4.2. Induction

In Remark 3.1.77 we concluded with a formulation of the *minimal condition* (MC) for any order $<$ as follows:

An order $<$ on a class $\mathbb{A}$ has MC is capture by the statement

For any "property", that is, formula $F[x]$ —recall that this notation, square brackets*, indicates our interest in* one *among the, possibly many, free variables of F— we have that the following is true*

$$(\exists a)F[a] \to (\exists a)\Big(F[a] \wedge \neg(\exists y)\big(y < a \wedge F[y]\big)\Big) \tag{1}$$

So let $<$ be the standard order on $\mathbb{N}$. We have used the fact that it is a total order (satisfies trichotomy) and that every nonempty subset of $\mathbb{N}$ has a minimal —hence unique minimum— element.

Pause. Why *unique* and *minimum*?◀

So let us fix in the rest of this section $<$ to be the "less than" order on $\mathbb{N}$, until we indicate otherwise.

Let us rewrite (1) for $\neg P[x]$ where $P[x]$ is arbitrary. We get the theorem

$$(\exists x)\neg P[x] \to (\exists x)\Big(\neg P[x] \wedge \neg(\exists y)\big(y < x \wedge \neg P[y]\big)\Big) \tag{2}$$

Using the equivalence theorem (p.90) and the 7, we obtain from (2)

$$\neg(\forall x)P[x] \to \neg(\forall x)\neg\Big(\neg P[x] \wedge (\forall y)\neg\big(y < x \wedge \neg P[y]\big)\Big)$$

and then (the tautology known as "contrapositive" is used) also

$$(\forall x)\neg\Big(\neg P[x] \wedge (\forall y)\neg\big(y < x \wedge \neg P[y]\big)\Big) \to (\forall x)P[x]$$

Using the tautology

$$\neg(A \wedge B) \equiv \neg A \vee \neg B$$

and the equivalence theorem, we transform the above to this theorem:

$$(\forall x)\Big(P[x] \vee \neg(\forall y)\big(\neg y < x \vee P[y]\big)\Big) \to (\forall x)P[x]$$

Again, this time using the tautology

$$\neg A \vee B \equiv A \to B$$

(twice) and the equivalence theorem, we transform the above to this theorem:

$$(\forall x)\Big((\forall y)\big(y < x \to P[y]\big) \to P[x]\Big) \to (\forall x)P[x] \tag{3}$$

(3) is the principle of *strong induction*, or *complete induction*, or *course-of-values induction* that you probably encountered at school, and the above work shows that *it is equivalent to the least principle*! (Clearly we can reverse all the steps we took above as all were equivalences!)

Let us render (3) more recognisable: By applying MP (elaborate this!) I can transform (3) in "rule of inference form", indeed I will write it like a rule that says, like all rules do, "if you proved my numerator, then my denominator is also proved!"

$$\frac{(\forall x)\Big((\forall y)\big(y < x \to P[y]\big) \to P[x]\Big)}{(\forall z)P[z]}$$

Dropping the $\forall$-prefix we have the rule in the form:

$$\frac{(\forall y)\big(y < x \to P[y]\big) \to P[x]}{P[z]} \tag{CVI}$$

"(CVI)" for Course-of-Values Induction. (CVI) says

To prove $P[x]$ (for all x is implied!) **do as follows**:

Step (a) Fix an **arbitrary** x-value. Now, **assume** $(\forall y)\big(y < x \to P[y]\big)$ for said x. We call the assumption **Induction Hypothesis**, for short, **I.H.**

Step (b) Next **prove** $P[x]$, for the same fixed unspecified x. This proof step we call the **Induction Step** or **I.S.**

Note that what is described by (a) and (b) is precisely an application of the Deduction theorem towards proving "If, for all $y < x$, $P[y]$ is true, then $P[x]$ is true", that is, **proving the implication on the numerator of (CVI) for any given** x.

Step (c) If you have done **Step** (a) and **Step** (b) above, then you **have proved** $P[x]$ (for all x is implied!)

Important.

- **Step** (a) above says "**arbitrary** x".

  So, I should *not* leave any x-value out of the proof!

  But how do I prove the I.S. for $x = 0$? There is no I.H. to rely one (no numbers below $x = 0$). No problem: The numerator implication in (CVI) now reads

  $$(\forall y)\big(y < 0 \to P[y]\big) \to P[0]$$

  The lhs of "$\to$" is true since $y < 0$ is false. Thus, to ensure the truth of the *implication* I must prove $P[0]$.

  This step was hidden in **Steps** (a) – (b) above. It is called the **Basis** of the induction!

- The I.H. is usually stated in English: **Assume** $P[y]$ (true), for *all* $y < x$.

Above we admitted much less than what we actually proved. $\mathbb{N}$ does *not* have the monopoly of the CVI methodology in proofs! So let us shift gear and have $<$ indicate in the corollary below an arbitrary order with MC on an arbitrary set A —*not* a set of numbers necessarily.

4.2.1 Corollary. *If $(A, <)$ is a POset with MC, then we can prove a property $P[x]$, for all $x \in A$, by doing precisely the steps of CVI:*

1. *Prove/verify $P[a]$, for every $<$-minimal member of A. This is the* Basis.
2. *Fix an arbitrary b and assume $P[x]$, for all $x < b$. This is the* I.H.
3. *Finally, do the* I.S.: *For the fixed b in 2. prove $P[b]$ using 1. and 2.*

Proof. Nothing changes in the derivation of the equivalence between MC and CVI above. Just forget the opening line "So let $<$ be the standard order on $\mathbb{N}$."!

The only change is in *applying* CVI in the general case is in the *Basis* step: Instead of proving/verifying $P[0]$ for the (unique) *minimum* element of $\mathbb{N}$, we prove/verify $P[x]$ for *all minimal* elements of A, which may be infinitely many! □

There is another simpler induction principle that we call it, well, *simple* induction:

$$\frac{P[0], P[x] \to P[x+1]}{P[x]} \tag{SI}$$

"(SI)" for Simple Induction. That is, to prove $P[x]$ for all x (denominator) do *three* things:

Step 1. Prove/verify $P[0]$

Step 2. **Assume** $P[x]$ for fixed ("frozen") x (unspecified!).

Step 3. **prove** $P[x+1]$ for that same x. The assumption is the I.H. for simple induction. The I.S. is the step that proves $P[x+1]$.

Note that what is described here is precisely an application of the Deduction theorem towards proving "$P[x] \to P[x+1]$", that is, **proving the implication for any given** x.

Step 4. If you have done **Step** 1. through **Step** 3. above, then you **have proved** $P[x]$ (for all x is implied!)

Is the principle (SI) *correct*? I.e., if I do all that the numerator of (SI) asks me to do (or **Steps** 1. – 3.), then do I *really* get that the denominator is true (for all x implied)?

4.2.2 Theorem. *The validity of (SI) is a consequence of MC on* $\mathbb{N}$*.*

Proof. Suppose (SI) is *not* correct. Then, for some property $P[x]$, despite having completed **Steps** 1. – 3., yet, $P[x]$ is *not true* for all x!

Well, if so, let $n \in \mathbb{N}$ be *smallest* such that $P[n]$ is *false*. Now, $n > 0$ since I *did* verify the truth of $P[0]$ (**Step** 1.). Thus, $n - 1 \geq 0$. But then, when I proved "$P[x] \to P[x+1]$ for all x (in $\mathbb{N}$)" —in **Steps** 2. and 3.— this includes **proving** the case

$$P[n-1] \to P[n] \tag{4}$$

But by the smallest-ness of n, $P[n-1]$ is *true*, hence $P[n]$ is true by the truth table of "$\to$". I have just got a contradiction! I conclude that no such smallest n exists, i.e., $P[x]$ is true (for all $x \in \mathbb{N}$). (SI) works! □

How do the simple and course-of-values induction relate? They are equivalent tools! Here is why:

4.2.3 Theorem. *From the validity of (SI) I can obtain the validity of (CVI).*

Proof. Suppose that I have

verified the numerator of (CVI), for $P[x]$, via **Steps** (a) and (b) p.93 (†)

but let me pretend that

I do not know if doing so guarantees the truth of the denominator, $P[x]$ (‡)

Let me show that it does, by doing simple induction SI using a related property, $Q[x]$.

I define $Q[x]$, for all x in $\mathbb{N}$, by

$$Q[x] \stackrel{Def}{\equiv} P[0] \wedge P[1] \wedge \ldots \wedge P[x] \tag{5}$$

Now, as we emphasised on p.92, "property" is colloquial for *formula*. But formulas do *not* have variable length! The length of $Q[x]$ above increases or decreases with the value of its input n. Well, (5) is also a colloquialism to keep things intuitively clear! The mathematically correct definition of Q is the following,

$$Q[x] \stackrel{Def}{\equiv} (\forall z)(z < x \to P[z]) \tag{5$'$}$$

but now that the point has been made, I will continue using the form (5).

So, my job is to show that

if for some property $P[x]$ I proved the truth of the numerator of (CVI), then

it is guaranteed that $P[x]$ is *true*, for all x (6)

I prove this by showing property $Q[x]$ is true, for all x, using SI.

To this end I have to do

SI 1) Verify $Q[x]$ for $x = 0$ (Basis). But $Q[0]$ —by (5)— is just $P[0]$, which I proved *true* as part of my due Basis for CVI (blue underlined if-clause above).

SI 2) For $x > 0$, show,

$$Q[x - 1] \to Q[x] \text{ is true} \tag{7}$$

I argue that I already showed (7) by proving the CVI numerator:

- I proved
$$P[0] \wedge P[1] \wedge \ldots \wedge P[x - 1] \to P[x]$$
- By tautological implication from the above I get also
$$P[0] \wedge P[1] \wedge \ldots \wedge P[x - 1] \to P[0] \wedge P[1] \wedge \ldots \wedge P[x - 1] \wedge P[x]$$
- But the above says $Q[x - 1] \to Q[x]$ is true. This is (7).

By SI, I have proved $Q[x]$ is true, for all x. But by (5), this trivially implies that $P[x]$ is true, for all x. I proved (6). □

4.2.4 Remark.

1. So, for $\mathbb{N}$, MC, CVI and SI **are all equivalent**. We have already indicated that MC and CVI are equivalent. The work on CVI vs. SI (4.2.3) and SI vs. MC (4.2.2) is summarised as
$$MC \Longrightarrow SI \Longrightarrow CVI \Longrightarrow MC$$
which establishes the equivalence claim about all three.
2. When do I use CVI and when SI? SI is best to use when to prove $P[x]$ (in the I.S.) I only need to know $P[x-1]$ is true. CVI is used when we need a more flexible I.H. that $P[n]$ is true for all $n < x$. See the examples below!
3. "0" is the boundary case if the claim we are proving is valid "for all $n \in \mathbb{N}$", or simply put, "for $n \geq 0$". If the claim is "for all $n \geq a$, $P[n]$ is true" then usually $P[n]$ is meaningless for $x < a$ and thus <u>the Basis is for $n = a$</u>. □

4.2.5 Example. This is the "classical first example of induction use" in the discrete math bibliography! Prove that

$$0 + 1 + 2 + \ldots + n = \frac{n(n + 1)}{2} \tag{1}$$

So, the property to prove is the entire expression (1). On must learn to not have to rename the "properties to use" as "$P[n]$".

<u>I will use SI</u>. So let us do the *Basis*. Boundary case is $n = 0$. We verify: $lhs = 0$. $rhs = (0.1)/2 = 0$. Good!

Fix n and tale the expression (1) as I.H.

Do the I.S. Prove:

$$0 + 1 + 2 + \ldots + n + (n + 1) = \frac{(n + 1)(n + 2)}{2}$$

Here it goes

$$\begin{aligned} 0 + 1 + 2 + \ldots + n + (n + 1) &\overset{\text{using I.H.}}{=} \frac{n(n + 1)}{2} + (n + 1) \\ &= (n + 1)(n/2 + 1) \\ &= \frac{(n + 1)(n + 2)}{2} \end{aligned}$$

□

I will write more concisely in the examples that follow.

4.2.6 Example. Same as above but doing away with the "0+". Again, I use SI.

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

- *Basis.* $n = 1$: (1) becomes $1 = (1.2)/2$. True.
- Take (1) as I.H. with fixed n.
- I.S.:

$$\begin{aligned} 1 + 2 + \ldots + n + (n + 1) &\overset{\text{using I.H.}}{=} \frac{n(n + 1)}{2} + (n + 1) \\ &= (n + 1)(n/2 + 1) \\ &= \frac{(n + 1)(n + 2)}{2} \end{aligned}$$

□

4.2.7 Example. Prove

$$1 + 2 + 2^2 + \ldots 2^n = 2^{n+1} - 1 \tag{1}$$

By SI.

- Basis. $n = 0$. $1 = 2^0 = 2^1 - 1$. True.
- As I.H. take (1) for fixed n.
- I.S.

$$\begin{aligned} 1 + 2 + 2^2 + \ldots + 2^n + 2^{n+1} &\overset{\text{using I.H.}}{=} 2^{n+1} - 1 + 2^{n+1} \\ &= 2 \cdot 2^{n+1} - 1 \\ &= 2^{n+2} - 1 \end{aligned}$$

□

4.2.8 Example. An inequality! I prove that

$$n < 2^n \tag{1}$$

for all $n \geq 0$.

I do SI on n.

- *Basis.* $0 < 2^0 = 1$ is true.
- As I.H. fix n and assume (1).
- For the I.S. we have $2^{n+1} = 2^n + 2^n$. By the I.H. $2^n > n$ but also $2^n \geq 1$. Thus, adding these two inequalities I get

$$2^{n+1} = 2^n + 2^n > n + 1$$

□

4.2.9 Example. (Euclid) Every natural number $n \geq 2$ is expressible as a product of primes.

A "product" includes the trivial case of **one** factor.

I do CVI (as you will see why!)

- *Basis*: For $n = 2$ we are done since 2 is a prime.[†]
- I.H. Fix an n and assume the claim for all k, such that $2 \leq k < n$.
- I.S.: Prove for n: Two subcases:
  1. If n is prime, then nothing to prove! Done.
  2. If not, then $n = a \cdot b$, where $a \geq 2$ **and** $b \geq 2$. By I.H.[‡] each of a and b are products of primes, thus so is $n = a \cdot b$. □

4.2.10 Example. (Euclid) Every natural number $n \geq 0$ is expressible base-10 as an expression

$$n = a_m 10^m + a_{m-1} 10^{m-1} + \cdots + a_1 10 + a_0 \tag{1}$$

$$\text{where each } a_i \text{ satisfies } 0 \leq a_i < 10 \tag{2}$$

Proof by CVI again. You will see why.

- *Basis.* For $n = 0$ the expression "0" has the form of the rhs of (1) *and* satisfies inequality (2).

[†]You will recall that a number $\mathbb{N} \ni n > 1$ is a *prime* iff its **only** factors are 1 and n.

[‡]You see? a and b cannot be both $n - 1$ to apply SI's I.H. In fact, if $n = (n-1)^2$, then $n = n^2 - 2n + 1$ or $n^2 - 3n + 1 = 0$. This equation has no natural number roots! So SI would *not* help with its rigid I.H.

- Fix an $n > 0$ and assume (I.H.) that if $k < n$, then k can be expressed as in (1).
- For the I.S. express the n of the I.H. using Euclid's theorem (3.1.47) as

$$n = 10q + r$$

where $0 \leq r < 10$. By the I.H. —since $q < n$— let

$$q = b_t 10^t + b_{t-1} 10^{t-1} + \cdots + b_1 10 + b_0$$

with $0 \leq b_j < 10$.

Then

$$\begin{aligned}
n &= 10q + r \\
n &= 10\Big(b_t 10^t + b_{t-1} 10^{t-1} + \cdots + b_1 10 + b_0\Big) + r \\
n &= b_t 10^{t+1} + b_t 10^t + \cdots + b_1 10^2 + b_0 10 + r
\end{aligned}$$

We see n has the right form since $0 \leq r < 10$. □

4.2.11 Example. Another inequality. Let p_n denote the n-th prime number, for $n \geq 0$. Thus $p_0 = 2$, $p_1 = 3$, $p_2 = 5$, etc.

We prove that

$$p_n \leq 2^{2^n} \tag{1}$$

I use CVI on n. This is a bit of a rabbit out of a hat if you never read Euclid's proof that there are infinitely many primes.

- Basis $p_0 = 2 \leq 2^{2^0} = 2^1 = 2$.
- Fix $n > 0$ and take (1) as I.H.
- The I.S.: I will work with the fixed n above and the expression (product of primes, plus 1; this is inspired from Euclid's proof quoted above).

$$p_0 p_1 p_2 \cdots p_n + 1$$

By the I.H. I have

$$\begin{aligned}
p_0 p_1 p_2 \cdots p_n + 1 &\leq 2^{2^0} 2^{2^1} 2^{2^2} \cdots 2^{2^n} + 1 && \text{by I.H.} \\
&= 2^{2^0 + 2^1 + 2^2 + \cdots + 2^n} + 1 && \text{algebra} \\
&= 2^{2^{n+1} - 1} + 1 && \text{by 4.2.7} \\
&= 2^{2^{n+1} - 1} + 2^{2^{n+1} - 1} && \text{smallest } n \text{ possible is } n = 1 \\
&= 2^1 \cdot 2^{2^{n+1} - 1} \\
&= 2^{2^{n+1}}
\end{aligned}$$

Now we have two cases on $q = p_0 p_1 p_2 \cdots p_n + 1$

1. q is a prime. Because of the "$+1$" q is different from all p_i in the product, so q is p_{n+1} or p_{n+2} or p_{n+3} or ...
   Since the sequence of primes is strictly increasing, p_{n+1} is the least that q can be.
   Thus
   $$p_{n+1} \leq p_0 p_1 p_2 \cdots p_n + 1 \leq 2^{2^n}$$
   in this case.
2. q is composite. By 4.2.9 some prime r divides q. Now, none of the
   $$p_0, p_1, p_2, \cdots, p_n$$
   divides q because of the "$+1$". Thus r is different from all of them, so it must be one of p_{n+1} or p_{n+2} or p_{n+3} or ...
   Thus,
   $$p_{n+1} \leq r < q = p_0 p_1 p_2 \cdots p_n + 1 \leq 2^{2^n}$$

Done! □

4.2.12 Example. Let

$$\begin{aligned} b_1 &= 3,\ b_2 = 6 \\ b_k &= b_{k-1} + b_{k-2}, \text{ for } k \geq 3 \end{aligned}$$

Prove by induction that b_n is divisible by 3 for $n \geq 1$. (Be careful to distinguish between what is *basis* and what are *cases* arising from the **induction step**! As you know, our text is careless about this.)

Proof. So the boundary condition is (from the underlined part above) $n = 1$. This is the *Basis*.

1. *Basis*: For $n = 1$, I have $a_1 = 3$ and this is divided by 3. We are good.
2. *I.H.* Fix n and **assume claim** for all $k < n$.
3. *I.S.* **Prove claim** for the above fixed n. There are two cases, as the I.H. is *not useable* for $n = 2$. Why? Because it would require entries b_0 and b_1. The red entry does not exist since the sequence starts with b_1. So,

   Case 1. $n = 2$. Then I am OK as $b_2 = 6$; it *is* divisible by 3.

   Case 2. $n > 2$. Is b_n divisible by 3? Well, $b_n = b_{n-1} + b_{n-2}$ in this case. By I.H. (valid for all $k < n$) I have that $b_{n-1} = 3t$ and $b_{n-2} = 3r$, for some integers t, r. Thus, $b_n = 3(t + r)$. Done! □

Here are a few additional exercises for you to try —please do try!

4.2.13 Exercise.

1. Prove that $2^{2n+1} + 3^{2n+1}$ is divisible by 5 for all $n \geq 0$.

2. Using induction prove that $1^3 + 2^3 + \ldots + n^3 = \left[\dfrac{n(n+1)}{2}\right]^2$, for $n \geq 1$.

3. Using induction prove that $\sum_{i=1}^{n+1} i2^i = n2^{n+2} + 2$, for $n \geq 0$.

4. Using induction prove that $\sqrt{n} < \dfrac{1}{\sqrt{1}} + \dfrac{1}{\sqrt{2}} + \ldots + \dfrac{1}{\sqrt{n}}$, for $n \geq 2$.

5. Let

$$
\begin{aligned}
b_0 &= 1,\, b_1 = 2,\, b_3 = 3 \\
b_k &= b_{k-1} + b_{k-2} + b_{k-3}, \text{ for } k \geq 3
\end{aligned}
$$

   Prove by induction that $b_n \leq 3^n$ for $n \geq 0$. (Once again, be careful to distinguish between what is *basis* and what are *cases* arising from the **induction step**!) □