7] In fact, this 1/2 result is nontrivial to prove because independence is defined in terms of finite sets of the random variables, and a limiting argument is needed to allow measurable functions of an *infinite* number of the observations. For a given prisoner i, we can define $\mathcal{E}_i$ as the set of all sequences that lead prisoner i to guess BLUE. Then we must show $Pr[\{c_i = RED\} \cap \mathcal{E}_i] = (1/2)Pr[\mathcal{E}_i]$. On the other hand, if $\mathcal{D}_i$ is any event that is completely defined by the random variables $\{c_1, \ldots, c_{i-1}, c_{i+1}, \ldots c_n\}$ for some finite integer $n > i$, then we immediately get $Pr[\{c_i = RED\} \cap \mathcal{D}_i] = Pr[c_i = RED]Pr[\mathcal{D}_i] = (1/2)Pr[\mathcal{D}_i]$ by definition of independence. So we need to describe $\mathcal{E}_i$ in terms of limits, unions, and intersections of events of type $\mathcal{D}_i$.

Notice that the above uses the fact that the expectation of a sum is the same as the sum of the expectation, regardless of whether or not the random variables are independent. Then, by the law of total expectation:

$$\begin{aligned} k/2 &= \mathbb{E}\left[N(k)|N(k) \leq k/4\right] Pr[N(k) \leq k/4] + \mathbb{E}\left[N(k)|N(k) > k/4\right](1 - Pr[N(k) \leq k/4]) \\ &\leq (k/4)Pr[N(k) \leq k/4] + k(1 - Pr[N(k) \leq k/4]) \\ &= -(3/4)kPr[N(k) \leq k/4] + k \end{aligned}$$

Thus, for all positive integers k, we have $Pr[N(k) \leq k/4] \leq 2/3$.

Now let N be the total number of executions (considering all of the infinite number of prisoners). Then $N(k) \leq N$ for all k, and so:

$$\{N \leq k/4\} \subseteq \{N(k) \leq k/4\}$$

Thus, for all positive integers k:

$$Pr[N \leq k/4] \leq 2/3$$

In particular, letting $k = 4n$ gives for all positive integers n:

$$Pr[N \leq n] \leq 2/3$$

Thus:

$$\lim_{n\to\infty} Pr[N \leq n] \leq 2/3$$

Formally, notice that:

$$\{N < \infty\} = \cup_{i=1}^{\infty}\{N \leq i\}$$

Thus, by standard limit theory for probability:

$$Pr[N < \infty] = \lim_{n\to\infty} Pr[\cup_{i=1}^{n}\{N \leq i\}] = \lim_{n\to\infty} Pr[N \leq n] \leq 2/3$$

So, with probabilistically measurable guesses, with probability at least 1/3, an infinite number of prisoners will be executed.