0.1. Notes on Arithmetic

Peano Arithmetic is a formal theory, kind of "applied logic", that talks about the behaviour of our familiar objects—the natural numbers $0, 1, 2, \ldots$—and the operations and relations on them *formally.* That is, purely syntactically, without mixing "expectations" with what has been *actually* assumed (axioms).

Of course, expectations lead us to the appropriate choice of axioms, but once the choice has been effected, the expectations are put aside, and give way to formal proving.

This "formal proving" is supposed to help us discover true statements about the "real" objects that our axiomatic system *simulates. Soundness* of all first order theories, and the fact that the "natural interpretation" of our axioms is valid over $\mathbb{N}$, guarantee that we will never prove any invalid formulas.

A celebrated theorem by Gödel, however, says that we cannot prove *all* true statements about the natural numbers, *no matter how we axiomatize their behaviour* in a "practical system".[*]

The nonlogical symbols are the following

1. "0" *When interpreted naturally*, it means "zero".
2. "S" *When interpreted naturally*, it means "+1", the successor function of one variable (i.e., $S(x)$ or more simply Sx is interpreted as $x + 1$).
3. "+" *When interpreted naturally*, it means "+", the "plus" function of two variables (i.e., the formal $x + y$ is interpreted as the "real" or informal $x + y$).
4. "·" *When interpreted naturally*, it means "·", the "times" function of two variables (i.e., the formal $x \cdot y$ is interpreted as the "real" or informal $x \cdot y$).
5. "<" *When interpreted naturally*, it means "<", the "less than" relation of two variables (i.e., the formal $x < y$ is interpreted as the "real" or informal $x < y$).

What about "1", "2"? What about $\leq$ and other things? Like exponentiation, etc? All these "other" things can be introduced *definitionally.* For example we can write $S0$ and $SS0$ as the formal counterparts of the real (i.e., whatever we, philosophically, think are ***really*** the numbers) 1 and 2 respectively. *Formal* abbreviations for $S0$ and $SS0$ are $\widetilde{1}$ and $\widetilde{2}$ respectively. Similarly, we write (formally) $\widetilde{n}$ as short for

$$\overbrace{S \ldots S}^{n \text{ copies}} 0$$

[*]A system of axioms is "practical" if we can tell in a finite number of steps whether a formula is an axiom or not. For example, our logical axioms Ax1–Ax6 are "practical". Even though Ax1 contains infinitely many formulas, we can certainly test in a finite number of steps whether or not a formula is a partial generalization of a tautology.

Note that due to a theorem of Church we cannot take all formulas that are valid in $\mathbb{N}$ as axioms, for this system, he proved, is *not* practical.

Of course this interprets *naturally* as the "real" natural number "n".

We are *not* going to use this rigid notation in Ch.12. Only at the beginning, briefly, so you can get a flavour of what formal arithmetic is *really* like. After that we will revert to the "quasi-formal-quasi-sloppy" text style (sigh ...)

The axioms of Peano Arithmetic (PA) are the *universal closures*† of the following:

(S)**1.** $\neg 0 = Sx$

(S)**2.** $Sx = Sy \Rightarrow x = y$ ("1-1ness of S")

$(+)$**1.** $x + 0 = x$

$(+)$**2.** $x + Sy = S(x + y)$

$(\cdot)$**1.** $x \cdot 0 = 0$

$(\cdot)$**2.** $x \cdot Sy = x \cdot y + x$

$(<)$**1.** $\neg x < 0$

$(<)$**2.** $x < Sy \equiv x < y \vee x = y$

$(<)$**3.** $x < y \vee x = y \vee y < x$

And the **Induction Schema**, one axiom for each formula A:

(Ind) $$A[x := 0] \wedge (\forall x)(A \Rightarrow A[x := Sx]) \Rightarrow A$$

The totality of the above axioms (i.e., the universal closures of what I have just listed) are going to be denoted by "PA".

If we keep all except (Ind) we have "Robinson's Arithmetic", in short, "ROB".

Writing $A[x]$ to indicate our interest in the free variable x of A‡ we rewrite (Ind) as follows:

(Ind$'$) $$A[0] \wedge (\forall x)(A[x] \Rightarrow A[Sx]) \Rightarrow A[x]$$

This yields, by *modus ponens* the derived rule of PA:

$$A[0] \wedge (\forall x)(A[x] \Rightarrow A[Sx]) \vdash A[x]$$

†If A is a formula, its universal closure is obtained by adding $(\forall x)$ in front of A for *each free variable of* A. The order of the various $(\forall x)$ is immaterial, since $\vdash (\forall x)(\forall y)A \equiv (\forall y)(\forall x)A$.

‡Recall that x need not be actually free. Compare: We write in Calculus "$f(x)$" to indicate a function's dependence of x. Maybe we can later prove that $f'(x)$—the derivative—equals 0 on an open interval (a, b). Then $f(x)$ is constant on (a, b), i.e., it does *not* depend on x after all. This does not stop us from writing "$f(x)$" nevertheless.

or, separating the premises[§]

(Ind$''$) $$A[0],\ (\forall x)(A[x] \Rightarrow A[Sx]) \vdash A[x]$$

The above is pretty close to the "practical" induction "protocol". A few more simplifications and we get to that: Note what (Ind$''$) offers us:

To prove
$$\mathrm{PA} \vdash A[x] \tag{1}$$
or, equivalently,
$$\mathrm{PA} \vdash (\forall x)A[x] \tag{2}$$
since PA consists of closed formulas, ***it suffices to do instead***

(a) $\mathrm{PA} \vdash A[0]$

and

(b) $\mathrm{PA} \vdash (\forall x)(A[x] \Rightarrow A[Sx])$

However, to do (b) we only need do simply

$$\mathrm{PA} \vdash A[x] \Rightarrow A[Sx]$$

We then get (b) by generalization! We can get away with this because our "Γ" here is PA, and we have ensured that it has *no free variables.* Thus, our "practical induction protocol" is the following *RULE*:

(Ind-rule) $$A[0],\ A[x] \Rightarrow A[Sx]\ \vdash\ A[x]$$

which is implemented (used) in practice, via the Deduction theorem, as follows:

(I) $\mathrm{PA} \vdash A[0]$. This is the *Basis.*

and

(II) Add $A[x]$ to PA as a new axiom ("auxiliary" or "temporary axiom"). This is the *Induction Hypothesis*, or "I.H."

(III) The "GoTo step" (goto "Sx", or goto "$x+1$"): Do

PA, $A[x] \vdash A[Sx]$

Jargon When we prove $\mathrm{PA} \vdash (\forall x)A[x]$ by induction, or—as we often omit the "$(\forall x)$"—$\mathrm{PA} \vdash A[x]$, we say that we do induction on x.

Let us sample a couple of applications that do not murky the waters by assuming things we never assumed (such as summations, for example).

[§]Legitimate by the tautological implications $P \wedge Q \models_{\mathrm{taut}} P$, $P \wedge Q \models_{\mathrm{taut}} Q$ and $P, Q \models_{\mathrm{taut}} P \wedge Q$.

0.1.1 Example. We prove $\mathrm{PA} \vdash 0 \le x$ by induction on x.

First off, "$\le$" is a *defined* symbol introduced by

$$x \le y \text{ stands for } x < y \vee x = y$$

As we know from class such "stands for" definitions give rise to tautologies

$$x \le y \equiv x < y \vee x = y$$

Onto our task. $A[x]$ is $0 \le x$.

Basis. I want $\mathrm{PA} \vdash 0 \le 0$.

$$\begin{aligned}
& 0 \le 0 \\
\equiv & \left\langle \text{Definition of } \le \right\rangle \\
& 0 < 0 \wedge 0 = 0
\end{aligned}$$

The last formula is an absolute theorem[†] (PA not used), hence so is $0 \le 0$.

<u>Add to PA the assumption $A[x]$ (I.H.)</u>

Goto step (prove $A[Sx]$).

$$\begin{aligned}
& 0 \le Sx \\
\equiv & \left\langle \text{Definition of } \le \right\rangle \\
& 0 < Sx \vee 0 = Sx \\
\equiv & \left\langle \text{Leib. plus axiom ``}(<)\mathbf{2.}\text{''} \right\rangle \\
& 0 < x \vee 0 = x \vee 0 = Sx \\
\equiv & \left\langle \text{Leib. plus def. of } \le \right\rangle \\
& 0 \le x \vee 0 = Sx
\end{aligned}$$

The last line is a tautological consequence of I.H., so we are done. $\square$

0.1.2 Example. We prove $\mathrm{PA} \vdash x < y \wedge y < z \Rightarrow x < z$ by induction on z.

That is, the formula "$x < y \wedge y < z \Rightarrow x < z$" is our "$A[z]$".

Basis. I want $\mathrm{PA} \vdash x < y \wedge y < 0 \Rightarrow x < 0$. This is a tautological consequence of $\neg y < 0$ (axiom $(<)$**1.**)

<u>Add to PA the assumption $A[z]$ (I.H.)</u>

[†] Specialization of axiom $(\forall z)(z = z)$, followed by tautological implication.

Goto step (prove $A[Sz]$).

$$
\begin{aligned}
& x < y \wedge y < Sz \Rightarrow x < Sz \\
\equiv & \Big\langle \text{Leib. plus axiom } (<)\mathbf{2.} \Big\rangle \\
& x < y \wedge (y < z \vee y = z) \Rightarrow x < Sz \\
\equiv & \Big\langle \text{Leib. plus tautology} \Big\rangle \\
& (x < y \wedge y < z) \vee (x < y \wedge y = z) \Rightarrow x < Sz \\
\equiv & \Big\langle \text{tautology} \Big\rangle \\
& (x < y \wedge y < z \Rightarrow x < Sz) \wedge (x < y \wedge y = z \Rightarrow x < Sz)
\end{aligned}
$$

The last line is a theorem by I.H. and axiom $(<)$**2**, so we are done.

You believe this? Here:
(1) $x < y \wedge y < z \Rightarrow x < z$ (The I.H.)
(2) $x < z \Rightarrow x < z \vee x = z$ (tautology)
(3) $x < z \Rightarrow x < Sz$ (taut. impl. from (2) plus axiom $(<)$**2.**)
(4) $x < y \wedge y < z \Rightarrow x < Sz$ ((1) plus (3) plus taut. impl.)
This justifies the left conjunct.

For the right conjunct, Ax6 tautologically yields $\vdash x < y \wedge y = z \Rightarrow x < z$. Now use (3) above to get $x < y \wedge y = z \Rightarrow x < Sz$ by taut. implication. □

0.1.3 Example. ("Strong" or "course-of-values" induction)

We now derive a useful induction principle (the one proposed up in front in GS) that goes by any of the above names. Its I.H. is "strong" and uses the "entire course of values (history)" of the induction variable. Naïvely, to prove $P(n)$ (n of "type" $\mathbb{N}$) you assume (I.H.) the claim for all $k < n$. Helped by this you prove $P(n)$. You then proclaim that you have proved $P(n)$ for all $n \in \mathbb{N}$. Here is the formal reason why this technique is "correct":

For every formula $A[x]$,

$$\mathrm{PA} \vdash (\forall x)\big((\forall z < x)A[z] \Rightarrow A[x]\big) \Rightarrow (\forall x)A[x] \qquad (CVI)$$

We have used the "bounded quantification" abbreviation above that is very common in the literature: We wrote "$(\forall z < x)A[z]$" for "$(\forall z)\big(z < x \Rightarrow A[z]\big)$".

To prove (CVI) (for any given formula A) we will employ the Deduction theorem to eliminate the rightmost $\Rightarrow$.

$$\text{Thus, we } \underline{\text{add } (\forall x)\big((\forall z < x)A[z] \Rightarrow A[x]\big) \text{ to PA.}} \qquad (0)$$

Next, we name the formula

$$(\forall z < x)A[z] \qquad (1)$$

by the short name $B[x]$.

Plan: Prove $\mathrm{PA} \vdash (\forall x)B[x]$. We use induction on x.

Basis. Do $\mathrm{PA} \vdash B[0]$. By (1), do

$$\mathrm{PA} \vdash (\forall z)(z < 0 \Rightarrow A[z])$$

We calculate this here:

$$\begin{aligned}
&(\forall z)(z < 0 \Rightarrow A[z])\\
\equiv &\Big\langle \text{axiom } (<)\mathbf{1.}\text{, redundant true, and sWLUS} \Big\rangle\\
&(\forall z)(\mathit{false} \Rightarrow A[z])
\end{aligned}$$

The last formula is in Ax1.

With the basis out of the way, add to PA the assumption $B[x]$. (I.H.)

For future use: (0) above, via specialization, yields

$$B[x] \Rightarrow A[x] \tag{2}$$

Goto step. (Prove $B[Sx]$)

$$\begin{aligned}
&(\forall z)(z < Sx \Rightarrow A[z])\\
\equiv &\Big\langle \text{axiom } (<)\mathbf{2.} \text{ and sWLUS} \Big\rangle\\
&(\forall z)(z < x \lor z = x \Rightarrow A[z])\\
\equiv &\Big\langle \text{WLUS and obvious tautology} \Big\rangle\\
&(\forall z)\Big((z < x \Rightarrow A[z]) \land (z = x \Rightarrow A[z])\Big)\\
\equiv &\Big\langle \text{distr. } \forall \text{ over } \land \Big\rangle\\
&(\forall z)(z < x \Rightarrow A[z]) \land (\forall z)(z = x \Rightarrow A[z])\\
\equiv &\Big\langle \text{Leib. plus 1-point-rule} \Big\rangle\\
&B[x] \land A[x]
\end{aligned}$$

The last line is a theorem, because $B[x]$ is (I.H.), and hence so is $A[x]$ by (2) and MP.

We now have that PA along with our first underlined assumption (the one just above (1)) proves

$$(\forall x)B[x]$$

By $\forall$-monotonicity and (2),

$$(\forall x)A[x]$$

By the Deduction theorem we are done! $\square$

0.1.4 Example. (Practicalities of (CVI)) We employ "strong" induction like this: To prove $\text{PA} \vdash (\forall x)A[x]$ just do the following two things.

1. Assume $(\forall z)(z < x \Rightarrow A[z])$. This is the I.H. We also say it informally: "Assume that, for all $z < x$, $A[z]$ holds".

2. Prove, using PA and the I.H., $A[x]$. This is the **Goto step**.

Indeed, steps 1. and 2. yield

$$\text{PA} \vdash (\forall z)(z < x \Rightarrow A[z]) \Rightarrow A[x]$$

by Deduction theorem. Since all the axioms of PA are closed, generalization is allowed and gives

$$\text{PA} \vdash (\forall x)\Big((\forall z)(z < x \Rightarrow A[z]) \Rightarrow A[x]\Big) \tag{3}$$

Since (CVI) is a PA theorem by our work in Example 0.1.3, $\text{PA} \vdash (\forall x)A[x]$ by (3) and modus ponens. $\square$

0.1.5 Example. (Hey! What about the basis?) The basis in (CVI) is part of the **Goto step**. Here's why: There is one case where the "Goto step" is not helped by the I.H. in 1. of the above example (0.1.4). That is when PA can prove the I.H. In that case we prove $A[x]$ just from PA for our "assumption" (I.H.) is a PA-theorem anyway—that is, it is not an *additional* assumption. So, when does this happen? That is if $x = 0$: Then $z < 0 \Rightarrow A[z]$ equivales to *true* in PA, since $z < 0$ equivales to *false* in PA by the first $<$-axiom. Thus we have to prove $A[0]$ from scratch.

In plain English (informally), since there is no z to the left of 0 in $\mathbb{N}$, there is no hypothesis "for all $z < 0$, $A[z]$ holds" to use when proving $A[0]$. The moral is: "You cannot escape from the Basis". $\square$

Both simple induction and CVI can be used to prove statements such as $(\forall n \geq n_0)A[n]$, where n_0 is a fixed positive number.

The two recipes are respectively:

(1) Simple: Do

(i) Verify $A[n_0]$ (Basis)

(ii) Assume $A[n]$ (for $n \geq n_0$)—this is the I.H.

(iii) Prove that $A[n+1]$ (for $n+1 \geq n_0$)—this is the "goto step".

(2) CVI: Do

(a) Assume that $A[m]$ holds for all m such that $n > m \geq n_0$—this is the I.H.

(b) Prove that $A[n]$ holds, ensuring that your argument is good for any $n \geq n_0$—this is the "goto step".

Note that doing (b) above you have to directly verify $A[n_0]$ because the I.H. is *true* and thus it does not help when you want to prove *true* $\Rightarrow A[n_0]$.[¶] We say that $n = n_0$ is the *boundary case*.

0.1.6 Example. Here we show the CVI in action. However, we retreat to informal mathematics, where we assume a lot. In particular we assume that we know what a prime is and that a factor of a factor of $n \in \mathbb{N}$ is a factor of n.[‖] Moreover, we will carry out the proof informally.

We want to prove:

$$\text{Every } n \geq 2 \text{ has a prime factor} \tag{1}$$

Following the recipe (a)–(b) above
1) **I.H.** Assume that if $2 \leq m < n$ then m has a prime divisor.
2) Now prove that $n \geq 2$ has a prime divisor.
There are two cases:

Case I. (The boundary case) If $n = 2$, then we are done: 2 has a prime factor, itself, since $2 = 2 \cdot 1$.

Case II. $n > 2$. Well, if n is prime we are done—just write $n = n \cdot 1$. If on the other hand it is composite, then $n = a \cdot b$ where $a > 1$ and $b > 1$. But $a > 1$ says $a \geq 2$. Thus $n > a \geq 2$ and the I.H. applies to a. That is, a has a prime factor. But any prime factor of a is a prime factor of n. Done.

□

[¶] The hypothesis is "for all m such that $n_0 > m \geq n_0$, $A[n_0]$ holds" or, in symbols, $n_0 > m \geq n_0 \Rightarrow A[n_0]$. But $n_0 > m \geq n_0$ equivales to *false*.

[‖] Well, OK: We are in $\mathbb{N}$. A *factor* or *divisor* of n is an m such that $(\exists x)n = m \cdot x$, where "We are in $\mathbb{N}$" means that whenever you write "$(\exists x)\ldots$" or "$(\forall x)\ldots$" you mean "$(\exists x \in \mathbb{N})\ldots$" and "$(\forall x \in \mathbb{N})\ldots$" respectively. Note that if $(\exists y)m = d \cdot y$ as well then $(\exists y)m = d \cdot y \wedge (\exists x)n = m \cdot x$, which is equivalent to $(\exists y)(\exists x)(m = d \cdot y \wedge n = m \cdot x)$ (you believe this?), hence (Ax6 and $\exists$-monotonicity), $(\exists y)(\exists x)(n = d \cdot y \cdot x)$, thus $(\exists z)n = d \cdot z$. d is a factor of n. Now, n is a *prime* if it is > 1 and all its factors are n and 1.