# Chapter 2: Set Theory

Notes by Lansothungg Humtsoe

## 1 Basics

We begin these notes on set theory with a definition of a set, and the basic notation we use to represent sets.

**Definition 1.** A *set* $X$ is a collection of elements from a known universe $\Omega$.

We have seen sets crop up here and there before. In discussion variables used in propositional formulae, we used the notation $x \in X$ to denote that $x$ is a member of a set $X$. We similarly used this notation in constructing $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$. We use the notation $x \notin X$ to denote that $x$ is not a member of a set $X$.

We have many different kinds of notations to describe sets. Certainly we have seen language used to describe sets, such as phrases like

> Let $\mathbb{N}$ denote the set of natural numbers, as constructed by the Peano Axioms.

In addition, we have notational structures for describing sets more generally. There are several ways to do this. In general, one point of agreement is that sets are always denoted inside of curly braces $\{\}$. This distinguishes sets from other collections of numbers, such as, say sequences, which are denoted inside parentheses $()$ and have an order imposed on the elements; sets are always unordered unless otherwise specified.

Let's look at some examples of defining sets. First, if possible, we could simply list all the elements in a set. Consider:

$$X = \{2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22\}.$$

This tells us that $X$ is the set whose elements are all the even numbers from 2 up to 22 (inclusive). As noted, sets are unordered objects, so we could equally well have written

$$X = \{2, 10, 18, 4, 12, 20, 6, 14, 22, 8, 16\},$$

and the set $X$ would be no different. In addition, each element of the universe is either in a set or not; its inclusion is binary. Hence, listing elements more than once also does not change the set. So, for example, we could write

$$X = \{2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22\},$$

and the set $X$ still is no different.

Now, suppose we wished to define a set $Y$ containing all the even numbers from 2 up to 5274. Obviously we would not like to list all these numbers! So we need a different way to describe $Y$. One obvious approach is to use ... here, so that we can yada-yada our way through. So we could write $Y$ as

$$Y = \{2, 4, 6, 8, \ldots, 5274\}.$$

In general, this is an acceptable description of $Y$, sometimes called an *implied list*. Also in general, I would warn you to be careful with implied lists, as context always matters. For example, in the above implied list,

you expect your reader to make the assumption that the yada-yada means to consider all even numbers up to 5274. However, perhaps what you really intended to list is "all even numbers that are either powers of 2 or divisible by 3." Or maybe what you are listing is really "integers between 2 and 5274 that are not odd primes." Certainly every number that is written in this presentation of $Y$ qualifies under either of these definitions, so it is of critical import that your specific intention is clear from context here.

To avoid this kind of potential pitfall, we have yet another notation that can be useful to describe sets, called *set-builder notation*. For set-builder notation, we describe a set in two parts: first, the universe from which the numbers come, and second, the rules for belonging to the set. For example, if we wanted to clarify that $Y$ is all the positive even numbers up to 5274, we could write

$$Y = \{x \in \mathbb{N} \mid x \text{ is even and } x \leq 5274\}.$$

Here, we see that the first part of the notation, $x \in \mathbb{N}$, describes where our numbers come from (that is, the positive integers). The second part of our notation describes the rules for being a member of $Y$: any member of $Y$ must be even and no larger than 5274. The central bar in this notation is sometimes written instead as a :, and is usually read as "such that," so that speaking this presentation, I would say "$Y$ is the set of natural numbers $x$ such that $x$ is even and $x$ is at most 5274."

Using set-builder notation, notice that the description to the right of the central bar is in fact a proposition about $x$, that given a member of the range $\mathbb{N}$ can be true or false. In general, this is how set-builder notation works. Given a proposition $p(x)$, where $x$ is a variable whose range is $\Omega$, we write

$$X = \{x \in \Omega \mid p(x)\}$$

to denote the set of elements in the universe $\Omega$ for which $p(x)$ is true. Using this type of notation, we have no ambiguity about what elements are in a set, since for each $x \in \Omega$ we clearly either have $p(x)$ true or false.

We note that in the first part of set-builder notation, the range specified need not be the entire universe of numbers from which we operate. It can be any set of numbers as well.

Before we go on to discuss more on this, let's have one more basic definition of an object that shows up all the time: the empty set.

**Definition 2.** The *empty set* $\emptyset$ is a set containing no elements; that is, $\forall x \in \Omega, x \notin \emptyset$.

While having a set of nothing may seem suspicious and strange, the empty set will appear from time to time as necessary. For example, we will see a version of "subtraction" for sets, and we need a way to denote what happens when you "subtract" everything. Same for "addition:" we need a set to act as an identity element for these types of operations. The empty set will perform that role for us.

## 2 Subsets

**Definition 3.** Let $A$ and $B$ be sets in universe $\Omega$. We say $A$ is a *subset* of $B$ if $\forall x \in \Omega, (x \in A) \Rightarrow (x \in B)$. We write $A \subseteq B$ to denote that $A$ is a subset of $B$. If $A$ is a subset of $B$, we call $B$ a *superset* of $A$, and write $B \supseteq A$ to denote that $B$ is a superset of $A$.

Let's look at an example of using set-builder notation to define a subset. Let's take $Y$, as above, to be the set

$$Y = \{x \in \mathbb{N} \mid x \text{ is even and } x \leq 5274\}.$$

Now, suppose that we wanted to isolate only those members of $Y$ that are divisible by 3. We could create a new set $Z$, explicitly to be a subset of $Y$, as

$$Z = \{x \in Y \mid x \text{ is divisible by } 3\}.$$

Here, we have used the set $Y$ as the range to consider when constructing the set $Z$, so only members of $Y$ can be elements of $Z$. By definition, we thus have $Z \subseteq Y$.

Let's consider an example of showing one set is a subset of another, following this definition.

**Example 1.** Let $X = \{x \in \mathbb{Z} \mid x \text{ is even}\}$ and let $Y = \{x \in \mathbb{Z} \mid x = 4k + 2 \text{ for some } k \in \mathbb{Z}\}$. Then $Y \subseteq X$.

**Proof.** Per the definition of subset, we wish to show that $x \in Y \Rightarrow x \in X$. We work by direct proof.
Suppose $x \in Y$, so that there exists $k \in \mathbb{Z}$ such that $x = 4k + 2$. Then $x = 2(2k + 2)$, and hence $x$ is even. By definition, then, $x \in X$. Therefore, $x \in Y \Rightarrow x \in X$, and thus $Y \subseteq X$. □

We note, moreover, that by definition, $\emptyset \subseteq X$ for every set $X$. This is indeed definitional, as $x \in \emptyset$ is always false, and hence $x \in \emptyset \Rightarrow x \in X$ is always true. This is perhaps an example of a larger truth about the empty set: given any proposition $p(x)$, the statement

$$\forall x \in \emptyset, p(x)$$

is always true. This is sometimes called *vacuous truth*; a statement about no numbers cannot ever be false. (You can think of the phrase "vacuous truth" as something like "truth in a vacuum," in which any statement can be made because there is nothing to make it about.)

We note that the symbol $\subset$ is sometimes used in place of $\subseteq$ to indicate that equality is impossible. As with $<$ and $\leq$, the difference between the two symbols is that in the latter case we allow the two things being compared to be the same, and in the former we force that they are different.

To dig into this a little further, let's consider what it means for two sets to be equal, and how we could prove they are equal if in fact they are. We start with a perhaps trivial definition, which we can then use to think about proof techniques for showing set equalities.

**Definition 4.** Let $A$ and $B$ be sets in universe $\Omega$. We say that $A = B$ if $\forall x \in \Omega, (x \in A) \Leftrightarrow (x \in B)$.

This definition should make sense: sets are equal if their elements are exactly the same. But let's break down that $\Leftrightarrow$ in the definition, recalling that we can think about it as two implicative statements. Let's run down some propositional logic here:

$$\begin{aligned}
A = B &\equiv \forall x \in \Omega, (x \in A) \Leftrightarrow (x \in B) \quad \text{(by definition)} \\
&\equiv \forall x \in \Omega, [(x \in A) \Rightarrow (x \in B)] \wedge [(x \in B) \Rightarrow (x \in A)] \quad \text{(see HW1)} \\
&\equiv [\forall x \in \Omega (x \in A) \Rightarrow (x \in B)] \wedge [\forall x \in \Omega, (x \in B) \Rightarrow (x \in A)] \\
&\equiv [A \subseteq B] \wedge [B \subseteq A]\,. \quad \text{(by definition)}
\end{aligned}$$

Hence, to show that two sets are equal to each other, we can show that each is a subset of the other. Let's take a look at an example of this kind of proof. Throughout these notes, we will use this structure to prove some of the theorems about sets that we care about.

**Example 2.** Let

$$A = \{x \in \mathbb{Z} \mid x = 4k + 3 \text{ for some } k \in \mathbb{Z}\},$$

and let

$$B = \{x \in \mathbb{Z} \mid x = 4k - 1 \text{ for some } k \in \mathbb{Z}\}.$$

Then $A = B$.

**Proof.** We show double containment, as described above.
First, to show that $A \subseteq B$, let $x \in A$. Then there exists $k \in \mathbb{Z}$ such that $x = 4k + 3$. But then $x = 4k + 3 = 4(k + 1) - 1$, and hence by definition $x \in B$. Therefore, $A \subseteq B$.

For the other containment, let $x \in B$. Then there exists $k \in \mathbb{Z}$ such that $x = 4k - 1$. But then $x = 4k - 1 = 4(k - 1) + 3$, and hence by definition $x \in A$. Therefore, $B \subseteq A$. $\square$

We note that thinking about sets from set-builder notation gives an alternative approach to thinking about proving subset relationships and equality. If we write $A = \{x \in \Omega \mid p(x)\}$ and $B = \{x \in \Omega \mid q(x)\}$, where $p(x)$ and $q(x)$ are logical formulae with $x$ in range $\Omega$, then we have

$$\begin{aligned} A \subseteq B &\equiv \forall x \in \Omega,\ (x \in A) \Rightarrow (x \in B) \\ &\equiv \forall x \in \Omega,\ p(x) \Rightarrow q(x). \end{aligned}$$

Likewise, applying the notion of double containment, we can rephrase $A = B$ as $\forall x \in \Omega,\ p(x) \Leftrightarrow q(x)$. If we approach set containment and equality from this perspective, we have a wealth of tools available to prove these types of propositions. Ultimately, these proofs will look quite similar to the types of proofs shown in the previous two examples.

Finally, we note that in some cases, thinking about what the possible subsets of a given set might look like can be interesting. We define this as follows:

**Definition 5.** Let $X$ be a set. The *power set* of $X$, denoted by $\mathcal{P}(X)$ or $2^X$ is the set whose elements are all the possible subsets of $X$. That is to say, $\mathcal{P}(X) = \{A \mid A \subseteq X\}$.

Here, in set-builder notation, we have not specified a universe for $A$. We note that the universe for $A$ is distinct from the universe for the elements of $X$; if $X$ contains, say, real numbers, then $\mathcal{P}(X)$ contains SETS of real numbers as its elements. For example, if

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

then we have

$$\begin{aligned} \mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}, \{1,2,3\}, \\ \{1,2,4\}, \{1,3,4\}, \{2,3,4\}, \{1,2,3,4\}\}. \end{aligned}$$

That is to say, $\mathcal{P}(X)$ is the set of all possible sets we can make out of the elements of $X$. Notice that, as we discussed above, since $\emptyset \subseteq X$, we must have $\emptyset \in \mathcal{P}(X)$, and likewise $X \in \mathcal{P}(X)$.

It is important to understand the distinction between the universe in which we find the elements of $X$ and the elements of $\mathcal{P}(X)$. For example, if $x \neq \emptyset$, it can NEVER be true that both $x \in X$ and $x \in \mathcal{P}(X)$. It is, however, always true that if $x \in X$, then $\{x\} \in \mathcal{P}(X)$.

# 3 Operations

Now that we have a sense of what sets look like, and how to think about subsets, let's dive into the kinds of operations we can do on sets. These operations are related to the kinds of operations we perform on propositions, which we shall observe as we go. At times, we will omit the specification of the universe in which we operate, assuming a common universe $\Omega$ for all sets unless otherwise specified.

**Definition 6.** Let $A, B$ be sets. Define the *intersection* of $A$ and $B$, denoted by $A \cap B$, to be the set

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

Notationally, and by the use of the word "and," this may remind you of a conjunction, and indeed it should. If we think of $A$ and $B$ using set builder notation, as follows:

$$A = \{x \mid p(x)\}, \quad B = \{x \mid q(x)\}, \tag{1}$$

then we have

$$A \cap B = \{x \mid p(x) \wedge q(x)\}.$$

Likewise, we can define a set operation that performs the same basic job as a disjunction:

**Definition 7.** Let $A, B$ be sets. Define the *union* of $A$ and $B$, denoted by $A \cup B$, to be the set

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

As intersection is related to conjunction, this is related to disjunction. Indeed, if $A$ and $B$ are as in (**??**), then we have

$$A \cup B = \{x \mid p(x) \vee q(x)\}.$$

As a result of these relationships, we can immediately lift some of the theorems we know about conjunction and disjunction into the world of set operations. In particular, since we know that for any propositions $p, q, r$, that $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$ and $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$, we immediately obtain the following theorem:

**Theorem 1.** *Let $A, B, C$ be sets. Then*

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$*, and*
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$*.*

To ensure that $A \cap B$ and $A \cup B$ are understood, see the Venn diagram and caption in Figure **??**. We shall occasionally use Venn diagrams to convey an understanding of sets and their relationships.

Figure 1: Imagine that the rectangle describes the possible universe $\Omega$. If $A$ is the red set, and $B$ is the blue set, then the purplish set where the two overlap is $A \cap B$. The set $A \cup B$ is all colored portions of the diagram.

Now, we have set operations that capture the propositional operations of conjunction and disjunction, but what about negation? Can we construct a set operation that performs this propositional operation as well? Well, of course. The operation is known as complementation.

**Definition 8.** Let $A$ be a set in the universe $\Omega$. Define the *complement* of $A$ to be the set $A^c$ defined by $x \in A^c \Leftrightarrow x \notin A$.

In the propositional sense, we have that if $A$ is defined as in (**??**), we have

$$A^c = \{x \in \Omega \mid \neg p(x)\}.$$

Note that under this definition, it is necessary to understand the universe from which $A$ comes. That is to say, if we wish to consider all the things that are not in $A$, we need to know all the things that matter. Indeed, if we allow the underlying universe to vary, we end up with a different understanding of complement, known as the relative complement.

**Definition 9.** Let $A, B$ be sets. The *relative complement* of $B$ in $A$, denoted by $A \backslash B$ is the set defined by

$$A \backslash B = \{x \mid x \in A \text{ and } x \notin B\}.$$

That is to say, the relative complement of $B$ in $A$ is the set of elements from $A$ that do not appear in $B$. This is sometimes read as "set minus," that is, $A \backslash B$ is read as "$A$ set-minus $B$." A warning here: there is no rule that $B$ must be contained in $A$ to "subtract" $B$ from $A$. We just take out whatever elements of $B$ *are* there, and ignore the rest. So $A \backslash B = A \backslash (A \cap B)$, since we only concern ourselves with those members of $B$ that are also members of $A$.

Since the operations of union, intersection, and complementation for sets have obvious connections to the operations of conjunction, disjunction, and negation for propositions, we immediately obtain a version of De Morgan's Laws for sets.

**Theorem 2** (De Morgan's Laws for Sets)**.** *Let $A, B$ be sets in the universe $\Omega$. Then*

- $(A \cup B)^c = A^c \cap B^c$, *and*
- $(A \cap B)^c = A^c \cup B^c$.

The proof of this theorem is left as an exercise; it can be approached either from the perspective of propositional logic, or from the perspective of showing two sets are equal by double containment.

Now, it is often the case that we wish to intersect or union more than just one set. To do so, we recursively define the following notation:

Given sets $A_1, A_2, \ldots, A_n$ in a universe $\Omega$, define

$$\bigcup_{i=m}^{k} A_i = \emptyset \text{ if } k < m; \quad \bigcup_{i=m}^{k} A_i = \left(\bigcup_{i=m}^{k-1} A_i\right) \cup A_k \text{ if } k \geq m, \text{ and}$$

$$\bigcap_{i=m}^{k} A_i = \Omega \text{ if } k < m; \quad \bigcap_{i=m}^{k} A_i = \left(\bigcap_{i=m}^{k-1} A_i\right) \cap A_k \text{ if } k \geq m.$$

This notation is similar to the recursive notation we defined for summations and products in the Induction notes. We note that under this definition, we can show the following:

**Proposition 1.** Let $A_1, A_2, \ldots, A_n$ be sets in a universe $\Omega$. Then we have

- $\bigcup_{i=1}^{n} A_i = \{x \in \Omega \mid \exists i \text{ with } 1 \leq i \leq n,\ x \in A_i\}$, and
- $\bigcap_{i=1}^{n} A_i = \{x \in \Omega \mid \forall i \text{ with } 1 \leq i \leq n,\ x \in A_i\}$.

Here, we will prove the first statement, and leave the second as an exercise.

**Proof.** Let $A_1, A_2, \ldots, A_n$ be sets in $\Omega$. We prove that $\bigcup_{i=1}^{n} A_i = \{x \in \Omega \mid \exists i \text{ with } 1 \leq i \leq n,\ x \in A_i\}$ by induction on $n$.

For the base case, when $n = 1$, we have that $\bigcup_{i=1}^{n} A_i = A_1$. On the other hand, $\{x \in \Omega \mid \exists i \text{ with } 1 \leq i \leq 1,\ x \in A_i\} = A_1$, since the only value $i$ can take is 1. Hence, the result holds in the case that $n = 1$.

Now, let us suppose that for some $k \geq 1$, it is true that $\bigcup_{i=1}^{k} A_i = \{x \in \Omega \mid \exists i \text{ with } 1 \leq i \leq k,\ x \in A_i\}$. Consider the case of $k + 1$. We have

$$\begin{aligned}
\bigcup_{i=1}^{k+1} A_i &= \bigcup_{i=1}^{k} A_i \cup A_{k+1} \quad \text{(by definition)} \\
&= \{x \in \Omega \mid \exists i \text{ with } 1 \leq i \leq k,\ x \in A_i\} \cup A_{k+1} \quad \text{(by the inductive hypothesis)} \\
&= \{x \in \Omega \mid \exists i \text{ with } 1 \leq i \leq k,\ x \in A_i \text{ or } x \in A_{k+1}\} \quad \text{(by definition of union)} \\
&= \{x \in \Omega \mid \exists i \text{ with } 1 \leq i \leq k+1,\ x \in A_i\} \quad \text{(since each } x \text{ is in one of } A_1, \ldots, A_k \text{ or in } A_{k+1}).
\end{aligned}$$

Hence, the result also holds for $k + 1$.

By induction, then, for any choice of $n$, we have that $\bigcup_{i=1}^{n} A_i = \{x \in \Omega \mid \exists i \text{ with } 1 \leq i \leq n,\ x \in A_i\}$. □

We can use this more general definition of a multiway union/intersection to develop a more sophisticated set of De Morgan's Laws for sets. The proof of this theorem is a homework exercise, but as with the first version of De Morgan's Laws, it can proven in multiple ways. Induction is an option, as is using Proposition **??** and showing double containment to demonstrate set equality.

**Theorem 3** (De Morgan's Laws for Sets, v. 2)**.** *Let $A_1, A_2, \ldots, A_n$ be sets in the universe $\Omega$. Then*

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

Our final operation on sets to define here is the Cartesian product. This operation is a little different, as the output of a Cartesian product does not live in the same universe as the original sets.

**Definition 10.** Let $A, B$ be sets, from possibly different universes $\Omega_1$ and $\Omega_2$. Define the *Cartesian product* of $A$ and $B$, denoted by $A \times B$, as the set

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

That is to say, the Cartesian product consists of all ordered pairs of elements, in which the first element comes from $A$ and the second element comes from $B$.

# 4 Well-Ordering Principle

We close this foray into basic set theory with a very useful result about natural numbers, known as the Well-Ordering Principle. To introduce the theorem, we first start with few definitions.

**Definition 11.** Let $S \subseteq \mathbb{R}$. We say that $\ell$ is a *lower bound* for $S$ if, for all $x \in S$, we have that $\ell \leq x$. We say that $u$ is an *upper bound* for $S$ if, for all $x \in S$, we have that $x \leq u$.

We note that lower and upper bounds for a set need not exist. For example, if we take $S = \mathbb{Z} \subseteq \mathbb{R}$, then obviously it has no lower bound or upper bound. In general, though, if we do have a lower or upper bound, there is an obvious question: what is the best lower bound we can have?

**Definition 12.** Let $S \subseteq \mathbb{R}$. We say that $m$ is the *greatest lower bound*, or *infimum* for $S$ if

- $m$ is a lower bound for $S$, and
- if $\ell$ is a lower bound for $S$, then $m \geq \ell$.

We write $m = \inf(S)$. If $m$ is an infimum for $S$, and $m \in S$, we say that $m$ is a *minimum* for $S$, and write $m = \min(S)$.

We can likewise define the *least upper bound*, or *supremum* for $S$ to be a number $M$ such that

- $M$ is an upper bound for $S$, and
- if $u$ is an upper bound for $S$, then $M \leq u$.

We write $M = \sup(S)$. If $M$ is a supremum for $S$ and $M \in S$, we say that $M$ is a *maximum* for $S$, and write $M = \max(S)$.

It is important, dealing with infimum and supremum, to ensure that these are well defined. That is, it is important to prove the following:

**Proposition 2.** Let $S \subseteq \mathbb{R}$. Then the infimum and supremum of $S$ are uniquely defined; that is, if $m_1, m_2$ are both infima for $S$, then $m_1 = m_2$, and if $M_1, M_2$ are both suprema for $S$, then $M_1 = M_2$.

**Proof.** We prove the result for the infimum, and note that the proof for supremum is symmetric (and is left as an exercise).

Suppose that $m_1, m_2$ are both infima for $S$. Note that by definition, then, both $m_1$ and $m_2$ are lower bounds for $S$. Moreover, we have that $m_1 \geq \ell$ for every lower bound $\ell$ for $S$. Since $m_2$ is an example of such an $\ell$, we have that $m_1 \geq m_2$. Likewise, since $m_2$ is also an infimum, we have that $m_2 \geq \ell$ for every lower bound $\ell$ for $S$. Taking the lower bound $\ell = m_1$, we thus have that $m_2 \geq m_1$. Hence, since $m_1 \leq m_2 \leq m_1$, we must have equality throughout, and $m_1 = m_2$. □

We note that in all of our four major sets of numbers, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$, we can never guarantee that suprema exist. Subsets in any of these can fail to be bounded above. The same is true for infima in $\mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$, but $\mathbb{N}$ is special in this regard. Certainly, $\mathbb{N}$ itself has a lower bound, so every set that is contained in $\mathbb{N}$ is also bounded below by, say, 1. But more than that, we have that every subset in $\mathbb{N}$ contains its own minimum. This property is extremely useful, as we shall see, and is codified as a theorem known as the Well-Ordering Principle.

**Theorem 4** (Well-Ordering Principle)**.** *Let $S \subseteq \mathbb{N}$ with $S \neq \emptyset$. Then $S$ contains a minimum.*

**Proof.** We work by contradiction. Suppose, then, that $S$ is a subset of $\mathbb{N}$, with $S \neq \emptyset$, and $S$ does not contain a minimum. We show, by strong induction, that $n \notin S$ for all $n \geq 1$.

For the base case, consider $n = 1$. Notice that $1 \leq k$ for every $k \in \mathbb{N}$, and hence it must be that $1 \leq k$ for every $k \in S$. Suppose, for the sake of contradiction, that $1 \in S$. Then if $t \in \mathbb{R}$ is a lower bound for $S$, then $t \leq 1$, and hence we have that 1 is both a lower bound for $S$ and is larger than any other lower bound for $S$, which implies that $1 = \inf(S)$. But then we have $1 = \min(S)$, since 1 is an infimum contained in the set. By supposition, though, $S$ does not contain a minimum, so it must be impossible that $1 \in S$. Therefore, $1 \notin S$.

Now, let us take for the strong inductive hypothesis that for some $n \geq 1$, we have that $k \notin S$ for all $1 \leq k \leq n$. Consider $n + 1$; suppose, for the sake of contradiction, that $n + 1 \in S$. Note that as $k \notin S$ for all $k < n + 1$, we therefore have that $n + 1 \leq k$ for all $k \in S$, so $n + 1$ is a lower bound for $S$. Moreover, if $t$ is a lower bound for $S$, then since $n + 1 \in S$ we must have $t \leq n + 1$, so $n + 1$ is a greatest lower bound for $S$. By definition, then $n + 1$ is a minimum for $S$. However, by hypothesis, $S$ contains no minimum, and hence it cannot be the case that $n + 1 \in S$. Therefore, $n + 1 \notin S$.

By induction, then, $n \notin S$ for all $n \in \mathbb{N}$, so $S^c = \mathbb{N}$ and thus $S = \emptyset$. But by assumption, $S \neq \emptyset$, and hence we have reached a contradiction.

Therefore, we must have that if $S$ is a subset of $\mathbb{N}$ with $S \neq \emptyset$, then $S$ contains a minimum. □

In general, we say that a universe $\Omega$ is *well-ordered* if, for any nonempty subset $S$ of $\Omega$, $S$ has a minimum element. Hence, the Well-Ordering Principle above is showing that $\mathbb{N}$ is a well-ordered set. It is the only well-ordered set among our favorite number sets $(\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R})$, but certainly is not the only well-ordered set in existence.

This may seem like a lot of work to prove something seemingly obvious, but the result is in fact quite useful. Consider, for example, the following result, which we have used a few times already this semester without formal proof: that rational numbers can be written in lowest terms.

**Theorem 5.** *Let $q \in \mathbb{Q}$, with $q > 0$. Then there exist nonnegative integers $a, b$ such that $q = \frac{a}{b}$, and $a$ and $b$ have no common factors.*

**Proof.** Let $q \in \mathbb{Q}$ with $q > 0$. Define $S \subseteq \mathbb{N}$ by

$$S = \{a \in \mathbb{N} \mid \exists b \in \mathbb{N} \text{ such that } q = \frac{a}{b}\},$$

that is to say, $S$ is the set of possible numerators for $q$.

Note that by definition, since $q \in \mathbb{Q}$, there exist integers $a, b$ such that $\frac{a}{b} = q$. If $a > 0$ then $a \in S$. If $a < 0$, then we also have $q = \frac{-a}{-b}$, so $-a \in S$. Moreover, as $q > 0$, we cannot have $a = 0$. Hence, in any case, we have that $S$ is not empty.

By the Well-Ordering Principle, then, $S$ contains a minimum element, say $a_0$. Let $b_0$ be such that $q = \frac{a_0}{b_0}$. We claim that $a_0$ and $b_0$ share no common factor. To prove this, we work by contradiction.

Suppose that $a_0$ and $b_0$ have a common factor, $d > 1$. Then there exist positive integers $k, j$ such that $a_0 = dk$ and $b_0 = dj$. We therefore have that $q = \frac{dk}{dj} = \frac{k}{j}$. Since $k, j$ are positive integers, then we have that $k \in S$. Moreover, as $d > 1$, we have that $k < a_0$. But $a_0$ is a minimum for $S$, so $k \geq a_0$. This is obviously impossible, and hence it must be the case that $a_0$ and $b_0$ have no common factors. □

The Well-Ordering Principle can also be used to prove statements that one hopes to be true for all $n \in \mathbb{N}$ using the method of contradiction. By the Well-Ordering Principle, if a statement of the form $\forall\ n \in \mathbb{N},\ p(n)$ is false, then there must be a smallest $n \in \mathbb{N}$ which is a counterexample. This is true because if we take $S = \{n \in \mathbb{N} \mid \neg p(n)\}$, then $S$ is a nonempty set of natural numbers, and hence it must have a minimum. This can be useful in the case of proof by contradiction, as in the following example. We note that this example could also have been proven by strong induction, so this is not the only approach here.

**Example 3.** Suppose we live in a world where there are exactly two values of postage stamps, namely 2¢ and 5¢ stamps. In addition, the lowest postage for any letter is 4¢. We say that a number $n$ is *postal* if we can affix stamps to a letter valuing exactly $n$¢. Prove that every number $n \geq 4$ is postal.

**Proof.** Suppose that the statement is not true, so that there exists some $n \geq 4$ with $n$ not postal. By the Well-Ordering Principle, then, there must be a least $n$ with $n$ not postal; call this minimum $n_0$.

Note that $4 \neq n_0$, since we can post a 4c letter with 2 stamps, each valued at 2¢.

Note that $5 \neq n_0$, since we can post a 5c letter with 1 stamp valued at 5¢.

For any $n > 5$, suppose that $n = n_0$. Then we note that $n - 2 \geq 4$, and $n - 2$ is postal, so there is some combination of stamps valuing $n - 2$¢. But then adding one 2¢stamp to this yields a combination of stamps valuing $n$¢, and hence $n \neq n_0$.

But then no choice of $n \geq 4$ can be a minimum counterexample, and hence no such counterexample exists. Therefore, every number $n \geq 4$ is postal. □

The critical point in the application of this technique occurs in the 4th paragraph, in which we note that since $n_0$ is a *minimal* counterexample, any smaller number is not a counterexample at all.

Again, a cautionary note: we can only apply the technique of contradiction by minimum counterexample to statements that apply to a well-ordered set like the natural numbers. Also, it is often true that rather than work by minimal counterexample, we can work by strong induction instead. Indeed, if you think about the minimal counterexample as $n_0$, we are implicitly assuming that the result holds for all $n < n_0$, as these are not counterexamples. So this is really the same type of technique as working by Strong Induction combined with contradiction. However, the technique is used frequently enough that it merits this small discussion, especially in proofs in discrete mathematics and graph theory.