A primitive recursive set theory and AFA : on the logical complexity of the largest bisimulation

Tim Fernando*

fernando@cwi.nl

Abstract. A subsystem of Kripke-Platek set theory proof-theoretically equivalent to primitive recursive arithmetic is isolated; Aczel's (relative) consistency argument for the Anti-Foundation Axiom is adapted to a (related) weak setting; and the logical complexity of the largest bisimulation is investigated.

1 Introduction

As every programmer understands, computing depends on coding: only what can be coded can be computed. Conversely, constraints on coding determine the sense in, and degree to which what can be coded can be computed. To a logician, it is natural to express these constraints in a formal system, and relate computation to proofs in that system.[2] Strong systems such as Zermelo-Fraenkel set theory with Choice, ZFC, allow all kinds of mathematical concepts to be coded as sets, however natural or unnatural these formulations might appear. Over a weaker set theory, coding can have greater computational significance, and even a somewhat odd and unlikely use of the notion of a set can sometimes prove fruitful. An instructive example is the set-theoretic coding of (countable) linguistic notions in Kripke-Platek set theory, KP (see Barwise [5]).[3]

More recently, the theory of non-well-founded sets presented in Aczel [4] has been applied in Barwise and Etchemendy [6] for a direct (i.e., natural?) set-theoretic coding of linguistic concepts with a possibly circular character. In this case, however, it is not clear that there is any need for infinite sets, and various alternatives[4] to Aczel's conception of a non-well-founded set have been put forward that suffice for finite sets. By contrast, the consistency proof for the *Anti-Foundation Axiom*, AFA, given in Aczel [4] is carried out relative to the system ZFC^- of ZFC minus foundation that supports a far richer notion of set than that of finite ones. So horrendously rich

* The author is gratefully indebted to Prof. S. Feferman for supervision, to CWI for refugè, and to the Netherlands Organization for the Advancement of Research (project NF 102/62-356, 'Structural and Semantic Parallels in Natural Languages and Programming Languages') for funding.

[2] One way of measuring the computational character of a formal theory is through the *provably recursive functions* given by its Π^0_2-theorems, which is used implicitly below.

[3] In what follows, KP is, as in Barwise [5], *not* assumed to contain the axiom of infinity; otherwise, see Jäger [12].

[4] Mislove, Moss and Oles [16], Abramsky [1], and Rutten [20].

a notion, in fact, that the problem becomes what notion of computation can be associated with these sets. The question of the computational character of AFA is of particular interest given its origins in the (computational) theory of transition systems (see chapter 8 of Aczel [4]). In that work on Milner's SCCS as well as on transition systems given in Plotkin's SOS-style (Rutten [19]), infinite sets are involved. Furthermore, if transition systems are to be related to first-order models (and some such steps are taken in Fernando [10, 11]), then the question of identifying a weak set theory supporting both transition systems and first-order models arises. In any case, *the present author's interest in analyzing* AFA *lies largely in its relation to the notion of a bisimulation — a notion fundamental to semantic attempts at explicating the dynamic nature of information.* For such semantic investigations, it is natural to appeal not only to the ordinary notions of computability and decidability familiar to computer scientists, but also to subtle, set-theoretic notions.[5]

Now, a logical analysis of AFA might proceed in various ways. Lindström [15] formalizes L. Hallnäs' conception of non-well-founded sets in Martin-Löf type theory, building on a constructive version of ZF given in Aczel [3], that, as it turns out, is equivalent to ZF over classical logic. This equivalence blocks a direct understanding in terms of proof-theoretic measures (that at present fall far short of ZF). And from a classical model-theoretic point of view, it would be natural to replace ZFC by a theory, say KP, with many interesting models, and investigate the question mark ? in the diagram

$$\begin{array}{ccc} Con(\mathsf{ZFC}^-) & \Longrightarrow & Con(\mathsf{ZFC}^- + \mathsf{AFA}) \\ \Downarrow & & \Downarrow \\ Con(\mathsf{KP}^-) & \overset{?}{\Longrightarrow} & Con(\mathsf{KP}^- + \mathsf{AFA}) \end{array}$$

where *Con* is a consistency statement formulated in terms of models. It bears repeating that the theory KP^- in the diagram might be enriched, so long as the models of interest are not ruled out. As will become clear below, *the issue here is not the consistency of* AFA, *but its computational requirements*. And these requirements are most clearly exposed in a theory more (directly) sensitive to constructive principles than ZFC^-.

This is not to say that ZFC^- is devoid of any intuitions about construction. The "limitations of size" principle behind the set-class distinction has been so widely accepted and developed that it is perhaps not terribly appropriate to apply the label "set theory" to a theory supporting the existence of a universal set. And there are sound foundational reasons to look at finer questions of size (through a theory of "counting") given that the object $\omega = \{0, 1, \ldots\}$ is *infinitely* more interesting (and complicated) than $\Omega = \{\Omega\}$. A comparison of these two sets suggests that some care must be exercised in pushing the intuition that a non-well-founded set is a limit of well-founded sets, particularly when it leads to a universal set (as is the case in Abramsky [1]).

[5] Having said this, it should be pointed out that ordinary recursion-theoretic questions about processes have been studied; the reader might consult Ponse [18] and the references cited therein.

The approach taken below is to carry out Aczel's relative consistency argument for AFA in a weak setting connected with a view of mathematics that, although called finitist, can nonetheless support infinite objects. The reader is referred to Feferman [9] for background on proof-theoretic and foundational reductions related to consistency arguments. For orientation, it is useful to note that the system PRA of primitive recursive arithmetic is commonly associated with finitism, and (reminiscent of KP's suitability for countable syntactic notions) is adequate for formulating elementary syntactic notions (involved, for example, in Gödel's incompleteness theorems). Briefly then, the next section describes a subsystem KP_1 of KP proof-theoretically equivalent to PRA (building on the correspondence between hereditarily finite sets and natural numbers, the theory of primitive recursive set functions in Jensen and Karp [14], and the reduction in Parsons [17] of Σ^0_1-IA to PRA). (The point here is that quantifier complexity for set theory is related but not identical to that for number theory.) Section 3 carries out Aczel's construction of a model of AFA in a primitive recursive framework provided by explicit mathematics (Feferman [8]) where a model of KP_1 can be defined. Complications arising from the problem of preserving restricted schemes of comprehension and collection motivate the discussion in section 4 of computational "counting" principles for the largest bisimulation.

2 A primitive recursive subsystem of KP

The analysis below rests on the well-known correspondence between the natural numbers ω and the hereditarily finite sets HF given by $c : HF \to \omega$

$$c\emptyset := 0$$
$$ca := \sum_{x \in a} 2^{cx}$$

and $d : \omega \to HF$

$$d0 := \emptyset$$
$$dn := \{di \mid i\text{th bit of } n \text{ is } 1\} .$$

Note that $\in$ (on HF) is a primitive recursive predicate

$$dm \in dn \Leftrightarrow odd([n/2]^m)$$

and accordingly is defined by

$$t_\in[m,n] = 0$$

for some primitive recursive term $t_\in(x,y)$ in the language $\mathcal{L}(\mathsf{PRA})$ of PRA. Now, we can describe an interpretation $-^*$ of $\mathcal{L}(\in)$ in $\mathcal{L}(\mathsf{PRA})$ by passing syntactically from $x \in y$ to $t_\in(x,y) = 0$, and semantically from an $\mathcal{L}(\mathsf{PRA})$-structure $\mathcal{M} = \langle M, \ldots\rangle$ to an $\mathcal{L}(\in)$-structure $\mathcal{M}^* = \langle M, E\rangle$ where

$$E := \{(m,n) \in M \times M \mid \mathcal{M} \models t_\in[m,n] = 0\} .$$

Observe that by the elementary closure properties of primitive recursive predicates, every Δ_0-formula $\varphi(\overline{x})$ in $\mathcal{L}(\in)$ $-^*$-translates to the form (provably equivalent in PRA) of an equation

$$t_\varphi(\overline{x}) = 0 \ .$$

Going the other direction, we have an interpretation $-^\circ$ of $\mathcal{L}(\mathsf{PRA})$ in $\mathcal{L}(\in)$ by the usual identification of natural numbers with finite ordinals. Note that the predicate $\omega(x)$ in $\mathcal{L}(\in)$ is Δ_0. Furthermore, the (numerical) primitive recursive functions can be extracted as restrictions to ω of primitive recursive set functions, to which we now turn.

The *primitive recursive set functions* are given in Jensen and Karp [14] as follows. Close the initial functions

$$\begin{aligned} P_{n,i}(\overline{x}) &= x_i \\ S^2(x_0, x_1) &= x_0 \cup \{x_1\} \\ C(x_0, x_1, x_2, x_3) &= \begin{cases} x_0 \text{ if } x_2 \in x_3 \\ x_1 \text{ otherwise} \end{cases} \end{aligned}$$

under substitution

$$F(\overline{x}) = G(H_1(\overline{x}), \ldots, H_k(\overline{x}))$$

and recursion

$$F(x, \overline{w}) = G(\bigcup_{u \in x} F(u, \overline{w}), x, \overline{w}) \ .$$

The *primitive recursive formulas* are the defining formulas for the set functions above. For example, the defining formula $\Phi(x, \overline{w}, y; \varphi(z, x, \overline{w}, y))$ for a function derived by recursion from a function G with defining formula $\varphi(z, x, \overline{w}, y)$ is

> there is a function h such that $h(x) = y$ and for all u in the domain of h, $u \subseteq$ domain h and
> $$\varphi(\bigcup_{v \in u} hv, u, \overline{w}, hu) \ .$$

Let PRS be the (classical) first-order theory in the language of set theory consisting of the axioms of extensionality, pairing, union, Δ_0-separation, induction on primitive recursive formulas φ

$$\forall z(\forall v \in z \varphi(v) \supset \varphi(z)) \supset \forall z \varphi(z)$$

and the *Σ_1-recursion rule*

$$\frac{\forall z, x, \overline{w}\ \exists! y\ \varphi(z, x, \overline{w}, y)}{\forall x, \overline{w}\ \exists! y\ \Phi(x, \overline{w}, y; \varphi(z, x, \overline{w}, y))}$$

where $\Phi(x, \overline{w}, y; \varphi(z, x, \overline{w}, y))$ is the defining formula for the function derived by recursion from a Σ_1-formula $\varphi(z, x, \overline{w}, y)$. Transitive models of PRS are *prim-closed* in the sense of Jensen and Karp [14].

Under suitable arithmetization, the collections of proofs in PRA and PRS are primitive recursive. Furthermore, a primitive recursive function can be constructed mapping (provably in PRA) axioms φ of PRS to PRA-proofs of φ^*. Consequently,

Proposition 1. (PRA) PRS $\vdash \varphi$ *implies* PRA $\vdash \varphi^*$.

The converse of Proposition 1 fails because the sets that PRA $-^*$-induces are "finite." (For a counter-example, take the $\mathcal{L}(\in)$-sentence that asserts that every non-empty a set has an $\in$-maximal element

$$\exists x \in a\ x \in a \supset \exists z \in a\ \forall x \in a\ z \notin x\ ;$$

its $-^*$-translation is a theorem of PRA.) We can, however, approximate a converse. As e° is a primitive recursive formula for every $\mathcal{L}$(PRA)-equation e, another inductive argument on the the length of a proof yields

Proposition 2. (PRS) PRA $\vdash \psi$ *implies* PRS $\vdash \psi^\circ$.

Furthermore, every model $\mathcal{M}$ of PRA can be embedded in a model of PRS, namely $\mathcal{M}^*$ via $\pi : \mathcal{M} \cong \mathcal{M}^{*\circ}$

$$\pi 0 := 0$$
$$\pi(n+1) := \sum_{m \leq n} 2^{\pi m}\ ,$$

whence

Proposition 3. PRS *is a conservative extension of* PRA.

Mention of primitive recursive formulas can be avoided altogether by asserting the principle of induction for all Σ_1-formulas. Set PRS$'$ to PRS with primitive recursive induction promoted to Σ_1-induction (Σ_1IA). Now, the $-^*$-translated content of the Σ_1-recursion rule does not change since Parsons [17] proved (in PRA) that if

$$\Sigma_1^0\text{-IA} \vdash \forall n \exists m R(n, m)$$

where R is primitive recursive, then

$$\text{PRA} \vdash R(n, fn)$$

for some primitive recursive function f. The arguments for PRS and PRA adapt readily to yield

Proposition 4. *1.* (PRA) PRS$'$ $\vdash \varphi$ *implies* Σ_1^0-IA $\vdash \varphi^*$.
2. (PRS) Σ_1^0-IA $\vdash \psi$ *implies* PRS$'$ $\vdash \psi^\circ$.
3. PRS$'$ *is a conservative extension of* Σ_1^0-IA.

As with Proposition 1, the converse to part 1 of Proposition 4 fails, which leads us to formulate

Lemma 5. *Let Φ be a primitive recursive collection of $\mathcal{L}(\in)$-formulas for which there is, provably in Σ_1^0-IA, a primitive recursive function f such that for every $\varphi \in \Phi$, $f\varphi$ is a Σ_1^0-IA-proof of φ^*. Then* PRS′ $+ \Phi$ *is proof-theoretically equivalent to* PRA.

Sieg [21] contains a wealth of information concerning Σ_1^0-IA, including "easy and helpful facts" (his words) such as

(a) Π_1^0-IA is equivalent to Σ_1^0-IA (p. 46), and
(b) Σ_1^0-collection[6] is contained in Σ_1^0-IA (p. 53).

Concerning point (b), it is interesting to note that PRS′ is a subsystem of the predicative set theories in Feferman [7], and hence does not imply Δ_0-collection:

$$\forall x \in a \exists y\ \varphi(x,y) \supset \exists z \forall x \in a \exists y \in z\ \varphi(x,y)$$

for Δ_0-formulas $\varphi(x,y)$. (It has the same transitive models as PRS, including sets that are *not* admissible.) Nevertheless, Φ in Lemma 5 can be taken to be Δ_0-collection, by adapting Sieg's argument for (b).[7] It is well-known that in the presence of Δ_0-collection, the distinction between Σ_1- and Σ-formulas (also called generalized or essentially Σ_1-formulas) evaporates. As for point (a), this allows us to conclude that, defining the subsystem KP_1 of KP as $\mathsf{KP}^- + (\Sigma_1 + \Pi_1)$IA (where KP^- is KP minus foundation)[8],

Theorem 6. KP_1 *is proof-theoretically equivalent to* PRA.

3 Aczel's AFA construction in a weak setting

To shed light on the infinitary demands of AFA, it is natural (as argued in section 1) to carry out Aczel [4]'s (relative) consistency argument for the axiom in a weaker

[6] These are arithmetic principles

$$\forall x < a\ \exists y\ \varphi(x,y) \supset \exists z \forall x < a\ \exists y < z\ \varphi(x,y)\ ,$$

where $\varphi(x,y)$ is Σ_1^0.

[7] Assume (in Σ_1^0-IA) that

$$(\forall x \in a \exists y\ \varphi(x,y))^*$$

where φ is Δ_0. Now, calling the formula

$$b \leq a \supset \exists z \forall x < b \exists y < z\ t_\in(x,b) = 0\ \supset\ (t_\in(y,z) = 0 \wedge \varphi^*(x,y))$$

$\psi(b)$, then as $\psi(0)$ and $\psi(b) \supset \psi(b+1)$, it follows by Σ_1^0-IA that (because $t_\in[m,n] = 0$ implies $m < n$)

$$(\exists z \forall x \in a \exists y \in z\ \varphi(x,y))^*\ .$$

[8] The Σ_1-recursion rule is a consequence (relative to KP^-) of $(\Sigma_1 + \Pi_1)$IA. If the existence of the transitive closure of a set is added to KP^- (as in work by Jäger), then Π_1IA is not necessary to justify the rule, although Π_1IA is useful for purposes other than proving the existence of transitive closures (see Barwise [5]), an example of which is given in section 4 below. The author does not see how to derive Π_1IA from Σ_1IA (in particular, how to adapt the argument in Sieg [21] reducing Π_1^0-IA to Σ_1^0-IA).

setting than ZFC^-. Accordingly, over a model $\langle S, \approx, \in \rangle$ of KP^- (i.e., KP minus foundation), define the following.

- A *graph* G is a pair $(N_G, \rightarrow_G)$ with $\rightarrow_G \subseteq N_G \times N_G$.
- A *decoration of a graph* G is a function d on N_G such that $da \approx \{db \mid a \rightarrow_G b\}$.
- The *Anti-Foundation Axiom*, AFA, is the assertion that every graph has a unique decoration.
- A *pointed graph (pg)* is a pair (G, a) consisting of a graph G and a set $a \in N_G$.[9]
- A *bisimulation between graphs G and G'* is a set R such that whenever bRb',

$$\forall x \leftarrow_G b \; \exists y \leftarrow_{G'} b' \; xRy \wedge \forall y \leftarrow_{G'} b' \; \exists x \leftarrow_G b \; xRy \; .$$

- Let

$$Bis(R, G, a, G', a') \Leftrightarrow \text{“}R \text{ is a bisimulation between } G \text{ and } G' \text{ such that } aRa'\text{”} \; ,$$

and let S_0 be the collection of all pg's, and $\approx_0$ be the subcollection of $S_0 \times S_0$ given by

$$(G, a) \approx_0 (G', a') \Leftrightarrow \exists R \; Bis(R, G, a, G', a') \; .$$

The preceding definitions all refer to sets (i.e., objects in S), except for the collections Bis, S_0 and $\approx_0$. These collections will serve as useful abbreviations, but where do they live? Rather than working in a framework where "limitations of size" lead, for example, to complications with quotients[10], it is possible instead to work in the framework of explicit mathematics (Feferman [8]), where

(1) theories of weak proof-theoretic strength can be formulated naturally, and
(2) the problem of quotients can be sidestepped by adopting Bishop's use of "equality" relations.

Concerning point (1), observe that a model of KP_1 can be defined (by numerically coding the hereditarily finite sets) in the theory $\mathsf{APP} + \mathsf{ECA} + \mathsf{Obj\text{-}ind}_N$ described in Jäger [13] (where it is stated, furthermore, to be proof-theoretically equivalent to PRA). As for point (2), this was anticipated above in isolating the interpretation $\approx$ of equality on S. To go along with $\approx_0$, define the subclass $\in_0$ of $S_0 \times S_0$ as follows

$$(G, a) \in_0 (G', a') \equiv \exists b \leftarrow_{G'} a' \;\; (G', b) \approx_0 (G, a) \; .$$

Theorem 7. [11] *If S, $\approx$ and $\in$ are* $(\mathsf{APP} + \mathsf{ECA})$*-classes such that*

$$\langle S, \approx, \in \rangle \models \mathsf{KP}^-$$

[9] The notion of an *accessible pointed graph* is avoided at some aesthetic cost to push through the argument in KP^-.

[10] Lemma 2.17 of Aczel [4] employs Scott's trick plus some choice principle to form a *strongly extensional quotient*.

[11] The author suspects (but has not had the energy to check all the necessary details) that the proof below can be formalized in APP+ECA, plus, if necessary, $\mathsf{Obj\text{-}ind}_N$. (Hence, the use of explicit mathematics.) But as it stands, the reduction claimed is model-theoretic, not proof-theoretic.

then (APP + ECA)*-classes* S_0, $\approx_0$ *and* $\in_0$ *can be defined (as above) such that*

$$\langle S_0, \approx_0, \in_0 \rangle \models \text{Extensionality} + \text{Pair} + \text{Union} + \text{AFA} .$$

Furthermore, the passage from $\langle S, \approx, \in \rangle$ *to* $\langle S_0, \approx_0, \in_0 \rangle$ *preserves satisfaction of (full)* Separation, *(full)* Collection, Infinity, Power, *and* Choice.

Proof. First, observe that ECA supports the (class) definitions of $S_0, \approx_0$, and $\in_0$ from $S, \approx$, and $\in$ above since the only terms that occur *qua* class in the defining formulas are $S, \approx$, and $\in$. Second, to see that $\approx_0$ is an equivalence relation is routine (assuming KP^-): clearly, $\approx_0$ is reflexive (since for every pg (G, a), the restriction of $\approx$ to N_G is a bisimulation on G), symmetric (since if R is a bisimulation between G and G', then R^{-1} is a bisimulation between G' and G), and transitive (since if R is a bisimulation between G and G', and R' is a bisimulation between G' and G'', then $R \circ R'$ is a bisimulation between G and G''). Third, although a quotient need not be formed, it is necessary to prove that $\in_0$ respects $\approx_0$. So suppose $(G, a), (G', a'), (G_1, a_1)$ and $(G'_1, a'_1) \in S_0$ satisfy

$$(G_1, a_1) \approx_0 (G, a) \in_0 (G', a') \approx_0 (G'_1, a'_1) ,$$

with the object of showing

$$(G_1, a_1) \in_0 (G'_1, a'_1) .$$

Since $(G', a') \approx_0 (G'_1, a'_1)$, there is a bisimulation R between G' and G'_1 relating a' to a'_1, and since $(G, a) \in_0 (G', a')$, there is a $b \leftarrow_{G'} a'$ with $(G', b) \approx_0 (G, a)$. Choose a $b_1 \in N_{G'_1}$ such that bRb_1 and $a'_1 \rightarrow_{G'_1} b_1$. Then $(G', b) \approx_0 (G'_1, b_1)$ (via R), whence $(G'_1, b_1) \approx_0 (G_1, a_1)$ (as required) since $(G', b) \approx_0 (G, a) \approx_0 (G_1, a_1)$.

Next, for the axioms, it is helpful to define the (class) map $[-]^0$ from S_0 to S given by

$$[(G, a)]^0 = \{(G, a') \mid \; a \rightarrow_G a'\} ,$$

and associate with every formula φ in the language $\{=, \dot{\in}\}$ of set theory the predicate $[\varphi]_0$ obtained by interpreting the quantifiers over S_0, the equality symbol $=$ by $\approx_0$, and the membership symbol $\dot{\in}$ by $\in_0$. A crucial property of the interpretation $\varphi \mapsto [\varphi]_0$ is that for every formula φ, and quantifier $Q \in \{\forall, \exists\}$,

$$[Qx\dot{\in}y \; \varphi]_0 \equiv Qx \in [y]^0 \; [\varphi]_0 .$$

This is a consequence of three facts: (1) $\approx_0$ is an equality for $\langle S_0, \approx_0, \in_0 \rangle$, (2) for every pg (G, a), and every $x \in_0 (G, a)$, there is a $y \in [(G, a)]^0$ where $x \approx_0 y$, and (3) for every pg (G, a), every $x \in [(G, a)]^0$ is $\in_0 (G, a)$.

Now, to establish the analog to Corollary 3.3 and Proposition 3.7 in Aczel [4] (implying the system V_c constructed there is *full*), define the predicate

$$Copy(G, a, A) \Leftrightarrow \forall (G', a') \in S_0 \quad (G', a') \in_0 (G, a) \; \equiv \; \exists (G'_1, a'_1) \in A \;\; (G'_1, a'_1) \approx_0 (G', a')$$

and assert

Lemma 8. (KP$^-$) *For every set (i.e., member of the class S) $A \subseteq S_0$, there is a $(G, a) \in S_0$ unique up to $\approx_0$ such that $Copy(G, a, A)$.*

Proof. Given such an A, pick an $a \notin \bigcup\{N_{G'} \mid (G', a') \in A\}$ (justified by an argument by contradiction using Δ_0-separation and the Russell set) and define (over $\langle S, \approx, \in\rangle$)

$$N_G := \{a\} \cup \bigcup\{N_{G'} \mid (G', a') \in A\}$$
$$\rightarrow_G := \{(a, a') \mid (G', a') \in A\} \cup \bigcup\{\rightarrow_{G'} \mid (G', a') \in A\} .$$

Then $Copy(G, a, A)$ holds. Moreover, to prove uniqueness up to $\approx_0$, suppose $Copy(G_1, a_1, A)$ were true. Constructing a bisimulation between G and G_1 relating a to a_1 appears simple enough — take

$$\{(a, a_1)\} \cup \{(b, b_1) \in N_G \times N_{G_1} \mid \exists (G', a') \in A \;\; \psi(G, b, G', a') \wedge \psi(G_1, b_1, G', a')\}$$

where $\psi(X, y, U, v)$ is $\exists R \; Bis(R, X, y, U, v)$. But the existential quantifier in ψ must be either bounded using Power, or else the definition cannot be justified by KP's limited separation principles. Fortunately, a more delicate argument is possible. Given an R satisfying

$$\forall b \leftarrow_G a \; \exists (G', a') \in A \;\; Bis(R, G, b, G', a') \wedge$$
$$\forall (G', a') \in A \; \exists b \leftarrow_G a \;\; Bis(R, G, b, G', a')$$

and an R_1 satisfying

$$\forall (G', a') \in A \; \exists b_1 \leftarrow_{G_1} a_1 \;\; Bis(R_1, G', a', G_1, b_1) \wedge$$
$$\forall b_1 \leftarrow_{G_1} a_1 \; \exists (G', a') \in A \;\; Bis(R_1, G', a', G_1, b_1) \; ,$$

compose R with R_1 and throw in (a, a_1) to form a bisimulation between G and G_1 relating a to a_1. It remains to show how to construct the required R and R_1. Observe that R can be obtained by applying Σ-collection (as given in Barwise [5], Theorem 4.4, p. 17) first to

$$\forall b \leftarrow_G a \; \exists R' \; \exists (G', a') \in A \;\; Bis(R', G, b, G', a') \; ,$$

then to

$$\forall (G', a') \in A \; \exists R' \; \exists b \leftarrow_G a \;\; Bis(R', G, b, G', a')$$

(which hold since $Copy(G, a, A)$), and forming the union. Constructing R_1 is similar. ⊣

Preservation of the axioms is proved à la Rieger's theorem (Aczel [4]) by applying Lemma 8 to a suitable A for the existence (or in the case of Extensionality, the uniqueness up to $\approx_0$) of a required $(G, a) \in S_0$.

Extensionality: given (G_0, a_0) and (G_1, a_1) in S_0 such that

$$\forall (G', a') \in S_0 \;\; (G', a') \in_0 (G_0, a_0) \;\equiv\; (G', a') \in_0 (G_1, a_1) \; ,$$

then $Copy(G_0, a_0, [(G,a)]^0)$ and $Copy(G_1, a_1, [(G,a)]^0)$, whence Lemma 8 yields $(G_0, a_0) \approx_0 (G_1, a_1)$.

Pair: given $(G,a), (G',a') \in S_0$, appeal to Lemma 8 with $A = \{(G,a),(G',a')\}$.

Union: given $(G,a) \in S_0$, appeal to Lemma 8 with $A = \bigcup\{[(G,a')]^0 \mid a \to_G a'\}$.

Separation: given a formula $\varphi(x)$ and $(G,a) \in S_0$, let A be the set

$$\{(G,a') \in [(G,a)]^0 \mid [\varphi((G,a'))]_0\}$$

(which exists, assuming $\langle S, \approx, \in \rangle$ satisfies Separation).

Collection: let $\varphi(x,y)$ be a formula and $(G,a) \in S_0$ such that

$$\forall x \in_0 (G,a) \; \exists y \in S_0 \; [\varphi(x,y)]_0 \; .$$

Assuming $\langle S, \approx, \in \rangle$ satisfies Collection, there is a set B such that

$$\forall x \in [(G,a)]^0 \; \exists y \in B \; [\varphi(x,y)]_0 \; .$$

Appeal to Lemma 8 with $A = B \cap S_0$.

AFA: Let g be a pg such that ["g is a graph"]$_0$. A first attempt would be to apply Lemma 8 to $\{B \in S_0 \mid [\psi]_0\}$ where ψ is

$$\exists x \dot{\in} N_g \;\; B = (x, \{z \mid x \to_g z\}) \; .$$

Unfortunately, this class is *not* a set, the problem being that given a pg (G,a), a proper class of pg's (G',a') may satisfy $(G',a') \in_0 (G,a)$. (Similarly, with $\approx_0$.) This suggests a different interpretation of the language of set theory, obtained by interpreting the quantifiers over S_0, $=$ by $\approx$, and $\dot{\in}$ by $\in [-]^0$. Write $[\varphi]^0$ for the result of applying this interpretation to some formula φ.[12] Now, if d is the result of applying Lemma 8 to $\{B \in S_0 \mid [\psi]^0\}$, then it follows that for every $x \in_0 N_g$,

$$\forall y \in_0 d(x) \; [x \to_g y]_0 \wedge \forall z \; [x \to_g z]_0 \supset z \in_0 d(x),$$

whence (by Extensionality),

$$\begin{aligned} x &\approx_0 d(x) \\ &\approx_0 \{z \mid x \to_g z\} \\ &\approx_0 \{d(z) \mid x \to_g z\} \; , \end{aligned}$$

[12] Ideally,

$$[\varphi(\overline{x})]_0 \equiv \exists \overline{y} \approx_0 \overline{x} \; [\varphi(\overline{y})]^0$$

would always hold; however, counter-examples such as $\neg \; x \dot{\in} x'$ are easy to find. Counter-examples not involving negation are more difficult to produce, and the author suspects that the above equivalence might hold for (Σ-formulas) φ constructed "non-negatively" (allowing for bounded quantification, which might be justified by Σ-collection). If so, then separation and collection over non-negative Δ_0-formulas might be true in $\langle S_0, \approx_0, \in_0 \rangle$.

that is, ["d is a decoration of g"]$_0$. Furthermore, if $d' \in S_0$ satisfies ["d' is a decoration of g"]$_0$, then $d \approx_0 d'$, since for every $x \in_0 N_g$, there is a bisimulation between $d(x)$ and $d'(x)$ (whence $d(x) \approx_0 d'(x)$).

Infinity: given an $a \in S$ that makes Infinity true (i.e., "$a = \omega$") in $\langle S, \approx, \in \rangle$, apply Lemma 8 to the set A obtained by applying Σ-collection to $\forall n \in a\ \exists b\ P(n, b)$ where

$$P(n,b) \Leftrightarrow \exists f\ function(f)\ \wedge\ dom(f) \approx n\ \wedge\ Q(f,b,n)\ \wedge\ \forall k \in n\ Q(f, f(k), k)$$

and

$$Q(f,x,y) \Leftrightarrow (\forall c \in [x]^0\ \exists z \in y\ f(z) \approx c)\quad \wedge\quad (\forall z \in y\ f(z) \in [x]^0)\ .$$

Power: given $(G, a) \in S_0$, apply Lemma 8 to a set A such that

$$\forall x \in Pow([(G,a)]^0)\ \exists (G', a') \in A\quad Copy(G', a', x)$$

(again obtained by Σ-collection, noting that $Copy$ can be put in Σ-form[13]).

Choice: choice functions can be produced in S_0 (assuming such exist in S) as in the proof of Rieger's theorem in Aczel [4].
⊣

The problem with preserving Δ_0-separation and Δ_0-collection is the unbounded existential quantifier in the definition of $\approx_0$ (whence the passage from φ to $[\varphi]_0$ does not preserve Δ-formulas). Since a bisimulation R between G and G' relating a to a' where $a \in N_G$ and $a' \in N_{G'}$ can be taken to be a subset of $N_G \times N_{G'}$, the problem is overcome by assuming Power.[14]

Corollary 9. *If S, $\approx$ and $\in$ are* (APP + ECA)*-classes such that*

$$\langle S, \approx, \in \rangle \models \mathsf{KP}^- + \mathsf{Power}$$

then the classes S_0, $\approx_0$ and $\in_0$ in Theorem 7 form a model of KP^- + Power + AFA.

As already mentioned, the theory

$$\mathsf{T} := \mathsf{APP} + \mathsf{ECA} + \mathsf{Obj\text{-}ind}_N$$

of explicit mathematics is proof-theoretically equivalent to PRA, according to Jäger [13]. Within T, classes $S, \approx, \in$ can be defined (as in the previous section) such that

$$\langle S, \approx, \in \rangle \models \mathsf{KP}_1 + \mathsf{Power}$$

[13] That is, $Copy(G, a, x)$ can be re-expressed as

$$\forall (G', a') \in [(G,a)]^0\ \exists (G'_1, a'_1) \in A\ \ (G', a') \approx_0 (G'_1, a'_1)\quad \wedge$$
$$\forall (G', a') \in A\ (G', a') \in_0 (G, a)\ .$$

[14] As Prof. Barwise has pointed out to the author, the notation KP^- + Power might be interpreted as requiring that the powerset predicate is taken to be Δ_0. This is not necessary, because all possible bisimulations referred to in $[-]_0$-translating a Δ_0-formula can be found in a set constructed from (the interpretations of) its free variables (since all quantifiers are bounded).

(where, by Lemma 5, $\mathsf{KP}_1 + \mathsf{Power} \equiv \mathsf{PRA}$). Combined with Corollary 9, this provides an illustration of the finitary character of AFA. From the point of view of admissible sets (and more generally, computational theories based on enumeration), however, Power is an undesirable axiom that is best avoided (or at least weakened to Power or Infinity), and its use for capturing the largest bisimulation suggests a certain impredicativity about AFA somewhat at odds with the claim that the axiom is finitistic. The question arises as to whether the existence of a bisimulation can be expressed by a Δ-predicate over a finitist (i.e., at most primitive recursive) subsystem of KP. A natural attempt at answering this question affirmatively would involve some principle of induction.[15]

4 Induction and the largest bisimulation

In practice, a good deal of the proof-theoretic strength of a set theory lies in its induction principles. Over a universe of possibly non-well-founded sets, however, such principles must be formulated carefully (given the difficulty in applying these globally). The approach taken below is to introduce a unary relation symbol *Ord* plus suitable axioms, and to relativize the induction principles to *Ord*. More precisely, let KP^{Ord} be KP plus

$$\begin{array}{rcl} & Ord(\emptyset) & \\ Ord(x) & \supset & Ord(x \cup \{x\}) \\ \forall y \in x\ Ord(y) & \supset & Ord(\bigcup x) \\ Ord(x) & \supset & \forall y \in x\ (Ord(y)\ \wedge\ \forall z \in y\ z \in x)\ . \end{array}$$

(Note that this is a conservative extension of KP since, under foundation, $Ord(x)$ can be given a Δ_0-definition, as in Barwise [5].) Now, let KP_1^{Ord} be the result of replacing foundation in KP^{Ord} by $(\Sigma_1 + \Pi_1)\mathsf{IA}$ relativized to *Ord*:

$$\forall^{Ord}\alpha(\forall\beta \in \alpha\varphi(\beta) \supset \varphi(\alpha)) \supset \forall^{Ord}\alpha\ \varphi(\alpha)$$

for Σ_1- and Π_1-formulas φ (where $\forall^{Ord}\alpha\ \chi$ is $\forall\alpha\ Ord(\alpha) \supset \chi$). KP_1^{Ord} supports the $(\Sigma_1\textit{-recursion rule})^{Ord}$

$$\frac{\forall z, x, \overline{w}\ \exists! y\ \psi(z, x, \overline{w}, y)}{\forall^{Ord}\alpha\ \forall\overline{w}\ \exists! y\ \Phi(\alpha, \overline{w}, y; \psi(z, \alpha, \overline{w}, y))}$$

where $\Phi(x, \overline{w}, y; \psi(z, x, \overline{w}, y))$ is the defining formula for the function derived by recursion from a Σ_1-formula $\psi(z, x, \overline{w}, y)$ (see section 2). This rule provides a constructive (i.e., ordinal) approach to the least and greatest fixed points of certain inductive definitions.[16]

[15] A negative answer for a particular subsystem T might proceed by choosing an appropriate model of T, with respect to which a Δ-definition of the largest bisimulation would lead to a contradiction. The interest in such a result would depend largely on the interest in the model used. The author has tried and failed to push through such an argument for $L(\omega_1^{CK})$.

[16] The reader is referred to Aczel [2] for far more background than is presently required.

Assume $\varphi(R,x)$ is a Δ predicate in which R occurs positively and for which there are Δ-predicates $\varphi_I(\alpha,x)$ and $\varphi_J(\alpha,x)$ satisfying for all α such that $Ord(\alpha)$,

$$\varphi_I(\alpha,x) \equiv \varphi(\exists\beta \in \alpha\ \varphi_I(\beta,-),x)$$

and for $\alpha > 0$,

$$\varphi_J(\alpha,x) \equiv \varphi(\forall\beta \in \alpha\ \varphi_J(\beta,-),x)\ .$$

Implicit here is the assumption that $\exists\beta \in \alpha\ \varphi_I(\beta,-)$ defines a set (i.e., $\{a \mid \exists\beta \in \alpha\ \varphi_I(\beta,a)\}$), as does, for $\alpha > 0$, $\forall\beta \in \alpha\ \varphi_J(\beta,-)$. The idea is to characterize the least fixed point I_φ and the greatest fixed point J_φ of $\varphi(R,x)$ by

$$I_\varphi(x) \equiv \exists^{Ord}\alpha\ \varphi_I(\alpha,x) \tag{1}$$
$$J_\varphi(x) \equiv \forall^{Ord}\alpha\ \varphi_J(\alpha,x)\ . \tag{2}$$

Typically, $\Leftarrow$ of (1) and $\Rightarrow$ of (2) are justified by induction principles, while the converses rest on a cardinality argument — i.e., a well-ordering principle that enables induction to be enforced globally (on all sets). But, how are the left hand sides defined in the first place? If we agree that

$$J_\varphi(x) \Leftrightarrow \exists R\ (\forall y \in R\ \varphi(R,y)\ \wedge\ x \in R)\ ,$$

then (2) makes $J_\varphi(x)$ Δ. Note that stipulating

$$I_\varphi(x) \Leftrightarrow \forall R\ (\forall y\ (\varphi(R,y) \supset y \in R)\ \supset\ x \in R)$$

not only fails to lower the complexity of $I_\varphi(x)$ given by (1), but is, in fact, incorrect, since the variable R must range over classes (including proper ones for choices of $\varphi(R,x)$ such as $x = x$). As for the corresponding definition above of $J_\varphi(x)$, the point is that the argument R in $\varphi(R,y)$ must be a set, which in the terminology of Aczel [4] induces a "set continuous" class operator

$$R \mapsto \{a \mid \exists r (\in V)\ r \subseteq R\ \wedge\ \varphi(r,a)\}\ .$$

We now apply these ideas to a concrete case.

Towards an inductive construction of a bisimulation between pointed graphs $(N,\to),a$ and $(N',\to'),a'$, define

$$\varphi(R,a,a';N,\to,N',\to') \Leftrightarrow a \in N\ \wedge\ a' \in N'\ \wedge$$
$$\forall x \leftarrow a\ \exists y \leftarrow' a'\ xRy\ \wedge\ \forall y \leftarrow' a'\ \exists x \leftarrow a\ xRy\ .$$

Lemma 10. *Over* KP_1^{Ord}, *a* Δ *predicate* $\varphi_J(\alpha,a,a';N,\to,N',\to')$ *can be constructed such that*

$$\varphi_J(0,a,a';N,\to,N',\to') \equiv a \in N\ \wedge\ a' \in N'\ ,$$

and for all $\alpha \neq 0$ *such that* $Ord(\alpha)$,

$$\varphi_J(\alpha,a,a';N,\to,N',\to') \equiv \varphi(\forall\beta \in \alpha\ \varphi_J(\beta,-,-';N,\to,N',\to'),a,a';N,\to,N',\to')\ .$$

Consider next what principles are needed to establish the following assertion: for all pointed graphs $(N, \to), a$ and $(N', \to'), a'$,

$$\exists R\ Bis(R, (N, \to), a, (N', \to'), a') \equiv \forall^{Ord}\alpha\ \varphi_J(\alpha, a, a'; N, \to, N', \to')\ , \qquad (3)$$

where recall from section 2 that $Bis(R, G, a, G', a')$ says that R is a bisimulation between G and G' relating a to a'. ($\Rightarrow$) follows from Π_1IA, even if relativized to Ord, applied to the slightly modified assertion: for all graphs $(N, \to)$ and $(N', \to')$, for every ordinal α, and for all $x \in N$ and $y \in N'$,

$$\exists R\ Bis(R, (N, \to), x, (N', \to'), y) \supset \varphi_J(\alpha, x, y; N, \to, N', \to')\ .$$

The base case $\alpha = 0$ is trivial, and the induction step is routine. Conversely, ($\Leftarrow$) of (3) is implied by the following assertion ψ: for all graphs $(N, \to)$ and $(N', \to')$, there is an ordinal $\hat{\alpha}$ such that

$$\forall x \in N\ \forall y \in N'\ \ \varphi_J(\hat{\alpha}, x, y; N, \to, N', \to')\ \supset\ \forall^{Ord}\beta\ \varphi_J(\beta, x, y; N, \to, N', \to')\ .$$

The reason is that from ψ, it follows that the set

$$R := \{(x, y) \in N \times N' \mid \varphi_J(\hat{\alpha}, x, y; N, \to, N', \to')\}$$

is a bisimulation between $(N, \to)$ and $(N', \to')$, and, assuming the right hand side of (3) holds, aRa'. Note that ψ^* (where $-^*$ is the translation from the language of set theory to the language of arithmetic given in section 2) is provable in Σ_1^0-IA, the point being that, under $-^*$, the "closure ordinal" $\hat{\alpha}$ can be computed primitive recursively from $(N, \to)$ and $(N', \to')$. By Lemma 5 and the preceding discussion of (3), it follows that

Theorem 11. $\mathsf{KP}_1^{Ord} + \psi$ *is a primitive recursive set theory, relative to which the largest bisimulation is* Δ. *More precisely,* $\mathsf{KP}_1^{Ord} + \psi$ *is proof-theoretically equivalent to* PRA, *and proves (3) for all pointed graphs* $(N, \to), a$ *and* $(N', \to'), a'$.

We leave to the interested reader the question of whether $\mathsf{KP}_1^{Ord} + \psi$ is contained in a theory T proof-theoretically equivalent to PRA, lying between KP^- and KP (plus, if necessary, a global well-ordering principle), for which the passage from $S, \approx, \in$ to $S_0, \approx_0, \in_0$ in the previous section sends models of T to models of $T+\mathsf{AFA}$. The author suggests taking T to be KP_1^{Ord}, although he has (alas) been unable to determine whether or not this works.

References

1. Samson Abramsky. Topological aspects of non-well-founded sets. Handwritten notes.
2. Peter Aczel. An introduction to inductive definitions. In J. Barwise, editor, *Handbook of Mathematical Logic*. North-Holland, Amsterdam, 1977.
3. Peter Aczel. The type-theoretic interpretation of constructive set theory. In *Logic Colloquium '77*. North-Holland, Amsterdam, 1978.
4. Peter Aczel. *Non-well-founded sets.* CSLI Lecture Notes Number 14, Stanford, 1988.
5. Jon Barwise. *Admissible sets and structures.* Springer-Verlag, Berlin, 1975.

6. Jon Barwise and John Etchemendy. *The liar: an essay on truth and circularity.* Oxford University Press, Oxford, 1987.
7. Solomon Feferman. Predicatively reducible systems of set theory. In *Proc. Symp. Pure Math.*, vol 13, Part II. Amer. Math. Soc., Providence, R.I., 1974.
8. Solomon Feferman. A language and axioms for explicit mathematics. In J.N. Crossley, editor, *Algebra and Logic*, LNM 450. Springer-Verlag, Berlin, 1975.
9. Solomon Feferman. Hilbert's program relativized: proof-theoretical and foundational reductions. *Journal of Symbolic Logic*, 53(2), 1988.
10. Tim Fernando. Transition systems over first-order models. Manuscript, 1991.
11. Tim Fernando. Parallelism, partial execution and programs as relations on states. Manuscript, 1992.
12. Gerhard Jäger. A version of Kripke-Platek set theory which is conservative over Peano arithmetic. *Zeitschr. f. math. Logik und Grundlagen d. Math*, 30, 1984.
13. Gerhard Jäger. Induction in the elementary theory of types and names. Preprint, 1988.
14. R.B. Jensen and C. Karp. Primitive recursive set functions. In *Proc. Symp. Pure Math.*, vol 13, Part I. Amer. Math. Soc., Providence, R.I., 1971.
15. Ingrid Lindström. A construction of non-well-founded sets within Martin-Löf's type theory. *Journal of Symbolic Logic*, 54, 1989.
16. M. Mislove, L. Moss, and F. Oles. Non-well-founded sets obtained from ideal fixed points. In *Fourth annual symposium on Logic in Computer Science*, 1989.
17. Charles Parsons. On a number-theoretic choice schema and its relation to induction. In J. Myhill, editor, *Intuitionism and Proof Theory*. North-Holland, Amsterdam, 1970.
18. Alban Ponse. Computable processes and bisimulation equivalence. Technical Report CS-R9207, Centre for Mathematics and Computer Science, 1992.
19. J.J.M.M. Rutten. Non-well-founded sets and programming language semantics. Technical Report CS-R9063, Centre for Mathematics and Computer Science, 1990.
20. J.J.M.M. Rutten. Hereditarily finite sets and complete metric spaces. Technical Report CS-R9148, Centre for Mathematics and Computer Science, 1991.
21. Wilfried Sieg. Fragments of arithmetic. *Annals of Pure and Applied Logic*, 28, 1985.

This article was processed using the LaTeX macro package with LLNCS style