

MODEL THEORY OF ARITHMETIC

Lecture 3: The Weak König Lemma

Tin Lok Wong

22 October, 2014

Within RCA_0 one can prove that [weak König's lemma] is equivalent to each of the following ordinary mathematical statements:

1. The Heine/Borel covering lemma [...].
 2. Every covering of a compact metric space by a sequence of open sets has a finite subcovering [...].
 3. Every continuous real-valued function on $[0, 1]$, or on any compact metric space, is bounded [...].
- [...]
12. Every countable formally real field has a (unique) real closure [...].
 13. Brouwer's fixed point theorem [...].
 14. The separable Hahn/Banach theorem [...].

Stephen Simpson [8, Theorem I.10.3]

3.1 The Ackermann interpretation

We prove the theorems stated at the end of last lecture. They can be paraphrased as saying induction and collection are equivalent to *separation* in set theory.

Theorem 2.7 (Harvey Friedman). For all $n \in \mathbb{N}$ and all $M \models \text{I}\Delta_0 + \text{exp}$, the following are equivalent.

- (a) $M \models \text{I}\Sigma_n$.
- (b) $S \upharpoonright a \in \text{Cod}(M)$ for every $S \in \Sigma_n\text{-Def}(M)$ and every $a \in M$.

Proof. Recall from Lemma 2.6 that nonempty coded sets have least elements. Hence (b) implies $M \models \text{L}\Sigma_n$ as shown in Figure 3.1, so that we get (a) from Theorem 2.3.

Conversely, suppose $M \models \text{I}\Sigma_n$. Take $a \in M$ and $S = \{x \in M : M \models \theta(x)\}$, where $\theta \in \Sigma_n(M)$. By considering $c = 2^a - 1 = \underbrace{11 \cdots 1}_a 2$, we see that

$$M \models \exists c \forall x < a \underbrace{(\theta(x) \rightarrow x \in \text{Ack}(c))}_{\Pi_n}. \quad (*)$$

Let c be a least witness to this, which exists by LII_n from Theorem 2.3. If we can find $x \in \text{Ack}(c)$ such that $M \models \neg\theta(x)$, then $c - 2^x$ would be smaller witness to (*), contradicting the minimality of c . Thus $S \upharpoonright a = \text{Ack}(c)$. \square

In particular, the previous theorem tells us that all bounded Δ_0 -definable sets are coded in models of $\text{I}\Delta_0 + \text{exp}$. In a sense, this provides a base step for the next theorem, and is one of the reasons why $\text{I}\Delta_0$ should be included in the $\text{B}\Sigma_n$'s.

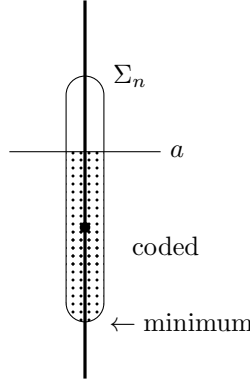


Figure 3.1: Proof that the coding of bounded Σ_n -definable sets implies $L\Sigma_n$

Theorem 2.8 (folklore). For all $n \in \mathbb{N}$ and all $M \models \text{I}\Delta_0 + \text{exp}$, the following are equivalent.

- (a) $M \models \text{B}\Sigma_{n+1}$.
- (b) $S \upharpoonright a \in \text{Cod}(M)$ for every $S \in \Delta_{n+1}\text{-Def}(M)$ and every $a \in M$.

Proof. For (a) \Rightarrow (b), we proceed by strong induction on n . Suppose $M \models \text{B}\Sigma_{n+1}$ and the implication is true for all smaller indices. Let $\varphi \in \Pi_n(M)$ and $\psi \in \Sigma_n(M)$ such that

$$M \models \forall x (\exists y \varphi(x, y) \leftrightarrow \forall y \psi(x, y)). \quad (1)$$

Then $S = \{x \in M : M \models \exists y \varphi(x, y)\} \in \Delta_{n+1}\text{-Def}(M)$. Let $a \in M$. Notice line (1) implies

$$M \models \forall x < a \exists y \underbrace{(\varphi(x, y) \vee \neg\psi(x, y))}_{\Pi_n}.$$

Using $\text{B}\Sigma_{n+1}$, find $b \in M$ such that

$$M \models \forall x < a \exists y < b (\varphi(x, y) \vee \neg\psi(x, y)). \quad (2)$$

Now, if $x < a$, then M satisfies

- $\exists y < b \varphi(x, y) \rightarrow \exists y \varphi(x, y)$;
- $\exists y \varphi(x, y) \rightarrow \forall y \psi(x, y)$ by (1);
- $\forall y \psi(x, y) \rightarrow \forall y < b \psi(x, y)$;
- $\forall y < b \psi(x, y) \rightarrow \exists y < b \varphi(x, y)$ by (2).

Since $M \models \text{B}\Sigma_n$, it follows that

$$S \upharpoonright a = \{x < a : M \models \exists y < b \varphi(x, y)\} = \{x < a : M \models \forall y < b \psi(x, y)\} \in \Delta_n\text{-Def}(M).$$

If $n > 0$, then the induction hypothesis implies $S \upharpoonright a \in \text{Cod}(M)$. If $n = 0$, then the previous theorem gives us the same conclusion.

Conversely, suppose (b) holds. Then as shown in Figure 3.1, we see that $M \models \text{L}\Pi_n$. So M satisfies $\text{I}\Sigma_n$ by Theorem 2.3, and thus also $\text{B}\Sigma_n$ by Theorem 2.2 if $n > 0$. Let $a \in M$ and $\varphi \in \Sigma_{n+1}(M)$ such that $M \models \forall x \leq a \exists y \varphi(x, y)$. Contracting quantifiers, we may assume $\varphi \in \Pi_n(M)$. (This is stated separately as Proposition 3.8 in the Further reading section.) Set $f(x) = (\min y)(\varphi(x, y))$, which exists in M for every $x \leq a$ by $\text{L}\Pi_n$. Consider the formula

$$x \leq a \wedge \forall x' \in [x, a] f(x') \leq f(x),$$

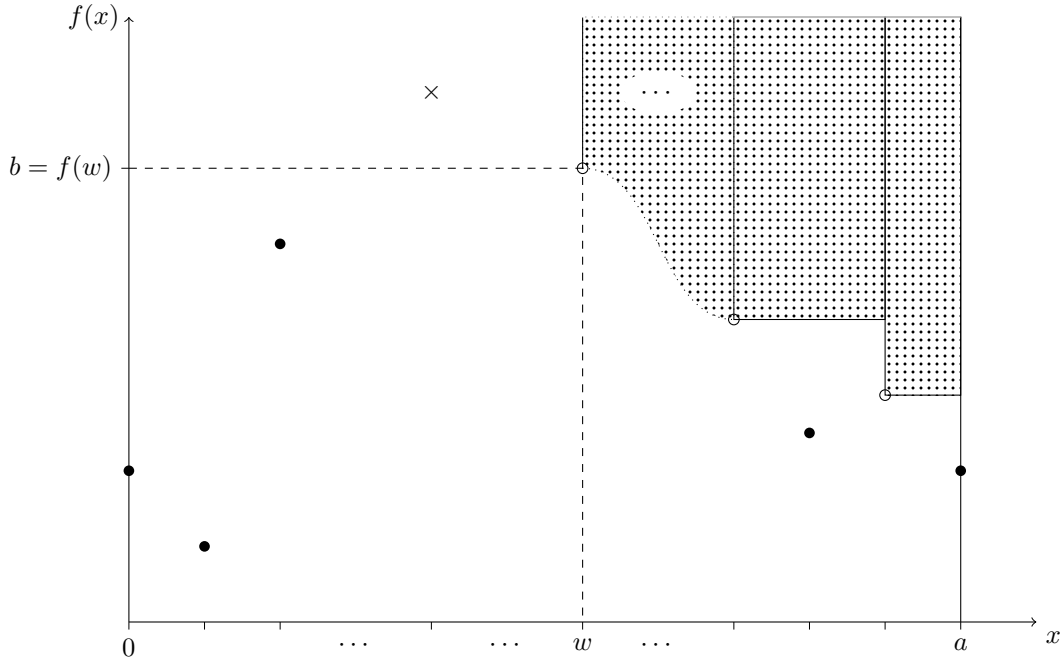


Figure 3.2: Proof that the coding of bounded Δ_{n+1} -definable sets implies $B\Sigma_{n+1}$

which we will denote by $\theta(x)$. This is equivalent in M to both

$$x \leq a \wedge \exists y \underbrace{(\varphi(x, y) \wedge \forall y' < y \neg \varphi(x, y'))}_{\Pi_n} \wedge \underbrace{\forall x' \in [x, a] \exists y' \leq y \varphi(x', y')}_{\Sigma_n \text{ over } B\Sigma_n} \underbrace{\varphi(x', y')}_{\Pi_n}$$

$\underbrace{\hspace{15em}}_{\Sigma_{n+1} \text{ over } B\Sigma_n}$

and

$$x \leq a \wedge \forall y (\varphi(x, y) \wedge \forall y' < y \neg \varphi(x, y') \rightarrow \forall x' \in [x, a] \exists y' \leq y \varphi(x', y')).$$

We do not need $B\Sigma_n$ above when $n = 0$, but in any case, the formula θ is Δ_{n+1} over M . So $S = \{x \leq a : M \models \theta(x)\} \in \text{Cod}(M)$ by (b). Since $M \models \theta(a)$, this set is nonempty and thus has a minimum, say w , by Lemma 2.6. Let $b = f(w)$ as evaluated in M . We claim that $M \models \forall x \leq a \exists y \leq b \varphi(x, y)$. If $x \in [w, a]$, then as $M \models \theta(w)$, we know $f(x) \leq b$ in M . So suppose

$$M \models \exists x < w \forall y \leq b \underbrace{\neg \varphi(x, y)}_{\Sigma_n}$$

$\underbrace{\hspace{10em}}_{\Sigma_n \text{ over } B\Sigma_n}$

By (b), the set of witnesses to this assumption is coded in M . So, being nonempty, it must have a maximum, say x , by Lemma 2.6. Splitting into points before and after w , one can verify that $M \models \theta(x)$. This contradicts the minimality of w . \square

The Ackermann membership makes a model of arithmetic into a model of set theory. These two theorems demonstrated how we can get the separation scheme. We will see in the Further exercises that all axioms of ZFC *except the axiom of infinity* hold in a sufficiently strong model of arithmetic under the Ackermann interpretation. Conversely, every model of set theory in which the axiom of infinity fails interprets arithmetic in a natural way, for example, via the ordinals. These lead to the slogan

$$\text{arithmetic} = \text{finite set theory.}$$

As a consequence, we can import the coding apparatus for finite objects from set theory. This provides one way to code sequences, formulas, proofs, etc. in arithmetic. Although there are usually

more arithmetic ways of coding these objects, it will not matter which coding method we choose as long as certain nicety conditions are satisfied, for example, regarding the complexity of the definition, and the size of codes.

3.2 The Weak König Lemma

Combinatorics come up very naturally in the model theory of arithmetic. The most prominent example is via *cuts*, a concept that goes back to Dedekind's definition of real numbers [2].

Definition. A *cut* of $M \models \text{PA}^-$ is a nonempty initial segment with no maximum. Write $I \subseteq_e M$ for ' I is a cut of M '.

If I is a cut of M , then we may alternatively view M as an *end extension* of I . This explains the subscript e in \subseteq_e . Different papers may have different definitions of cuts, e.g., they may additionally require cuts to be proper, or be closed under addition, multiplication, etc. Unless otherwise stated, we only require cuts to be closed under successor.

Definition. An *exponential cut* is a cut that is closed under $x \mapsto 2^x$.

In general, there may be some ambiguity in saying a cut is closed under a certain definable function f when the cut itself is a model of arithmetic, because f may be interpreted differently in the cut and in the universe. We will see in the next lecture that this ambiguity does not arise if the graph of f is Δ_0 -definable.

Example 3.1. Every $M \models \text{PA}^-$ contains the *standard cut* \mathbb{N} . This is because such M realizes all the closed \mathcal{L}_A -terms $0, 1, 1 + 1, 1 + 1 + 1, \dots$ and the following lemma holds.

Lemma 3.2. $\text{PA}^- \vdash \forall x, y (y > x \rightarrow y \geq x + 1)$.

Proof. Let $y > x$. Then we find z such that $y = x + z + 1$ by axiom (xii). Since $z \geq 0$ by axiom (xv), we conclude $y \geq x + 1$ by axiom (xi). \square

One of the most basic facts about cuts is that proper cuts can never be definable in models with induction.

Proposition 3.3. The following are equivalent for all $M \models \text{PA}^-$ and all $n \in \mathbb{N}$.

- (a) $M \models \text{I}\Sigma_n$.
- (b) No proper cut of M is Σ_n -definable.

Proof sketch. For (a) \Rightarrow (b), recall $\text{I}\Sigma_n$ says every Σ_n -definable set that contains 0 and is closed under successor must contain all numbers. Every cut contains 0 (because it is nonempty and closed downwards) and is closed under successor (because it has no maximum). So if a cut is Σ_n -definable, then it cannot be proper under $\text{I}\Sigma_n$.

For (b) \Rightarrow (a), when given $\theta \in \Sigma_n$, consider $\forall x' \leq x \theta(x')$. \square

Remark 3.4. This proposition remains true if we replace Σ_n by Π_n , with essentially the same proof.

We usually use Proposition 3.3 in the form of *overspill*.

Overspill (A. Robinson [6]). Let $n \in \mathbb{N}$ and $I \subsetneq_e M \models \text{I}\Sigma_n$. If $\theta \in \Sigma_n(M)$ such that

$$M \models \theta(x) \quad \text{for all } x \in I,$$

then $M \models \theta(x)$ for arbitrarily small $x \in M \setminus I$.

Proof. Otherwise $x < b \wedge \theta(x)$ defines I for some $b \in M \setminus I$. \square

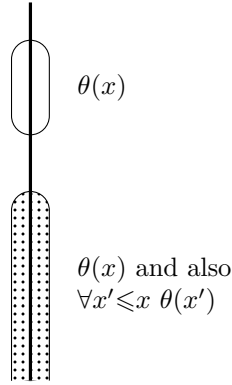


Figure 3.3: The non-definability of cuts implies induction

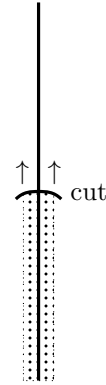


Figure 3.4: Overspill

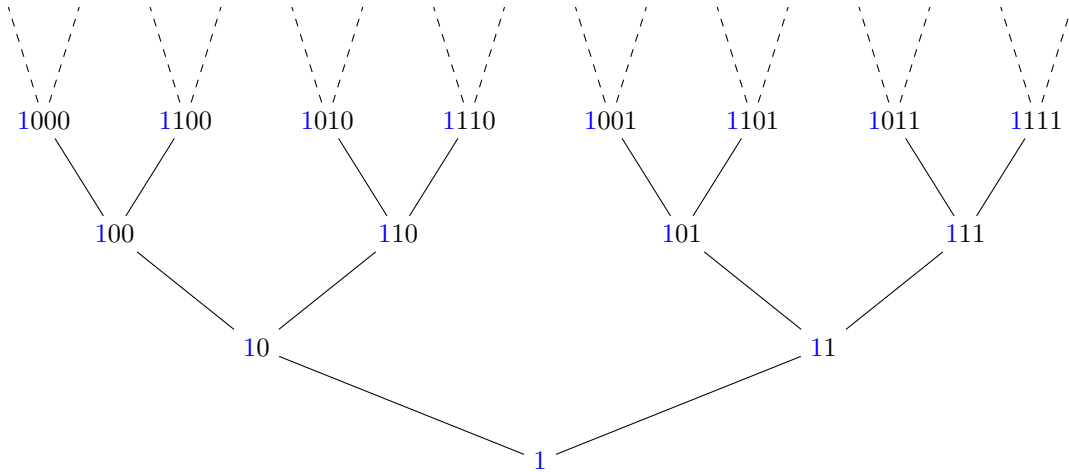


Figure 3.5: Coding of 0–1 sequences

It is customary to use the word *overspill* both as a verb and a noun. Overspill is sometimes called *overflow*. It admits many variants. For example, some versions only have cofinally many x in the cut satisfying θ in the hypothesis [3]. An upside-down version, commonly referred to as *underspill* or *underflow*, infers from the satisfaction of a formula above the cut to satisfaction within the cut. We will see underspill in Lecture 6.

The non-definability of cuts turns out to be less negative than it sounds. Recall from Lemma 2.6 that nonempty coded sets always have maximum elements. In other words, they are bounded or ‘finite’. As a result, we cannot use coded sets *per se* to model unbounded or ‘infinite’ object. One way round this problem is to restrict coded sets to a cut.

Definition. Let $I \subseteq_e M \models \text{I}\Delta_0$. If $c \in M$, then $\text{Ack}(c/I) = I \cap \text{Ack}(c)$. A set of this form is called a *coded subset of I in M* . Set $\text{Cod}(M/I) = \{\text{Ack}(c/I) : c \in M\}$.

In the literature, there are many alternative names for $\text{Cod}(M/I)$, including $\text{SSy}_I(M)$ and $\mathfrak{R}_I(M)$. The acronym SSy comes from *standard system*.

We will use the coded subsets of cuts to formulate the *Weak König Lemma*.

Weak König Lemma. Every infinite 0–1 tree has an infinite branch.

This is ‘weak’ in the sense that the usual König Lemma allows the tree to be finitely branching.

We may use the coding from set theory for trees, but in this particular case, it would be much more convenient to use binary expansions of numbers. Logarithms are always to the base 2.

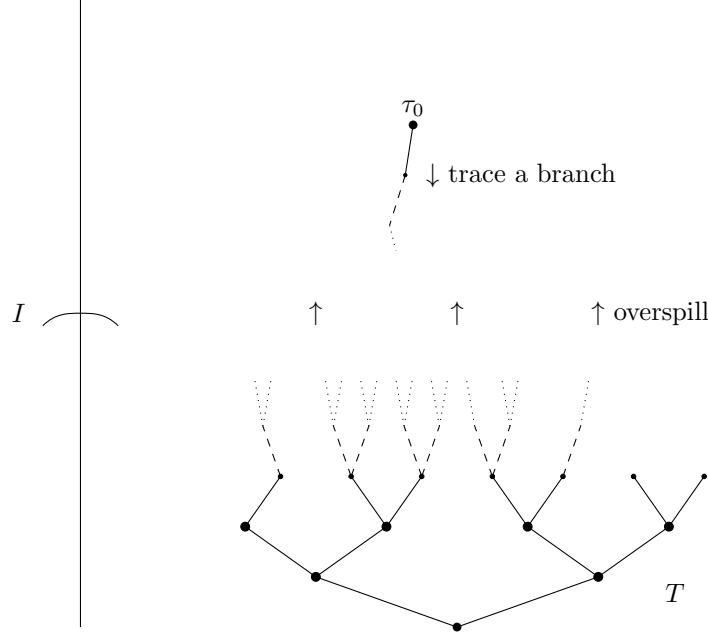


Figure 3.6: Proof that the coded subsets of an exponential cut satisfy the Weak König Lemma

Definition. Let $I \subseteq_e M \models \text{I}\Delta_0$.

- $\text{len } x = \lfloor \log x \rfloor$. In other words $y = \text{len } x$ stands for $2^{y+1} > x \wedge 2^y \leq x$, or more precisely

$$\exists w (w = 2^y \wedge 2w > x \wedge w \leq x).$$

We do not define $\text{len } 0$.

- $\sigma \subseteq_p \tau$ means ‘ σ is an initial part of τ ’, i.e.,

$$\text{len } \sigma \leq \text{len } \tau \wedge \forall i < \text{len } \sigma (i \in \text{Ack}(\sigma) \leftrightarrow i \in \text{Ack}(\tau)).$$

- $\sigma \upharpoonright \ell$ denotes ‘the restriction of σ to ℓ ’. Alternatively, we can define $\sigma \upharpoonright \ell$ by

$$\text{Ack}(\sigma \upharpoonright \ell) = \{\ell\} \cup \{i \in \text{Ack}(\sigma) : i < \ell\}.$$

- $T \subseteq I$ is a *binary tree* if $\sigma \subseteq_p \tau \in T$ implies $\sigma \in T$.
- A *branch* in a binary tree $T \subseteq I$ is a binary tree $B \subseteq T$ such that $\sigma \subseteq_p \tau$ or $\tau \subseteq_p \sigma$ for all $\sigma, \tau \in B$.
- A binary tree is *unbounded* (in I) if $\{\text{len } \tau : \tau \in T\} = I$.

With all these coding, we can now formulate the Weak König Lemma for cuts.

Theorem 3.5 (essentially Scott [7]). Let $M \models \text{I}\Delta_0 + \text{exp}$ and I be a proper exponential cut of M . If T is an unbounded binary tree in $\text{Cod}(M/I)$, then it has an unbounded branch $B \in \text{Cod}(M/I)$.

The idea of the proof is as follows. Since T is unbounded, it has a node at every level $\ell \in I$. By overspill, it must also have a node, say τ_0 , at some level above I . Then $B = \{\tau \in I : \tau \subseteq_p \tau_0\}$ must be an unbounded branch in T . We need the cut to be exponential because the size of a node of length ℓ is exponential in ℓ , so that to have such a node in I , we must also have 2^ℓ in I .

On the contrary, it is not necessary to require $M \models \text{exp}$ in this theorem, but our proof will invoke Theorem 2.7, which needs exponentiation as stated. To eliminate this redundant hypothesis, we simply need to keep our minds clear about which exponentials need to exist while going through the proof of Theorem 2.7. Fact 3.6(c) below will help.

We will see the proof of Theorem 3.5 in detail in the next lecture, together with a converse. Some properties of our coding will be needed then. Let us first list some facts about exponentiation that we will need to establish these properties.

Fact 3.6. $\text{I}\Delta_0$ proves

- (a) $\forall x, y (y = 2^x \rightarrow y > 0 \wedge y > x)$;
- (b) $\forall x, y, x', y' (y = 2^x \wedge y' = 2^{x'} \wedge x < x' \rightarrow y < y')$; and
- (c) $\forall x, y (y = 2^x \rightarrow \forall x' \leq x \exists y' \leq y (y' = 2^{x'}))$.

Proof. All straightforward Δ_0 -inductions. A very useful observation when applying induction is that the string of quantifiers $\forall \bar{x} \dots$ is equivalent to $\forall b \forall \bar{x} < b \dots$ over PA^- . \square

Lemma 3.7. (a) $y = \text{len } x$ is Δ_0 .

- (b) $\text{I}\Delta_0 \vdash \forall x > 0 \exists! \ell (\ell = \text{len } x)$.
- (c) $\text{I}\Delta_0 \vdash \forall x (\text{len } x < x \wedge x < 2^{1+\text{len } x})$.
- (d) $\sigma \subseteq_p \tau$ and $\tau = \sigma \upharpoonright \ell$ are Δ_0 .
- (e) $\text{I}\Delta_0 \vdash \forall \sigma, \ell \exists! \tau (\tau = \sigma \upharpoonright \ell)$.
- (f) $\text{I}\Delta_0 \vdash \forall \sigma, \ell (\ell \leq \text{len } \sigma \rightarrow \text{len}(\sigma \upharpoonright \ell) = \ell \wedge \sigma \upharpoonright \ell \subseteq_p \sigma)$.
- (g) $\text{I}\Delta_0 \vdash \forall \sigma_1, \sigma_2, \sigma_3 (\sigma_1 \subseteq_p \sigma_2 \wedge \sigma_2 \subseteq_p \sigma_3 \rightarrow \sigma_1 \subseteq_p \sigma_3 \wedge \sigma_1 \leq \sigma_3)$.
- (h) $\text{I}\Delta_0 \vdash \forall \tau \forall \sigma_1, \sigma_2 \subseteq_p \tau (\sigma_1 \subseteq_p \sigma_2 \vee \sigma_2 \subseteq_p \sigma_1)$.

Essentially, these are all we need to know about the coding of finite sequences. To put it in another way, our results will not be affected if some other coding method is adopted as long as it satisfies these properties. As a consequence, the actual proof of this lemma, being specific to our coding, is not so important. So I only include an outline of it here.

Proof. (b) Uniqueness is straightforward using Fact 3.6(b). For existence, notice the set $\{y < x : 2^y \leq x\}$ is nonempty since $x \geq 1 = 2^0$. Imitating the proof of Lemma 2.6(b), one finds a maximum of this set, which must be equal to $\text{len } x$.

- (c) This is a reformulation of Lemma 3.7(a).
- (e) Uniqueness follows from extensionality in the Further exercises. For existence, follow the proof of Theorem 2.7, but use Fact 3.6(c) instead of exp to ensure enough exponentials exist.
- (f) As $\ell \in \text{Ack}(\sigma \upharpoonright \ell)$, we know $2^\ell \leq \sigma \upharpoonright \ell$. Notice $\text{Ack}(\sigma \upharpoonright \ell) \subseteq \text{Ack}(2^{\ell+1} - 1)$. So Further exercise (d) implies $\sigma \upharpoonright \ell \leq 2^{\ell+1} - 1 < 2^{\ell+1}$. Putting these together gives $\text{len}(\sigma \upharpoonright \ell) = \ell$. The remaining part is straightforward.
- (g) Proceed directly.
- (h) Without loss, suppose $\sigma_1 < \sigma_2$. Then for all $i < \text{len } \sigma_1$,

$$i \in \sigma_1 \Leftrightarrow i \in \tau \Leftrightarrow i \in \sigma_2. \quad \square$$

Further exercises

Let us look at some set-theoretic axioms in a model $M \models \text{I}\Delta_0$ under the Ackermann interpretation.

- (a) Show that $M \models \exists x \forall i (i \notin \text{Ack}(x))$. This is the axiom of empty set.
- (b) Let $x > 0$ in M . Recall from Lemma 3.7(b) that $\text{len } x$ must exist. Show $\text{len } x \in \text{Ack}(x)$.
- (c) Deduce that $M \models \forall x, y (\forall i (i \notin \text{Ack}(x) \wedge i \notin \text{Ack}(y)) \rightarrow x = y)$. This says there is at most one set that is empty.
- (d) Show by Δ_0 -induction that $M \models \forall x, y (\forall i \in \text{Ack}(x) i \in \text{Ack}(y) \rightarrow x \leq y)$. This implies the axiom of extensionality. Lemma 2.6(a) and the comment in the proof of Fact 3.6 may help.
- (e) Show $M \models \neg \exists x (\exists i (i \in \text{Ack}(x)) \wedge \forall i \in \text{Ack}(x) \exists j \in \text{Ack}(x) i \in \text{Ack}(j))$ using Lemma 2.6. This *refutes* the axiom of infinity.

Further comments

We ignored $B\Sigma_0$ in our analysis of collection schemes because it is actually equivalent to $B\Sigma_1$. This is a special case of a more general observation.

Proposition 3.8. Fix $n \in \mathbb{N}$. Denote by $\text{Coll}(\Pi_n)$ the set of all sentences of the form

$$\forall \bar{z} \forall a (\forall \bar{x} < a \exists \bar{y} \varphi(\bar{x}, \bar{y}, \bar{z}) \rightarrow \exists b \forall \bar{x} < a \exists \bar{y} < b \varphi(\bar{x}, \bar{y}, \bar{z})),$$

where $\varphi \in \Pi_n$. Then $\text{Coll}(\Sigma_{n+1})$ is equivalent to $\text{Coll}(\Pi_n)$.

It is known [5] that $I\Delta_0 \not\equiv B\Sigma_1$, and $B\Sigma_1 \not\equiv \text{exp}$. The following is one of the major open questions in weak arithmetic.

Question 3.9 (Wilkie–Paris [9]). Does $I\Delta_0 + \neg\text{exp} \vdash B\Sigma_1$?

See Adamowicz–Kołodziejczyk–Paris [1] for the most recent status of this question and the connections with complexity theory.

Further reading

We mentioned near the end of Section 3.1 that arithmetic and finite set theory are *bi-interpretable*. If one formulates finite set theory more carefully, then actually the two are *synonymous*. See my paper with Kaye [4] for a precise formulation of this, including the definition of bi-interpretability and synonymy.

References

- [1] Zofia Adamowicz, Leszek Aleksander Kołodziejczyk, and Jeff B. Paris. Truth definitions without exponentiation and the Σ_1 collection scheme. *The Journal of Symbolic Logic*, 77(2):649–655, June 2012.
- [2] Richard Dedekind. *Stetigkeit und irrationale Zahlen*. Verlag von Friedrich Vieweg und Cohn, Braunschweig, 1872.
- [3] Constantinos Dimitracopoulos. Overspill and fragments of arithmetic. *Archive for Mathematical Logic*, 28:173–179, 1989.
- [4] Richard Kaye and Tin Lok Wong. On interpretations of arithmetic and set theory. *Notre Dame Journal of Formal Logic*, 48(4):497–510, October 2007.
- [5] Jeff B. Paris and Laurence A. S. Kirby. Σ_n -collection schemas in arithmetic. In Angus Macintyre, Leszek Pacholski, and Jeff B. Paris, editors, *Logic Colloquium '77*, volume 96 of *Studies in Logic and the Foundations of Mathematics*, pages 199–209. North-Holland Publishing Company, Amsterdam, 1978.
- [6] Abraham Robinson. *Non-standard Analysis*, volume 42 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Company, Amsterdam, 1966.
- [7] Dana Scott. Algebras of sets binumerable in complete extensions of arithmetic. In Jacob C.E. Dekker, editor. *Recursive Function Theory*, volume V of *Proceedings of Symposia in Pure Mathematics*, pages 117–121. American Mathematical Society, Providence, Rhode Island, 1962.
- [8] Stephen G. Simpson. *Subsystems of Second Order Arithmetic*, volume 1 of *Perspectives in Logic*. Cambridge University Press, Cambridge, second edition, 2009.
- [9] Alex J. Wilkie and Jeff B. Paris. On the existence of end extensions of models of bounded induction. In Jens Erik Fenstad, Ivan T. Frolov, and Risto Hilpinen, editors. *Logic, Methodology and Philosophy of Science VIII*, volume 126 of *Studies in Logic and the Foundations of Mathematics*, pages 143–161. North-Holland Publishing Company, Amsterdam, 1989.