1 Mathematical Preliminaries

1.1 Set Theory

Definition 1 (Set)**.** A *set* is collection of distinct elements, where the order in which the elements are listed does not matter. The size of a set S, denoted $|S|$, is known as its *cardinality* or *order*. The members of a set are referred to as its elements. We denote membership of x in S as $x \in S$. Similarly, if x is not in S, we denote $x \not\in S$.

Example 1. Common examples of sets include the set of real numbers $\mathbb{R}$, the set of rational numbers $\mathbb{Q}$, and the set of integers $\mathbb{Z}$. The sets $\mathbb{R}^+, \mathbb{Q}^+$ and $\mathbb{Z}^+$ denote the strictly positive elements of the reals, rationals, and integers respectively. We denote the set of natural numbers $\mathbb{N} = \{0, 1, \ldots\}$. Let $n \in \mathbb{Z}^+$ and denote $[n] = \{1, \ldots, n\}$.

We now review several basic set operations, as well as the power set. It is expected that students will be familiar with these constructs. Therefore, we proceed briskly, recalling definitions and basic examples intended solely as a refresher.

Definition 2. Set Union Let A, B be sets. Then the *union* of A and B, denoted $A \cup B$ is the set:

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

Example 2. Let $A = \{1, 2, 3\}$ and $B = \{4, 5, 6\}$. Then $A \cup B = \{1, 2, 3, 4, 5, 6\}$.

Example 3. Let $A = \{1, 2, 3\}$ and $B = \{3, 4, 5\}$. So $A \cup B = \{1, 2, 3, 4, 5\}$. Recall that sets do not contain duplicate elements. So even though 3 appears in both A and B, 3 occurs exactly once in $A \cup B$.

Definition 3. Set Intersection Let A, B be sets. Then the *intersection* of A and B, denoted $A \cap B$ is the set:

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

Example 4. Let $A = \{1, 2, 3\}$ and $B = \{1, 3, 5\}$. Then $A \cap B = \{1, 3\}$. Now let $C = \{4\}$. So $A \cap C = \emptyset$.

Definition 4 (Symmetric Difference)**.** Let A, B be sets. Then the *symmetric difference* of A and B, denoted $A \triangle B$ is the set:

$$A \triangle B := \{x : x \in A \text{ or } x \in B, \text{ but } x \not\in A \cap B\}$$

Example 5. Let $A = \{1, 2, 3\}$ and $B = \{1, 3, 5\}$. Then $A \triangle B = \{2, 5\}$.

For our next two definitions, we let U be our *universe*. That is, let U be a set. Any sets we consider are subsets of U.

Definition 5 (Set Complementation)**.** Let A be a set contained in our universe U. The *complement* of A, denoted A^C or $\overline{A}$, is the set:

$$\overline{A} := \{x \in U : x \not\in A\}$$

Example 6. Let $U = [5]$, and let $A = \{1, 2, 4\}$. Then $\overline{A} = \{3, 5\}$.

Definition 6 (Set Difference)**.** Let A, B be sets contained in our universe U. The *difference* of A and B, denoted $A \setminus B$ or $A - B$, is the set:

$$A \setminus B = \{x : x \in A \text{ and } x \not\in B\}$$

Example 7. Let $U = [5]$, $A = \{1, 2, 3\}$ and $B = \{1, 2\}$. Then $A \setminus B = \{3\}$.

Remark: The Set Difference operation is frequently known as the *relative complement*, as we are taking the complement of B relative to A rather than with respect to the universe U.

Definition 7 (Cartesian Product)**.** Let A, B be sets. The *Cartesian product* of A and B, denoted $A \times B$, is the set:

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

Example 8. Let $A = \{1, 2, 3\}$ and $B = \{a, b\}$. Then $A \times B = \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\}$.

Definition 8 (Power Set)**.** Let S be a set. The *power set* of S, denoted 2^S or $\mathcal{P}(S)$, is the set of all subsets of S. Formally:

$$2^S := \{A : A \subset S\}$$

Example 9. Let $S = \{1, 2, 3\}$. So $2^S = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$.

Remark: For finite sets S, $|2^S| = 2^{|S|}$; hence, the choice of notation.

Definition 9 (Subset)**.** Let A, B be sets. A is said to be a *subset* of B if for every $x \in A$, we have $x \in B$ as well. This is denoted $A \subset B$ (equivocally, $A \subseteq B$). Note that B is a *superset* of A.

Example 10. Let $A = [3], B = [6], C = \{2, 3, 5\}$. So we have $A \subset B$ and $C \subset B$. However, $A \not\subset C$ as $1 \not\in C$; and $C \not\subset A$, as $5 \not\in A$.

Remark: Let S be a set. The subset relation forms a partial order on 2^S. To show two sets A and B are equal, we must show $A \subset B$ and $B \subset A$. We demonstrate how to prove two sets are equal below.

Proposition 1.1. *Let $A = \{6n : n \in \mathbb{Z}\}, B = \{2n : n \in \mathbb{Z}\}, C = \{3n : n \in \mathbb{Z}\}$. So $A = B \cap C$.*

Proof. We first show that $A \subset B \cap C$. Let $n \in \mathbb{Z}$. So $6n \in A$. We show $6n \in B \cap C$. As 2 is a factor of 6, $6n = 2 \cdot (3n) \in B$. Similarly, as 3 is a factor of 6, $6n = 3 \cdot (2n) \in C$. So $6n \in B \cap C$. We now show that $B \cap C \subset A$. Let $x \in B \cap C$. Let $n_1, n_2 \in \mathbb{Z}$ such that $x = 2n_1 = 3n_2$. As 2 is a factor of x and 3 is a factor of x, it follows that $2 \cdot 3 = 6$ is also a factor of x. Thus, $x = 6n_3$ for some $n_3 \in \mathbb{Z}$. So $x \in A$. Thus, $B \cap C \subset A$. Thus, $A = B \cap C$, as desired. □

Proposition 1.2. *Let A, B, C be sets. Then $A \times (B \cup C) = (A \times B) \cup (A \times C)$.*

Proof. Let $(x, y) \in A \times (B \cup C)$. If $y \in B$, then $(x, y) \in (A \times B)$. Otherwise, $y \in C$ and so $(x, y) \in (A \times C)$. Thus, $A \times (B \cup C) \subset (A \times B) \cup (A \times C)$. Now let $(d, f) \in (A \times B) \cup (A \times C)$. Clearly, $d \in A$. So f must be in either B or C. Thus, $(d, f) \in A \times (B \cup C)$, which implies $(A \times B) \cup (A \times C) \subset A \times (B \cup C)$. We conclude that $A \times (B \cup C) = (A \times B) \cup (A \times C)$. □

1.2 Relations and Functions

Definition 10 (Relation)**.** Let X be a set. A k-ary relation on X is a subset $R \subset X^k$.

Example 11. The notion of equality $=$ over $\mathbb{R}$ is the canonical example of a relation. It is perhaps the most well-known instance of an *equivalence relation*, which will be discussed later.

Intuitively, a k-ary relation R contains k-tuples of elements from X that share common properties. Computer scientists and mathematicians are interested in a number of different relations, including the adjacency relation (graph theory), equivalence relations, orders (such as partial orders), and functions. In this section, functions, asymptotics, and equivalence relations will be discussed.

1.2.1 Functions

The notion of a *function* will be introduced first. Functions are familiar mathematical objects, which appear early on in mathematics education with the notion of an input-output machine. Roughly speaking, a function takes an input and produces an output. Some common examples include the linear equation $f(x) = ax + b$ and the exponential $f(x) = 2^x$.

We denote a function as follows. Let X and Y be sets. A function is a map $f : X \to Y$ such that for every $x \in X$, there is a unique $y \in Y$ where $f(x) = y$. We say that X is the *domain* and Y is the *codomain*. The *range* or *image* is the set $f(X) = \{f(x) : x \in X\}$. More formally, a function is defined as follows:

Definition 11. Function Let X and Y be sets. A function f is a subset (or 1-place relation) of $X \times Y$ such that for every $x \in X$, there exists a unique $y \in Y$ where $(x, y) \in f$.

Let's consider some formal functions and one example of a relation that is not a function.

Example 12. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x$. This is known as the *identity map*.

Example 13. Let $g : \mathbb{R} \to \mathbb{R}$ be given by $g(x) = 3x^2$.

Example 14. Let $h : \mathbb{R} \to \mathbb{R}$ given by:

$$h(x) = \begin{cases} x & : x \neq 3 \\ -3, 2 & : x = 3 \end{cases}$$

Note that h is *not* a function as $(3, -3) \in h$ and $(3, 2) \in h$. The definition of a function states that there must be a *unique* y such that $(3, y) \in h$. If we revise h such that $h(3) = -3$ *only*, then h satisfies the definition of a function.

From a combinatorial perspective, special types of functions known as *injections* and *surjections* are of great importance. The idea is that if we have two sets X and Y and know the cardinality of X, then an injection or surjection from X to Y yields results about Y's cardinality. We can deduce similar results about Y's cardinality.

An injection is also known as a *one-to-one* function. Recall the definition of a function states that the map $f : X \to Y$ maps each $x \in X$ to a unique $y \in Y$. It allows for functions such as $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = 0$. Clearly, every $x \in \mathbb{R}$ maps to the same y-coordinate: $y = 0$. An injection disallows functions such as these. The idea is that each $y \in Y$ can be paired with at most one $x \in X$, subject to the constraint that each element in X must be mapped to some element from Y. So there can be unmapped elements in Y, but not in X.

We define this formally as follows.

Definition 12 (Injection)**.** A function $f : X \to Y$ is said to be an injection if $f(x_1) = f(x_2) \implies x_1 = x_2$. Equivocally, f is an injection if $x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$.

Let's consider examples of functions that are injections, as well as those that fail to be injections.

Example 15. Let X be a set. Recall the identity map $\text{id} : X \to X$ given by $\text{id}(x) = x$. This function is an injection. Let $\text{id}(x_1) = \text{id}(x_2)$. Then we have $\text{id}(x_1) = x_1 = \text{id}(x_2) = x_2$, which implies that $x_1 = x_2$.

Example 16. Consider the function $g : \mathbb{R} \to \mathbb{R}$ be given by $g(x) = x^2$. Observe that g fails to be an injection. Let $g(x_1) = g(x_2) = 4$. We have $x_1 = 2$ and $x_2 = -2$, both of which map to 4. If we instead consider $h : \mathbb{R}^+ \to \mathbb{R}$ by $h(x) = x^2$, we have an injection since we only consider the positive real numbers. Observe as well that both g and h do not map to any element less than 0.

Remark: Let's reflect on what we know about injections. An injection is a function, in which any mapped element in the codomain is mapped to exactly once. There may be elements in the codomain which remain unmapped. As a result, for two sets X and Y, it is defined that $|X| \leq |Y|$ if there exists an injection $f : X \to Y$. Intuitively speaking, an injection pairs each element from the domain with an element in the codomain, allowing for leftover elements in the codomain. Hence, X has no more elements than Y if there exists an injection $f : X \to Y$.

Surjections or onto functions center around the codomain, rather than the domain. Recall the earlier example of $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = x^2$, which satisfies $g(x) \geq 0$ for every $x \in \mathbb{R}$. So the negative real numbers will never be maped under g. Surjections exclude functions like g. Intuitively, a function is a surjection if every element in the codomain is mapped. Any element of the codomain can have multiple domain points mapping to it, as long as each has at least one domain point mapping to it. We define this formally as follows.

Definition 13 (Surjection)**.** Let X and Y be sets. A function $f : X \to Y$ is a surjection if for every $y \in Y$, there exists an $x \in X$ such that $f(x) = y$.

We have already seen an example of a function that is not a surjection. Let us now consider a couple examples of functions that are surjections.

Example 17. Recall the identity map $\text{id} : X \to X$. For any $x \in X$, we have $\text{id}(x) = x$. So the identity map is a surjection.

Example 18. Let $X = \{a, b, c, d\}$ and let $Y = \{1, 2, 3\}$. Define $f : X \to Y$ by $f(a) = f(b) = 1, f(c) = 2$ and $f(d) = 3$. This function is a surjection, as each $y \in Y$ is mapped under f. Observe that there are more X elements than Y elements. If X instead had two elements, then f would not be a surjection because at most two of the three elements in Y could be mapped.

Remark: Similarly, let's now reflect upon what we know about surjections. A surjection is a function in which every element of the codomain is mapped at least once. Some elements in the codomain may have multiple elements in the domain mapping to them. Therefore, if there exists a surjection $f : X \to Y$, then $|X| \geq |Y|$. We now introduce the notion of a bijection. A function is a bijection if it is both an injection and a surjection. Intuitively, a bijection matches the elements in the domain and codomain in a one-to-one manner. That is, each element in the domain has precisely one mate in the codomain and vice versa. For this reason, two sets X and Y are defined to have the same cardinality if there exists a bijection $f : X \to Y$. Combinatorialists use bijections to ascertain set cardinalities. The idea is that given sets X and Y, with $|Y|$ known, can we construct a bijection $f : X \to Y$? If the answer is yes, then $|X| = |Y|$.

Definition 14 (Bijection)**.** Let X and Y be sets. A bijection is a function $f : X \to Y$ that is both an injection and a surjection.

Example 19. Some examples of bijections include the identity map, as well as the linear equation $f(x) = mx + b$.

We conclude by showing that the composition of two injective functions are injective, and that the composition of two surjective functions are surjective. This implies that the composition of two bijections is itself a bijection, which is an important fact when working with permutations (which we shall see later).

Proposition 1.3. *Let $f : X \to Y$ and $g : Y \to Z$ be injective functions. Then $g \circ f$ is also injective.*

Proof. Let $x_1, x_2 \in X$ be distinct. As f is injective, $f(x_1) \neq f(x_{2)}$. Similarly, as g is injective, $g(f(x_1)) \neq g(f(x_2))$. So $(g \circ f)(x_1) \neq (g \circ f)(x_2)$, as desired. As x_1, x_2 were arbitrary, we conclude that $g \circ f$ is injective. □

Proposition 1.4. *Let $f : X \to Y$ and $g : Y \to Z$ be surjective functions. Then $g \circ f$ is also surjective.*

Proof. Let $z \in Z$. As g is surjective, there exists $y \in Y$ such that $g(y) = z$. Now as f is surjective, there exists $x \in X$ such that $f(x) = y$. Thus, $(g \circ f)(x) = z$. As z was arbitrary, it follows that $g \circ f$ is surjective. □

1.2.2 Equivalence Relations

Equivalence relations are of particular importance in mathematics and computer science. Intuitively, an equivalence relation compares which elements in a set X share some common property. The goal is to then partition X into equivalence classes such that all the elements in one of these parts are all equivalent to each other. This allows us to select an arbitrary distinct representative from each equivalence class and consider only that representative.

This idea of partitioning comes up quite frequently. The integers modulo n, denoted $\mathbb{Z}/n\mathbb{Z}$, is a canonical example. Big-Theta is another important equivalence relation. Equivalence relations allow us to prove powerful theorems such as Fermat's Little Theorem from Number Theory and Cauchy's Theorem from Group Theory, as well as to construct a procedure to minimize finite state automata via the Myhill-Nerode Theorem.

In order to guarantee such a partition, an equivalence relation must satisfy three properties: reflexivity, symmetry, and transitivity. We define these formally below, restricting attention to binary relations.

Definition 15 (Reflexive Relation)**.** A relation R on the set X is said to be reflexive if $(a, a) \in R$ for every $a \in X$.

Definition 16 (Symmetric Relation)**.** A relation R on the set X is said to be symmetric if $(a, b) \in R$ if and only if $(b, a) \in R$ for every $a, b \in X$.

Definition 17 (Transitive Relation)**.** A relation R on the set X is said to be transitive if for every $a, b, c \in X$ satisfying $(a, b), (b, c) \in R$, then $(a, c) \in R$.

Definition 18 (Equivalence Relation)**.** An equivalence relation is a reflexive, symmetric, and transitive relation.

Let us break each of these definitions down and compare them to how the equality relation on $\mathbb{R}$ behaves. Intuitively, a real number is equal to itself; i.e., $3 = 3$, $0 = 0$ and $1 \neq 2$. The properties of an equivalence relation reflect this behavior. The reflexive axiom states that $(a, a) \in R$ for every $a \in X$. Intuitively, reflexivity captures this notion that an element is equivalent to itself.

Now consider the definition of a symmetric relation: for every $a, b \in X$, $(a, b) \in R$ if and only $(b, a) \in R$. Suppose in a high school algebra problem we deduce that that for the variables x and y, we have $x = y$. Does it make sense that $y \neq x$? Of course not. Equivalence relations must capture this property as well, which is the purpose of the symmetry axiom. Elements in the same equivalence class must be pairwise equivalent.

The last axiom is transitivity. We refer back to the example of the high school algebra problem. Suppose this time we have three variables x, y and z satisfying $x = y$ and $y = z$. Over the real numbers, it makes perfect sense that $x = z$. So from an intuitive perspective, it is important that equivalence relations enforce this property. However, from a more technical perspective, transitivity implies that the equivalence classes are pairwise disjoint. In other words, transitivity is really the driving force in partitioning the set into equivalence classes. This will be proven later.

The congruence relation $a \equiv b \pmod{n}$ is a canonical example of an equivalence relation. We prove this below.

Definition 19 (Congruence Relations)**.** Let $n \geq 1$ be an integer. The congruence relation modulo n is a binary relation on $\mathbb{Z}$ given by: $a \equiv b \pmod{n}$ (read as: a is congruent to b modulo n) if and only if n divides $b - a$.

Proposition 1.5. *Let $n \geq 1$ be an integer. The relation $a \equiv b \pmod{n}$ is an equivalence relation.*

Proof. We show that the congruence relation modulo n is reflexive, symmetric, and transitive.

- **Reflexivity.** Let $a \in \mathbb{Z}$. We show that $a \equiv a \pmod{n}$. So n divides $a - a = 0$. Thus, $a \equiv a \pmod{n}$. So the congruence relation modulo n is reflexive.

- **Symmetry.** Let $a, b \in \mathbb{Z}$ such that $a \equiv b \pmod{n}$. We show that $b \equiv a \pmod{n}$. Let $q \in \mathbb{Z}$ such that $nq = a - b$. Thus, $n(-q) = b - a$, so n divides $b - a$. Thus, $b \equiv a \pmod{n}$. So the congruence relation modulo n is symmetric.

- **Transitivity.** Let $a, b, c \in \mathbb{Z}$ such that $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$. We show that $a \equiv c \pmod{n}$. By the definition of the congruence relation, n divides $(a - b)$ and n divides $(b - c)$. Let $h, k \in \mathbb{Z}$ such that $nh = a - b$ and $nk = b - c$. So $nh + nk = n(h + k) = a - c$. Thus, n divides $a - c$, so $a \equiv c \pmod{n}$. It follows that the congruence relation modulo n is transitive.

We conclude that the congruence relation modulo n is an equivalence relation. □

We now formalize the notion of an equivalence class, with the goal of showing that an equivalence relation *partitions* a set. Informally, a partition of a set X is a collection of disjoint subsets of X, whose union is precisely X. This is formalized as follows.

Definition 20 (Partition)**.** Let X be a set. A *partition* of X is a set $\mathcal{P}$ satisfying the following.

- Each member $P \in \mathcal{P}$ is a non-empty subset of X.

- For any two distinct $P_1, P_2 \in \mathcal{P}$, $P_1 \cap P_2 = \emptyset$.

- $\displaystyle\bigcup_{P \in \mathcal{P}} P = X$.

Definition 21 (Equivalence Class)**.** Let X be a set, and let $\equiv$ be an equivalence relation on X. Fix $x \in X$. The *equivalence class* of x is the set $[x] = \{s \in X : x \equiv s\}$. That is, $[x]$ is the set of elements that are equivalent to x.

Remark: Note that $[x] = [s]$ for any $s \in [x]$. This follows from the transitivity of $\equiv$, which will be proven shortly.

Example 20. Fix $n \geq 1$. The equivalence classes of the congruence relation modulo n are the classes $[0], [1], \ldots, [n-1]$. Some additional tools from number theory are required to justify this, so we omit a proof. Informally, the congruence relation modulo n is represented by the remainder classes upon division by n.

As an equivalence relation is reflexive, every element of the set belongs to some equivalence class. Thus, in order for an equivalence relation to partition the set, it suffices to show that the equivalence classes are pairwise disjoint.

Proposition 1.6. *Let $\equiv$ be an equivalence relation on the set X. Let $[x], [y]$ be distinct equivalence classes under $\equiv$. Then $[x] \cap [y] = \emptyset$.*

Proof. Suppose to the contrary that $[x] \cap [y] \neq \emptyset$. As $[x]$ and $[y]$ are distinct and have non-empty intersection, there exists $z \in [y] - [x]$. Without loss of generality, suppose $y \in [x] \cap [y]$. So $x \equiv y$. Since $z \in [y]$, we have $y \equiv z$. By transitivity, $x \equiv z$, which implies $z \in [x]$, a contradiction. □

1.3 Proof by Induction

Many theorems of interest ask us to prove a proposition holds for all natural numbers. Verifying such statements for all natural numbers is challenging, due to the fact that our domain is not just large but infinite. Informally, proof by induction allows us to verify a small subset of base cases. Together, these base cases imply the subsequent cases. Thus, the desired theorem is proven as a result.

Intuitively, we view the statements as a sequence of dominos. Proving the necessary base cases knocks (i.e., proves true) the subsequent dominos (statements). It is inescapable that all the statements are knocked down; thus, the theorem is proven true.

The most basic form of induction is the Principle of Weak Induction.

Definition 22 (Principle of Weak Induction)**.** Let $P(n)$ be a proposition regarding an integer n, and let $k \in \mathbb{Z}$ be fixed. If:

(a) $P(k)$ holds; and

(b) for every $m \geq k$, $P(m)$ implies $P(m+1)$,

then for every $n \geq k$, $P(n)$ holds.

The definition of the Principle of Weak Induction in fact provides a format for structuring proofs.

- First, we verify a single base case: for some $k \in \mathbb{N}$, the statement $P(k)$ holds.
- Second, we assume that $P(m)$ holds for some integer $m \geq k$. This step is known as the **inductive hypothesis**, and it is indispensible for a proof by induction. We must and do use the fact that $P(m)$ is true when proving that $P(m+1)$ holds.
- In our final step, we show that for an arbitrary $m \geq k$ that $P(m)$ implies $P(m+1)$. This is known as the **inductive step**.

We illustrate this proof technique with the following example.

Proposition 1.7. *Fix $n \in \mathbb{N}$. We have:* $\displaystyle\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$.

Proof. We prove this theorem by induction on $n \in \mathbb{N}$.

- **Base Case.** Our first step is to verify the base case: $n = 0$. In this case, we have $\sum_{i=0}^{n} i = 0$. Note as well that $\frac{0 \cdot 1}{2} = 0$. Thus, the proposition holds when $n = 0$.

- **Inductive Hypothesis.** Now for our inductive hypothesis: fix $k \geq 0$, and suppose $\displaystyle\sum_{i=0}^{k} i = \frac{k(k+1)}{2}$.

- **Inductive Step.** We prove true for the $k+1$ case. Consider:

  $$\sum_{i=0}^{k+1} i = (k+1) + \sum_{i=0}^{k} i$$

  By the inductive hypothesis, $\sum_{i=0}^{k} i = \frac{k(k+1)}{2}$. Now we have:

  $$\begin{aligned}
  &(k+1) + \sum_{i=0}^{k} i \\
  &= (k+1) + \frac{k(k+1)}{2} \\
  &= \frac{2(k+1) + k(k+1)}{2} \\
  &= \frac{(k+1)(k+2)}{2}.
  \end{aligned}$$

So by the Principle of Weak Induction, the result follows. □

Remark: Notice that the inductive hypothesis was imperative in the inductive step. Once we used that $\displaystyle\sum_{i=0}^{k} i = \frac{k(k+1)}{2}$, it was a matter of algebraic manipulation to obtain that $\displaystyle\sum_{i=0}^{k+1} i = \frac{(k+1)(k+2)}{2}$. As we have verified a base case when $n = 0$ and proven that $S(k)$ implies $S(k+1)$ for an arbitrary $k \geq 0$, the Principle of Weak Induction affords us that the proposition is true.

We now examine a second example applying the Principle of Weak Induction.

Proposition 1.8. *For each $n \in \mathbb{N}$ and each $x > -1$, $(1+x)^n \geq 1 + nx$.*

Proof. The proof is by induction on n.

- **Base Case.** Consider the base case of $n = 0$. So we have $(1+x)^n = 1 \geq 1 + 0x = 1$. So the proposition holds at $n = 0$.

- **Inductive Hypothesis.** Fix $k \geq 0$ and suppose that $(1+x)^k \geq 1 + kx$.

- **Inductive Step.** We have that $(1+x)^{k+1} = (1+x)^k(1+x)$. By the inductive hypothesis, $(1+x)^k \geq (1+kx)$. So:

  $$(1+x)^k(1+x) \geq (1+kx)(1+x) = 1 + (k+1)x + kx^2 \geq 1 + (k+1)x.$$

  The last inequality follows from the fact that kx^2 is non-negative; so removing it from the right hand side will not increase that side.

So by the Principle of Weak Induction, the result follows. □

We next introduce the Principle of Strong Induction. Intuitively, strong induction is useful in proving theorems of the form "for all n, $P(n)$" where $P(k)$ alone does not neatly lend itself to forcing $P(k+1)$ to be true. Instead, it may be easier to leverage some subset of $\{P(0), \ldots, S(k)\}$ to force $P(k+1)$ to be true. Strong induction allows us to use any or all of $P(0), \ldots, P(k)$ to prove that $P(k+1)$ is true. The Principle of Strong Induction is formalized as follows.

Definition 23 (Principle of Strong Induction)**.** Let $P(n)$ be a proposition regarding an integer n, and let $k \in \mathbb{Z}$ be fixed. If:

- $P(k)$ is true; and
- for every $m \geq k$, $[P(k) \wedge P(k+1) \wedge \ldots \wedge P(m)]$ implies $P(m+1)$,

then for every $n \geq k$, the statement $P(n)$ is true.

Just as with the Principle of Weak Induction, the Principle of Strong Induction provides a format for structuring proofs.

- First, we verify for all base cases k, the statement $P(k)$ holds. This ensures that subsequent cases which rely on these early base cases are sound. For example, strong inductive proofs regarding graphs or recurrence relations may in fact have several base cases, which are used in constructing subsequent cases.
- Second, we assume that for some integer $m \geq k$, $P(k), \ldots, S(m)$ **all** hold. Notice our inductive hypothesis using strong induction assumes that **each** of the previous cases are true, while the inductive hypothesis when using weak induction only assumes $P(m)$ to be true. Strong induction assumes the extra cases because we end up using them.
- In our final step (the inductive step), we show that for an arbitrary $m \geq k$ that $P(k) \wedge P(k+1) \wedge \ldots \wedge P(m)$ implies $P(m+1)$.

Remark: The Principle of Weak Induction and the Principle of Strong Induction are equally powerful. That is, any proof using strong induction may be converted to a proof using weak induction, and vice versa. In practice, it may be easier to use strong induction, while weak induction may be clunky to use.

We illustrate how to apply the Principle of Strong Induction with a couple examples.

Proposition 1.9. *Let $f_0 = 0, f_1 = 1$; and for each natural number $n \geq 2$, let $f_n = f_{n-1} + f_{n-2}$. We have $f_n \leq 2^n$ for all $n \in \mathbb{N}$.*

Proof. The proof is by strong induction on $n \in \mathbb{N}$. Observe that $f_0 = 0 \leq 2^0 = 1$. Similarly, $f_1 = 1 \leq 2^1 = 2$. So our base cases of $n = 0, 1$ hold. Now fix $k \geq 1$; and suppose that for all $n \in \{0, \ldots, k\}$, $f_n \leq 2^n$. We now prove that $f_{k+1} \leq 2^{k+1}$. By definition of our sequence, $f_{k+1} = f_k + f_{k-1}$. We now apply the inductive hypothesis to f_k and f_{k+1}. Thus:

$$\begin{aligned} f_{k+1} &\leq 2^k + 2^{k-1} \\ &= 2^k \left(1 + \frac{1}{2}\right) \\ &\leq 2^k \cdot 2 = 2^{k+1} \end{aligned}$$

As desired. So by the Principle of Strong Induction, the result follows. □

Proposition 1.10. *Every positive integer can be written as the product of a power of 2 and an odd integer.*

Proof. The proof is by strong induction on $n \in \mathbb{Z}^+$. We have our base case $n = 1$. So $n = 1 \cdot 2^0$. Thus, the proposition holds for $n = 1$. Now fix $k \geq 1$, and suppose the proposition holds true for all $n \in [k]$. We prove true for the $k + 1$ case. We have two cases:

- Case 1: Suppose $k + 1$ is odd. Then $k + 1 = (k + 1) \cdot 2^0$, and we are done.
- Case 2: Suppose instead $k + 1$ is even. Then $k + 1 = 2h$ for some $h \in \mathbb{Z}^+$. By the inductive hypothesis, $h = m2^j$ for some odd integer m and $j \in \mathbb{N}$. Thus, $k + 1 = m2^{j+1}$, and we are done.

So by the Principle of Strong Induction, the result follows. □