Mathematical Induction

The following three principles governing $\mathbb{N}$ are equivalent.

- **Ordinary Induction Principle.** *If S is a set of natural numbers having $0 \in S$, and $(n + 1) \in S$ whenever $n \in S$, then $S = \mathbb{N}$.*

- **Complete Induction Principle.** *Suppose S is a set of natural numbers such that for every natural number n, if $m \in S$ for all $m < n$ then also $n \in S$. Then $S = \mathbb{N}$.*

- **Least Element Principle.** *If S is a nonempty subset of $\mathbb{N}$, then S has a least element.*

The Ordinary Induction Principle is the basis for the so-called *proof by weak induction*. The Complete Induction Principle is the basis for the so-called *proof by strong induction*. The Least Element Principle is the basis for the *proof by minimum counterexample*.

Any statement of the form

$$\forall n \in \mathbb{N}.\varphi(n)$$

can be proved by any one of these three proof methods. Assume for the rest of this handout that the statement we want to prove is of this form, unless stated otherwise. We let $S = \{n \in \mathbb{N} : \varphi(n)\}$; we thus need to show that $S = \mathbb{N}$.

Proof by Weak Induction

We desire to apply the Ordinary Induction Principle; we therefore need to show that all the hypotheses of the principle are satisfied. This means we have to show two things:

1. $\varphi(0)$,

2. for all $n \in \mathbb{N}$, $\varphi(n) \implies \varphi(n + 1)$.

Showing item 1 is usually called *proving the base case* and showing item 2 is usally called doing the *inductive step* (or *induction step*). The hypothesis $\varphi(n)$ is usally called the *inductive hypothesis* (or *induction hypothesis*). Let's look at an example.

Suppose we want to prove that *"if G is a graph of size m, then $\sum_{v \in V(G)} \deg v = 2m$."* As stated in the handout on Logic, this statement omits the universal quantifier. Rewriting it out in full we get,

Theorem (Theorem 2.1, CZ). *For every (natural number) m and for every G, if G is a graph of size m, then $\sum_{v \in V(G)} \deg_G v = 2m$.*

Proof. Let G be any graph of size m, where $m = 0$. We have $2m = 2 \cdot 0 = 0$. Since G has no edge, every vertex of it has degree 0. Thus,

$$\sum_{v \in V(G)} \deg_G v = \sum_{v \in V(G)} 0 = 0.$$

Therefore, $\sum_{v \in V(G)} \deg_G v = 0 = 2m$ when $m = 0$.

Now suppose that m is any positive integer and assume inductively that any graph H of size $m - 1$ satisfies $\sum_{v \in V(H)} \deg_H v = 2(m - 1)$. Let G be any graph of size m. Since $m > 0$, the graph G has at least one edge. Let $e = xy$ be an edge in G. Consider the graph $G - e$. It has $m - 1$ edges. Thus, the inductive hypothesis is applicable to it, i.e., we have that $\sum_{v \in V(G-e)} \deg_{G-e} v = 2(m - 1)$. We compute

$$
\begin{aligned}
\sum_{v \in V(G)} \deg_G v &= \sum \{\deg_G v : v \in V(G), v \neq x, v \neq y\} + \deg_G x + \deg_G y \\
&= \sum \{\deg_{G-e} v : v \in V(G), v \neq x, v \neq y\} + (\deg_{G-e} x + 1) + (\deg_{G-e} y + 1) \\
&= \sum \{\deg_{G-e} v : v \in V(G), v \neq x, v \neq y\} + \deg_{G-e} x + \deg_{G-e} y + 2 \\
&= \sum \{\deg_{G-e} v : v \in V(G)\} + 2 \\
&= 2(m - 1) + 2 \\
&= 2m - 2 + 2 \\
&= 2m.
\end{aligned}
$$

The theorem is thus proven by induction. $\square$

Proof by Minimum Counterexample

The strategy here is to prove by contradiction. Instead of attempting to show $S = \mathbb{N}$, we try to show $S \neq \mathbb{N} \implies \textit{False}$. Since $S \subseteq \mathbb{N}$ by our definition of S, the only way that $S \neq \mathbb{N}$ can hold is that $\mathbb{N} \subset S$, i.e., $\mathbb{N} \setminus S \neq \emptyset$, i.e., the set $X \stackrel{\text{def}}{=} \{n \in \mathbb{N} : \neg\varphi(n)\}$ is nonempty. So we begin the proof by supposing that X is nonempty. The Least Element Principle then guarantees that X has a least element. We then proceed to show that this leads to a contradiction. Let's prove the last theorem again by minimum counterexample.

Theorem (Theorem 2.1, CZ). *For every (natural number) m and for every G, if G is a graph of size m, then $\sum_{v \in V(G)} \deg_G v = 2m$.*

Proof. Suppose that there exists some natural number m and there is some graph of size m such that the sum of all its vertex degrees does not equal $2m$. Let G be a graph of size k, and k is least possible, such that $\sum_{v \in V(G)} \deg_G v \neq 2k$. Since any graph H of size 0 certainly satisfies $\sum_{v \in V(H)} \deg_H v = 0 = 2m$, we conclude that $k > 0$. This means that G has some edge. Let $e = xy$ be an edge in G. Consider the graph $G - e$. It has $k - 1$ edges. Since $G - e$ has fewer edges than G we see that $\sum_{v \in V(G-e)} \deg_{G-e} v = 2(k - 1)$ (Why?). We compute

$$
\begin{aligned}
\sum_{v \in V(G)} \deg_G v &= \sum \{\deg_G v : v \in V(G), v \neq x, v \neq y\} + \deg_G x + \deg_G y \\
&= \sum \{\deg_{G-e} v : v \in V(G), v \neq x, v \neq y\} + (\deg_{G-e} x + 1) + (\deg_{G-e} y + 1) \\
&= \sum \{\deg_{G-e} v : v \in V(G), v \neq x, v \neq y\} + \deg_{G-e} x + \deg_{G-e} y + 2 \\
&= \sum \{\deg_{G-e} v : v \in V(G)\} + 2 \\
&= 2(k - 1) + 2 \\
&= 2k - 2 + 2 \\
&= 2k.
\end{aligned}
$$

However, this contradicts our assumption on G that $\sum_{v \in V(G)} \deg_G v \neq 2k$. $\qquad\square$

Proof by Strong Induction

This proof method is similar to the proof by weak induction. We need to prove the base case and do the inductive step. The difference is in the inductive hypothesis and the inductive step. For strong induction, the inductive hypothesis is

$$\forall m \in \mathbb{N}(m < n \implies \varphi(m))$$

and the inductive step is

$$\forall n \in \mathbb{N}[(\forall m \in \mathbb{N}(m < n \implies \varphi(m))) \implies \varphi(n)].$$

Let's look at an example.

Theorem (Theorem 4.3, CZ). *Every tree of order n has size $n - 1$.*

Proof. Consider any tree of order $n = 1$. It has no edge, i.e., its size is $0 = n - 1$.

Now suppose n is a positive integer > 1 and assume inductively that for any positive integer k, if $k < n$ then any tree of order k has size $k - 1$. Let T be any tree of order n. Let m be the size of T. Since $n > 0$, it follows that T has some edge. Let e be any edge in T. The graph $T - e$ is then a union of two trees (Why?). Say that $T - e = T_1 \cup T_2$, and let n_1 (n_2) be the order of T_1 (T_2), and let m_1 (m_2) be the size of T_1 (T_2). Since $n = n_1 + n_2$ and both n_1 and n_2 are positive, we see that $0 < n_1 < n$ and $0 < n_2 < n$. Using the inductive hypothesis, we conclude that $m_1 = n_1 - 1$ and $m_2 = n_2 - 1$. Since $T - e$ results from deleting an edge from tree T, and since $T - e = T_1 \cup T_2$, it follows that $m = m_1 + m_2 + 1$. Thus,

$$
\begin{aligned}
m &= m_1 + m_2 + 1 \\
  &= (n_1 - 1) + (n_2 - 1) + 1 \\
  &= (n_1 + n_2) - 1 \\
  &= n - 1.
\end{aligned}
$$

The theorem is thus proven by induction. $\qquad\square$

Remarks

- Notice that the base case in the last proof is when $n = 1$, not when $n = 0$. How then is an induction principle applicable? It is because we can restate the statement of the theorem as *"for any natural number n and for any tree T, if $n \geq 0$ and T is of order n then T has size $n - 1$."* Since no tree of order 0 exists, the implication in the paraphrase is true vacuously when $n = 0$.

  This is a general pattern. The size of a smallest mathematical object in a collection of our interest may be positive. To prove by induction that all objects in this collection satisfy some property, we have to choose as base cases the sizes that are meaningful for these objects. In these cases, we actually apply an induction principle to prove some paraphrase of the original instead of the original. However, the paraphrase must imply the original.

- A statement of the form

  $$\varphi(n) \implies \varphi(n + 1) \text{ for all natural number } n$$

  and

  $$\varphi(n - 1) \implies \varphi(n) \text{ for all positive integer } n$$

  means exactly the same thing. Some authors use one form while other authors use the other. When reading proof, do not get bogged down by this seeming difference.

- Some proof by induction has more than one base case, and its inductive step seems incorrect. Let's look at the structure of one such hypothetical proof. First, the author proves two bases cases, viz., $\varphi(0)$, and $\varphi(1)$. Second, the author proves in the inductive step that for all $n \in \mathbb{N}[\varphi(n) \implies \varphi(n + 2)]$. Last, the author concludes that her theorem follows by mathematical induction.

  Why does the author need to prove two base cases? The answer is because her inductive step only proves that if φ holds for some odd (even) natural number n, then φ holds for the next odd (even) natural number $n + 2$.

  And what makes her proof a valid proof by mathematical induction? The answer is because she is secretly applying the Complete Induction Principle; she is actually proving by strong induction! (Do you see how?)

- In writing an induction proof, one may, like I have done, prove the base case(s) first and then do the inductive step second. One can also switch this order by doing the inductive step first and then proving the base case(s) second. Some authors prefer this latter ordering on the ground that the inductive step is usally where most of the real work is, while proving the base cases is often trivial. These authors may even skip the proof of the base case(s) entirely! This occurs mostly in journal prose, and I consider it a bad practice. Your induction proofs in this class must contain proof of the base case(s).

  No matter how you decide to present your induction proof, it must be clear from your writeup where your base case is, what your inductive hypothesis is, and where your inductive step is.

- You may come across *structural induction* in more advanced computer science courses. Structural induction is a technique for proving theorems about recursively defined objects like syntax trees and binary trees. A proof by structural induction does not contain a variable that ranges over the natural numbers. The reason it is still called induction is because its proof process is very similar to induction on $\mathbb{N}$. In fact, it is possible to define a variable n that ranges over $\mathbb{N}$ and paraphrase the statement we are proving by structural induction in terms of n. In that sense structural induction is normal induction in disguise. However, it offers a definite advantage over normal induction whenever it is applicable.