MA5219 LOGIC AND FOUNDATION OF MATHEMATICS I

Lecture 11: Elementary Arithmetic

Tin Lok Wong

17 September, 2018

The aim of this lecture is to describe how one can prove the key properties of the formulas defined in the previous lecture in a theory called $\mathrm{I}\Delta_0(\exp)$.

Definition. The $\mathcal{L}_{\mathrm{A}}(\exp)$ theory $\mathrm{I}\Delta_0(\exp)$ consists of the elements of $\mathrm{Q}(\exp)$ and an induction axiom

$$\forall \bar{z}\ \big(\theta(0,\bar{z}) \wedge \forall x\ \big(\theta(x,\bar{z}) \to \theta(x+1,\bar{z})\big) \to \forall x\ \theta(x,\bar{z})\big)$$

for each $\Delta_0(\exp)$ formula $\theta(x,\bar{z})$.

The choice of the base theory $\mathrm{Q}(\exp)$ is usually not important mathematically, as long as it is finite (and true in $\mathbb{N}$). Philosophically, one would also like the base theory to be of a purely algebraic nature. For example, some authors use the $\mathcal{L}_{\mathrm{A}}(\exp)$ theory PA^- as their base theory, which axiomatizes the non-negative parts of discretely ordered rings. The two choices of base theory make equivalent definitions for $\mathrm{I}\Delta_0(\exp)$.

Lemma 11.1. $\mathrm{I}\Delta_0(\exp)$ proves the $\mathcal{L}_{\mathrm{A}}(\exp)$ theory PA^-, which consists of the following sentences.

(assoc+)	$\forall x,y,z$	$(x+y)+z = x+(y+z)$.
(comm+)	$\forall x,y$	$x+y = y+x$.
(assoc×)	$\forall x,y,z$	$(x\times y)\times z = x\times(y\times z)$.
(comm×)	$\forall x,y$	$x\times y = y\times x$.
(distr)	$\forall x,y,z$	$x\times(y+z) = (x\times y)+(x\times z)$.
(zero-one)	$\forall x$	$(x+0=x \wedge x\times 0 = 0 \wedge x\times 1 = x)$.
(irrefl)	$\forall x$	$x \not< x$.
(tran)	$\forall x,y,z$	$(x<y \wedge y<z \to x<z)$.
(lin)	$\forall x,y$	$(x<y \vee x=y \vee y<x)$.
(+/<)	$\forall x,y,z$	$(x<y \to x+z<y+z)$.
(×/<)	$\forall x,y,z$	$(x<y \wedge z>0 \to x\times z<y\times z)$.
(<+)	$\forall x,y$	$(x<y \to \exists z\ \ x+z=y)$.
(discrete)	$\forall x$	$(x>0 \to x=1 \vee x>1) \quad \wedge \quad 0<1$.
(non-neg)	$\forall x$	$(x=0 \vee x>0)$.

Proof. We omit the proof because, if one has doubt, then one can include PA^- as part of the base theory for $\mathrm{I}\Delta_0(\exp)$. □

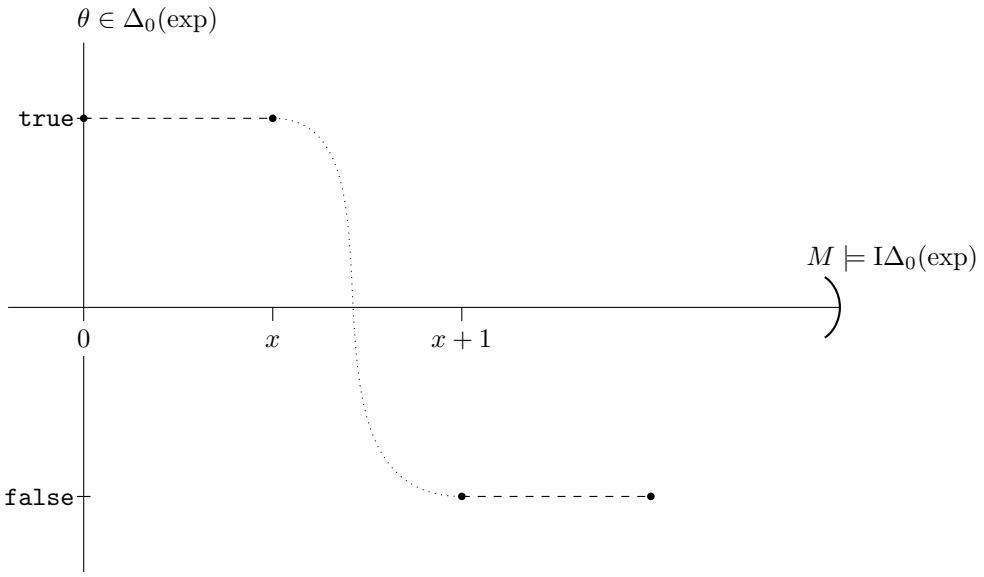

Figure 11.1: Induction as an intermediate-value property

We will use (assoc+), (comm+), (assoc×), (comm×), (distr), (zero-one), (tran) and (lin) so frequently that references to them are often omitted.

Induction axioms are what distinguish arithmetic from algebra. They say that, if the base case is true and the induction step holds for a formula, then this formula is true universally. Model-theoretically, the induction axiom for a formula θ asserts that if the set of elements satisfying θ contains 0 and is closed under adding-one, then it must contain all elements. Algebraically, the induction axiom for θ tells us if θ is true at 0 but false at some other point, then there must be a point at which the truth value of θ changes from true to false in one step; see Figure 11.1.

The way we stated induction is sometimes called *step induction*. Induction can appear in other guises, for example, as strong induction (in which the induction hypothesis involves all previous cases), and as the least number principle (which asserts the existence of least witnesses). These alternative versions are sometimes more convenient to use. Nevertheless, it is fact that they are all equivalent at the $\Delta_0(\exp)$ level.

The theory $\mathrm{I}\Delta_0(\exp)$ and its numerous equivalent formulations are widely studied. It is sometimes called *Elementary Arithmetic* or *Exponential Function Arithmetic*. It is intended to axiomatize the part of finite mathematics that does not involve the use of objects of superexponential sizes. It is capable of proving many arithmetic facts. Let us start with some basic ones. Most syntactic entailments we see in this module are proved in the semantic way via an implicit use of the Completeness Theorem. Notice that (<S) says our non-standard definition of $x \leqslant y$ as $x < y + 1$ agrees with the usual one over $\mathrm{I}\Delta_0(\exp)$.

Remark 11.2. If there is no risk of ambiguity, then when we instantiate a quantifier by an element of a structure, we may use the same letter for the element and the quantified variable to make the connection visible in the notation. This is *not* intended to suggest that elements of our structures are variables, or vice versa.

Lemma 11.3. $\mathrm{I}\Delta_0(\exp)$ proves

$(+^n)$ $\forall x$ $\quad x \times (\underline{n+1}) = \underbrace{x + x + \cdots + x}_{(n+1)\text{-many } x\text{'s}}$ for all $n \in \mathbb{N}$;

(S>) $\forall x$ $\quad x + 1 > x$;

(<+S) $\forall x, y$ $\quad (x < y \leftrightarrow \exists z \;\; x + z + 1 = y)$;

(<S) $\forall x, y$ $\quad (y < x + 1 \leftrightarrow y = x \vee y < x)$;

(exp×) $\forall x, y \quad 2^x \times 2^y = 2^{x+y}$;

(exp>) $\forall x \quad 2^x > x$;

(exp/<) $\forall x, y \quad (x < y \to 2^x < 2^y)$;

Proof sketch. Work in an arbitrary $M \models \mathrm{I}\Delta_0(\exp)$.

($+^n$) First rewrite $\underline{n+1}$ as $\underbrace{1+1+\cdots+1}_{(n+1)\text{-many 1's}}$ using (R+) and (R1). Then apply (distr) and (zero-one).

(S>) Use the second conjunct in (discrete) and (+/<).

(<+S) Take $x, y \in M$. First, suppose $x < y$. Then (<+) gives $w \in M \models x + w = y$. Notice that (zero-one) and (irrefl) imply $w \neq 0$. So we are done by (QS_0). Conversely, let $z \in M \models x + z + 1 = y$. In view of (non-neg), either $z = 0$ or $z > 0$. If $z = 0$, then $y > x + z = x$ by (S>). If $z > 0$, then $y > x + z > x$ by (S>) and (+/<). Thus (irrefl) tells us $x \neq y$ in either case. From this we deduce that $x < y$ by (Q<).

(<S) The $\leftarrow$ direction follows from (S>) and (tran). For the $\rightarrow$ direction, use (<+S), (QS_1), then (Q<).

(exp×) Fix $x \in M$. We show by induction on y in M that $M \models \forall y\ (2^x \times 2^y = 2^{x+y})$. The base case when $y = 0$ follows from (Qexp) and (zero-one). For the induction step, apply (Qexp_1), (distr), the induction hypothesis, and then (Qexp_1) to $2^x \times 2^{y+1}$ to get 2^{x+y+1}.

(exp>) Proceed by induction on x in M. For the base case $x = 0$, use (Rexp) and (R<). For the induction step, split into two cases. If $x = 0$, then one deduces $2^{x+1} > x + 1$ using (R1), (Rexp) and (R<). If $x \neq 0$, then $x \geqslant 1$ by (non-neg) and (discrete), and thus $2^{x+1} = 2^x + 2^x > x + x \geqslant x + 1$ by (Qexp), (+/<), and (+/<) respectively.

(exp/<) Fix $x \in M$. We show by induction on z in M that $M \models \forall z\ (2^x < 2^{x+z+1})$. This is sufficient in view of (<+S). We only show the induction step here because the base case can be shown similarly. Let $z \in M \models 2^x < 2^{x+z+1}$. Recall ($\mathrm{Qexp}_1$) tells us that $2^{x+(z+1)+1} = 2^{x+z+1} + 2^{x+z+1}$, which is at least 2^{x+z+1} by (non-neg) and (+/<). Then apply the induction hypothesis. □

All functionality and boundness properties of the formulas introduced in the previous lecture can be proved in $\mathrm{I}\Delta_0(\exp)$. The verification is long and routine, but one should convince oneself that this is possible. For most people, it is more enlightening to write than to read such verifications. So we only show one here as a demonstration.

Lemma 11.4. $\mathrm{I}\Delta_0(\exp) \vdash \forall s\ \exists!\ell\ \mathrm{len}(s) = \ell$.

Proof. Fix $s \in M \models \mathrm{I}\Delta_0(\exp)$. If $s = 0$, then it is easy because $M \models \forall \ell\ \big(\mathrm{len}(s) = \ell \leftrightarrow \ell = 0\big)$ by the definition of len. So suppose $s \neq 0$.

Existence Notice $2^{4s} > 4s = s + s + s + s \geqslant s + 0 = s$ by (exp>), ($+^n$), (non-neg), and ($\mathrm{Q}+_0$) respectively. In particular, we know $M \models \exists x\ (s < 2^{4x})$. Apply the contrapositive of $\Delta_0(\exp)$ induction in M to find $k \in M$ such that

$$M \models s < 2^{4\times 0} \vee (2^{4k} \leqslant s \wedge s < 2^{4(k+1)}).$$

Since we are in the case when $s \neq 0$, we know $s \not< 1$ by (RInit). So the first disjunct above is false because (R×) and (Rexp) implies $2^{4\times 0} = 1$. Hence the second disjunct holds. Let $\ell = k + 1$. Then $k = \ell - 1$ and so $M \models \mathrm{len}(s) = \ell$.

Uniqueness Let $\ell \in M \models \mathrm{len}(s) = \ell$. Take any $\ell' \in M$. Assume $\ell' < \ell$. Then (non-neg) implies $0 \leqslant \ell' < \ell$ and so $\ell \geqslant 1$ by (discrete). Thus $\ell - 1$ exists in M. Now

$$\begin{array}{lll} & \ell' < \ell = (\ell - 1) + 1 & \text{by the definition of } \ell - 1; \\ \therefore & \ell' \leqslant \ell - 1 & \text{by } (<\mathrm{S}); \\ \therefore & 2^{\ell'} \leqslant 2^{\ell - 1} < s & \text{by } (\exp/<) \text{ and the definition of len, as } s \neq 0; \\ \therefore & s \not\leqslant 2^{\ell'} & \text{by (tran) and (irrefl)}; \\ \therefore & M \models \mathrm{len}(s) \neq \ell' & \text{by the definition of len, as } s \neq 0. \end{array}$$

By symmetry, we see that $\ell < \ell'$ implies $M \models \mathrm{len}(s) \neq \ell'$ too. So if $M \models \mathrm{len}(s) = \ell'$, then we must have $\ell' = \ell$ by (lin). □

In view of Lemma 11.4, in every $M \models \mathrm{I}\Delta_0(\exp)$, the formula $\mathrm{len}(s) = \ell$ defines the graph of a function $M \to M$. When there is no risk of ambiguity, we denote this function also by len. The same applies to other functions we define in a theory.

Many other key properties of the formulas introduced in the previous lecture are also provable in $\mathrm{I}\Delta_0(\exp)$. The following is one example.

Lemma 11.5. $\mathrm{I}\Delta_0(\exp)$ proves

(1) $\forall t, t' \ \big(t \neq 0 \to \mathrm{len}(t \mathbin{\widehat{}} t') = \mathrm{len}(t) + \mathrm{len}(t') + 1\big)$; and

(2) $\forall t, t' \ \big(\forall j{<}\mathrm{len}(t') \ \langle t \mathbin{\widehat{}} t' \rangle_j = \langle t' \rangle_j \wedge \langle t \mathbin{\widehat{}} t' \rangle_{\mathrm{len}(t')} = 0 \wedge \forall i{<}\mathrm{len}(t) \ \langle t \mathbin{\widehat{}} t' \rangle_{\mathrm{len}(t')+i+1} = \langle t \rangle_i\big)$.

Assignment 11.6. Prove Lemma 11.5(2) from Lemma 11.3 and Lemma 11.5(1). You may assume that all the functionality and boundedness properties stated in the previous lecture are provable in $\mathrm{I}\Delta_0(\exp)$. [7 points]

Let us see a general way of utilizing $\mathrm{I}\Delta_0(\exp)$ to build numbers with specific hexadecimal representations. It can give us a feeling of the strength of $\mathrm{I}\Delta_0(\exp)$. It is similar to the Axiom of Separation in set theory, except that instead of two truth values for membership, we have 16.

Notation. If $n \in \mathbb{N}$, then $\bigvee_{d<n} \theta_d = \theta_0 \vee \theta_1 \vee \cdots \vee \theta_{n-1}$.

Bounded $\Delta_0(\exp)$ comprehension (hexadecimal version)**.** Let $\theta_0(x, \bar{u}), \theta_1(x, \bar{u}), \ldots, \theta_{15}(x, \bar{u})$ be $\Delta_0(\exp)$ formulas. Then

$$\mathrm{I}\Delta_0(\exp) \vdash \forall \bar{u} \ \forall a$$
$$\left(\forall x{<}a \ \Big(\bigvee_{d<16} \theta_d(x, \bar{u}) \wedge \bigwedge_{\substack{d,e<16 \\ d \neq e}} \neg\big(\theta_d(x, \bar{u}) \wedge \theta_e(x, \bar{u})\big)\Big) \to \exists s \ \forall x{<}a \ \bigwedge_{d<16} \big(\langle s \rangle_x = \underline{d} \leftrightarrow \theta_d(x, \bar{u})\big)\right).$$

Proof. Fix parameters $\bar{c} \in M \models \mathrm{I}\Delta_0(\exp)$. It is sufficient to show

$$M \models \forall a \left(\begin{array}{l} \forall x{<}a \ \Big(\bigvee_{d<16} \theta_d(x, \bar{c}) \wedge \neg \bigwedge_{\substack{d,e<16 \\ d \neq e}} \big(\theta_d(x, \bar{c}) \wedge \theta_e(x, \bar{c})\big)\Big) \\ \\ \to \exists s{<}2^{4a} \ \forall x{<}a \ \bigwedge_{d<16} \big(\langle s \rangle_x = d \leftrightarrow \theta_d(x, \bar{c})\big) \end{array}\right). \qquad (*)$$

We proceed by $\Delta_0(\exp)$ induction on a in M. The base case when $a = 0$ holds trivially by (RInit).

For the induction step, let $a \in M$ satisfying the subformula between the big brackets in $(*)$. Assume, in addition, that

$$M \models \forall x{<}a+1 \ \Big(\bigvee_{d<16} \theta_d(x, \bar{c}) \wedge \neg \bigwedge_{\substack{d,e<16 \\ d \neq e}} \big(\theta_d(x, \bar{c}) \wedge \theta_e(x, \bar{c})\big)\Big). \qquad (\dagger)$$

Use the induction hypothesis and (S>) to find $s \in M \models s < 2^{4a} \wedge \forall x{<}a \bigwedge_{d<16} \big(\langle s \rangle_x = d \leftrightarrow \theta_d(x, \bar{c})\big)$. Take $d < 16$ such that $M \models \theta_d(a, \bar{c})$. By (†) and (<S), we know such d exists and is unique. Define $s' = 2^{4a}d + s$. Then

$$\begin{aligned} s' &= 2^{4a}d + s < 2^{4a}d + 2^{4a} && \text{by } (+/<), \text{ as } s < 2^{4a}; \\ &= 2^{4a}(d+1) \leqslant 2^{4a} \times 2^4 && \text{by } (\mathrm{R}<) \text{ and } (\times/<), \text{ as } d < 2^4; \\ &= 2^{4a+4} = 2^{4(a+1)} && \text{by } (\exp\times). \end{aligned}$$

Pick $x < a+1$. If $x = a$, then $s' = 2^{4(x+1)} \times 0 + 2^{4x}d + s$, where $s < 2^{4a} = 2^{4x}$, and hence $\langle s' \rangle_x = d$ by the definition of $\langle s' \rangle_x$. So suppose $x \neq a$. Then $x < a$ by (<S). Use the definition of $\langle s \rangle_x$ to find $z, w \in M \models 2^{4(x+1)}z + 2^{4x}\langle s \rangle_x + w \wedge w < 2^{4x}$. Notice $x < a$ implies $x + 1 \leqslant a$ by $(+/<)$ and (<S). So $a - (x+1)$ exists in M. Now

$$s' = 2^{4a}d + s = 2^{4a}d + 2^{4(x+1)}z + 2^{4x}\langle s \rangle_x + w = 2^{4(x+1)}(2^{4(a-(x+1))}d + z) + 2^{4x}\langle s \rangle_x + w.$$

Thus $\langle s' \rangle_x = \langle s \rangle_x$, where $x < a < a + 1$ by (S>). The choice of s then implies $M \models \theta_{\langle s' \rangle_x}(x, \bar{c})$. Notice that no $e \neq \langle s' \rangle_x$ can make $M \models \theta_e(x, \bar{c})$ by (†). □

Intuition. The theory $\mathrm{I}\Delta_0(\exp)$ is able to carry out constructions and reason about notions that

(A) are defined by $\Delta_0(\exp)$ formulas; and

(B) do not involve objects $g(x)$ that grow faster than

$$\exp^{(n)}(x) = 2^{2^{\cdot^{\cdot^{\cdot^{2^x}}}}} \Big\} n\text{-many 2's}$$

for all $n \in \mathbb{N}$.

Roughly speaking, condition (B) provides a bounding term, under which we can apply bounded $\Delta_0(\exp)$ comprehension to the formula given by (A) to obtain what we want.

Building on what we developed, we will verify Löb's derivability conditions in $\mathrm{I}\Delta_0(\exp)$ in the next lecture.