CHALLENGES TO PREDICATIVE FOUNDATIONS OF ARITHMETIC

by

Solomon Feferman[1] and Geoffrey Hellman

Introduction. This is a sequel to our article "Predicative foundations of arithmetic" (1995), referred to in the following as [PFA]; here we review and clarify what was accomplished in [PFA], present some improvements and extensions, and respond to several challenges. The classic challenge to a program of the sort exemplified by [PFA] was issued by Charles Parsons in a 1983 paper, subsequently revised and expanded as Parsons (1992). Another critique is due to Daniel Isaacson (1987). Most recently, Alexander George and Daniel Velleman (1996) have examined [PFA] closely in the context of a general discussion of different philosophical approaches to the foundations of arithmetic.

The plan of the present paper is as follows. Section 1 reviews the notions and results of [PFA], in a bit less formal terms than there and without the supporting proofs, and presents an improvement communicated to us by Peter Aczel. Then Section 2 elaborates on the structuralist perspective which guided [PFA]. It is in Section 3 that we take up the challenge of Parsons. Finally, Section 4 deals with the challenges of George and Velleman, and thereby, that of Isaacson as well. The paper concludes with an appendix by Geoffrey Hellman, which verifies the predicativity, in the sense of [PFA], of a suggestion credited to Michael Dummett for another definition of the natural number concept.

1. Review. In essence, what [PFA] accomplished was to provide a formal context based on the notions of finite set and predicative class and on *prima-facie* evident principles for such, in which could be established the existence and categoricity of a natural number structure. The following reviews, in looser formal terms than [PFA], the notions and results therein prior to any discussion of their philosophical significance. Three formal systems were introduced in [PFA], denoted EFS, EFSC and EFSC*, resp. All are formulated within classical logic. The language *L*(EFS), has two kinds of variables:

Individual variables: *a,b,c,u,v,w,x,y,z,* ..., and

Finite set variables: *A,B,C, F,G,H*,....

[1] This paper was written while the first author was a Fellow at the Center for Advanced Study in the Behavioral Sciences (Stanford, CA) under grants from the Andrew W.Mellon Foundation and the National Science Foundation, whose facilities and support have been greatly appreciated.

The intended interpretation is that the latter range over *finite sets of individuals*. There is one binary operation symbol (,) for a *pairing function* on individuals, and *individual terms* *s,t,...* are generated from the individual variables by means of this operation. We have two relation symbols, = and $\in$, by means of which *atomic formulas* of the form $s = t$ and $s \in A$ are obtained. *Formulas* φ, ψ,... are generated from these by the propositional operations $\neg$, &, $\vee$, $\rightarrow$, and by the quantifiers $\forall$ and $\exists$ applied to either kind of variable. The language *L(*EFSC)*,* which is the same as that of EFSC*, adds a third kind of variable:

Class variables: **X,Y,Z,**....[2]

In this extended language, we also have a membership relation between individuals and classes, giving further atomic formulas of the form $s \in \mathbf{X}$. Then formulas in *L*(EFSC) are generated as before, allowing in addition, quantification over classes. A formula of this extended language is said to be *weak second-order* if it contains no bound class variables. The intended range of the class variables is the collection of weak second-order definable classes of individuals. We could consider finite sets to be among the classes, but did not make that identification in [PFA]. Instead we write $A = \mathbf{X}$ if A and **X** have the same extension. Similarly, we explain when a class is a subclass of a set, and so on. A class **X** is said to be *finite* and we write Fin(**X**) if $\exists A(A = \mathbf{X})$.

The *Axioms of EFS* are denoted (Sep), (FS-I), (FS-II), (P-I) and (P-II), resp.; these are explained as follows. The separation scheme (Sep) asserts that any definable subset of a finite set is finite, i.e. for each formula φ of *EFS*, $\{x \in A \mid \varphi(x) \}$ is a finite set B when A is a given finite set. (FS-I) asserts the existence of an empty (finite) set, and (FS-II) tells us that if A is a finite set and a is any individual then $A \cup \{a\}$ is a finite set. The pairing axioms (P-I) and (P-II) respectively say that pairing is one-one and that there is an urelement under pairing; it is convenient to introduce the symbol 0 for an individual which is not a pair.

The *Axioms of EFSC* augment those of EFS by the scheme (WS-CA) for weak second-order comprehension axiom, which tells us that $\{x \mid \varphi(x) \}$ is a class **X** for any weak second-order φ. In this language, we allow the formula φ in (Sep) to contain free class variables; then it can be replaced by the assertion that any subclass of a finite set is finite.

[2] The class variables are given in boldface, to distinguish them from the finite set variables.

The following theorem (numbered 1 in [PFA]) is easily proved by a model-theoretic argument, but can also be given a finitary proof-theoretic argument.

METATHEOREM. *EFSC is a conservative extension of EFS*.

In the language of EFSC, (binary) relations are identified with classes of ordered pairs, and functions, for which we use the letters **f**,**g**,...,[3] are identified with many-one relations; n-ary functions reduce to unary functions of n-tuples. Then we can formulate the notion of *Dedekind finite class* as being an **X** such that there is no one-one map from **X** to a proper subclass of **X**. By the axiom (Card) is meant the statement that every (truly) finite class is Dedekind finite. The *Axioms of EFSC** are then the same as those of EFSC, with the additional axiom (Card).

Now, working in EFSC, we defined a triple $\langle \mathbf{M}, a, \mathbf{g} \rangle$ to be a *pre-N-structure* if it satisfies the following two conditions:

(N-I) $\forall x \in \mathbf{M}[\ \mathbf{g}(x) \neq a\ ]$, and

(N-II) $\forall x,y \in \mathbf{M}\ [\ \mathbf{g}(x) = \mathbf{g}(y) \rightarrow x = y\ ]$.

These are the usual first two Peano axioms when a is 0 and **g** is the successor operation. By an *N-structure* is meant a pre-N-structure which satisfies the *axiom of induction* in the form:

(N-III) $\forall \mathbf{X} \subseteq \mathbf{M}\ [\ a \in \mathbf{X}\ \&\ \forall x\ (x \in \mathbf{X} \rightarrow \mathbf{g}(x) \in \mathbf{X}\ ) \rightarrow \mathbf{X} = \mathbf{M}\ ]$

It is proved in EFSC that we can define functions by primitive recursion on any N-structure; the idea is simply to obtain such as the union of finite approximations. This union is thus definable in a weak second-order way. From that, we readily obtain the following theorem (numbered 5 in [PFA]):

THEOREM (*Categoricity, in EFSC*) *Any two N-structures are isomorphic*.

[3] As a point of difference with [PFA], function variables here are given in boldface in order to indicate that they are treated as special kinds of classes.

Now to obtain existence of N-structures, in [PFA] we began with a specific pre-N-structure $\langle \mathbf{V}, 0, \mathbf{s} \rangle$, where $\mathbf{V} = \{x \mid x = x\}$ and $\mathbf{s}(x) = x' = (x,0)$; that this satisfies (N-I) and (N-II) is readily seen from the axioms (P-II) and (P-I), resp. Next, define

(1) $$\mathrm{Clos}^{-}(A) \leftrightarrow \forall x[\, x' \in A \rightarrow x \in A\,],$$

and

(2) $$y \leq x \leftrightarrow \forall A\, [x \in A \;\&\; \mathrm{Clos}^{-}(A) \rightarrow y \in A\,].$$

In words, $\mathrm{Clos}^{-}(A)$ is read as saying that A is closed under the predecessor operation (when applicable), and so $y \leq x$ holds if y belongs to every finite set which contains x and is closed under the predecessor operation. Let

(3) $$\mathrm{Pd}(x) = \{y \mid y \leq x\}.$$

The next step in [PFA] was to cut down the structure $\langle \mathbf{V}, 0, \mathbf{s} \rangle$ to a special pre-N-structure:

(4) $$\mathbf{M} = \{\, x \mid \mathrm{Fin}(\mathrm{Pd}(x)) \;\&\; \forall y[\, y \leq x \rightarrow y = 0 \vee \exists z(y = z')\,]\,\}.$$

This led to the following theorem (numbered 8 in [PFA]):

THEOREM *(Existence, in EFSC*)* $\langle \mathbf{M}, 0, \mathbf{s} \rangle$ *is an N-structure.*

To summarize: in [PFA], categoricity of N-structures was established in EFSC and existence in EFSC*. Following publication of this work we learned from Peter Aczel of a simple improvement of the latter result obtained by taking in place of $\mathbf{M}$ the following class:

(5) $$\mathbf{N} = \{x \mid \mathrm{Fin}(\mathrm{Pd}(x)) \;\&\; 0 \leq x\}.$$

THEOREM (Aczel). *EFSC proves that* $\langle \mathbf{N}, 0, \mathbf{sc} \rangle$ *is an N-structure.*

We provide the proof of this here, using facts established in THEOREM 2 of [PFA].

(i) $0 \in \mathbf{N}$, because $Pd(0) = \{0\}$ and $0 \le 0$.

(ii) $x \in \mathbf{N} \to x' \in \mathbf{N}$, because $Pd(x') = Pd(x) \cup \{x'\}$, and $0 \le x \to 0 \le x'$.

(iii) If $\mathbf{X}$ is any subclass of $\mathbf{N}$ and $0 \in \mathbf{X} \wedge \forall y[\, y \in \mathbf{X} \to y' \in \mathbf{X}\,]$ then $\mathbf{X} = \mathbf{N}$. For, suppose that there is some $x \in \mathbf{N}$ with $x \notin \mathbf{X}$. Let $A = \{\, y \mid y \le x \;\&\; y \notin \mathbf{X} \,\}$; A is finite since it is a subclass of the finite set Pd(x). Moreover, A is closed under predecessor, so A contains every $y \le x$; in particular, $0 \in A$, which contradicts $0 \in \mathbf{X}$.

The theorem follows from (i)-(iii), since the axioms (N-I) and (N-II) hold on **V** and hence on **N**.

It was proved in [PFA] that EFSC* is of the same (proof-theoretic) strength as the system PA of Peano Axioms and is a conservative extension of the latter under a suitable interpretation. The argument was that EFSC* is interpretable in the system ACA_0, which is a well-known second-order conservative extension of PA based on the arithmetical comprehension axiom scheme together with induction axiom in the form (N-III). Conversely, we can develop PA in EFSC* using closure under primitive recursion on any N-structure. Since any first-order formula of arithmetic so interpreted then defines a class, we obtain the full induction scheme for PA in EFSC*. Now, using the preceding result, the whole argument applies *mutatis mutandis* to obtain the following:

METATHEOREM (Aczel). *EFSC is of the same (proof-theoretic) strength as PA and is a conservative extension of PA under the interpretation of the latter in EFSC.*

This result also served to answer QUESTION 1 on p.13 of [PFA].

Incidentally, it may be seen that the definition of **N** in (5) above is equivalent to the following:

(6) $$x \in \mathbf{N} \leftrightarrow \forall A[\, x \in A \;\&\; Clos^{-}(A) \to 0 \in A\,] \;\&\; \exists A[\, x \in A \;\&\; Clos^{-}(A)\,]\,.$$

For, the first conjunct here is equivalent to the statement that $0 \le x$, and the second to Fin(Pd(x)). In this form, Aczel's definition is simply the same as the one proposed by

George (1987), p.515.[4] Part of the progress that is achieved by this work in our framework is to bring out clearly the assumptions about finite sets which are needed for it and which are *prima-facie* evident for that notion.

There is one further improvement in our work to mention. It emerged from correspondence with Alexander George and Daniel Velleman that the remark in Footnote 5 on p.16 of [PFA] asserting a relationship of our work with a definition of the natural numbers credited to Dummett was obscure. The exact situation has now been clarified by Geoffrey Hellman in the Appendix to this paper, where it is shown that Dummet's definition also yields an N-structure, provably in EFSC.

[4] That,in turn, was a modification of a definition of the natural numbers proposed by Quine (1969) using only the first conjunct in (6), which is adequate when read in strong second-order form, but not when read in weak second-order form; cf. George (1987), p.515, and George and Velleman (1996), n.10.