Notre Dame Journal of Formal Logic
Volume 41, Number 2, 2000

A SINGLE AXIOM FOR SET THEORY

DAVID BENNETT

Abstract Axioms in set theory typically have the form $\forall z \exists y \forall x(x \in y \leftrightarrow Fxz)$, where F is a relation which links x with z in some way. In this paper we introduce a particular linkage relation L and a single axiom based on L from which all the axioms of Z (Zermelo set theory) can be derived as theorems. The single axiom is presented both in informal and formal versions. This calls for some discussion of pertinent features of formal and informal axiomatic method and some discussion of pertinent features of the system S of set theory to be erected on the single axiom. S is shown to be somewhat stronger than Z, but much weaker than ZF (Zermelo-Fraenkel set theory).

1. Introduction

The Cantorian notion of a set has dominated much of the mathematics of the twentieth century while remaining highly problematic. The following are only a few of many possible ways to approach this notion:

1. One may choose to use the Cantorian notion with great freedom, convinced that problems arise only in remote regions of set theory which need not be entered. Many mathematicians who take this approach can scarcely be said to subscribe to any "theory" of sets, toward which in fact they tend to be somewhat hostile.

2. One may be convinced that a cumulative hierarchical notion of set, still essentially Cantorian, can settle the problems. The wide acceptance of this approach is due in significant measure to the influential advocacy of it by Gödel in [4].

3. Following the early lead of Zermelo, one may choose to rein in the set notion axiomatically, still with the attitude that there are basic truths about sets, some of which can be captured by the axioms. This approach will normally be combined with (2).

Received March 23, 1999; revised February 16, 2001; printed August 30, 2002
2001 Mathematics Subject Classification: Primary, 03E30; Secondary, 00A30
Keywords: set, Cantorian, axiom, links, descendant, single, foundation, basic

4. Acknowledging that the axioms of set theory admit of different models, one may abandon the notion that a proposition about sets is simply true or false and take up instead the notion that it will be true in one model, false in another. The Cantorian notion of a set then becomes weakened, by questioning whether there can be a single domain of all sets, but may still play a central role in that the domain of each of the models will itself be a Cantorian set.

5. One may see axiomatic set theory as a way of *defining* the set notion. Given that the Continuum Hypothesis, for example, is both consistent with and independent from the other axioms of set theory, someone following this approach would be free to adjoin either CH or its negation to the other axioms. It is difficult to find anything more than a heuristic role for the Cantorian notion of a set in this approach.

The axiomatic method one adopts along with (3), (4), or (5) may be either formal or informal, and some people have strong preferences favoring formal over informal method, or vice versa. The complexities of the situation are further compounded by the fact that mathematicians, set theorists, logicians, and philosophers have quite different and frequently incompatible motives for taking an interest in sets. Accordingly, an attempt to reach readers of all these sorts, as I hope to do, is rather perilous. My strategy for mitigating the peril is to acknowledge, up front, that some of the ideas introduced below are controversial. Whenever I perceive an idea to be problematic in some way I will attempt to exhibit its drawbacks alongside its virtues and let readers decide for themselves which are which. Finally, readers are advised to skim or skip material which they perceive to be addressed primarily to other readers.

Here now, in outline, is a preview of the basic content of the seven sections of the paper which follow: in Section 2 we define and study a novel relation called linkage and show how the linkage relation can function as a restriction on an inconsistent "naïve axiom" for set theory; in Section 3 we show how all the axioms of Zermelo set theory (Z) except the Axiom of Infinity can be derived from a single axiom based on the linkage relation, using formal axiomatic method; in Section 4 we show how all the axioms of Z, including the Axiom of Infinity, can be derived from a stronger single axiom, using either informal or formal axiomatic method; in Section 5 we investigate basic features of the system S of set theory which results from adopting the new axiom and show to what extent S is stronger than Z; in Section 6 we discuss particular features of informal and formal axiomatic method which pertain to our axioms; in Section 7 we discuss the Axiom of Replacement and the Axiom of Foundation as these axioms bear on S while not having been included in S; in Section 8 I express my personal opinions about a few of the controversial matters the paper discusses.

The basic motivation leading to the discovery of the single axiom of S was the idea of starting with a naïve (inconsistent) comprehension axiom which could be suitably restricted by exploiting the hierarchical or wellfoundedness notion of a set. Linkage turned out to be the sought for restriction. Ironically, while it is possible to show in the metalanguage of S that if S is consistent no infinite descending $\in$-chains can be defined in S, no such result can be proved within S itself. So, while S was strongly motivated by considerations embodied in the Axiom of Foundation, this axiom is not provable in S and remains peripheral to S in the same way it is peripheral to Z or to ZF: both Z and ZF are strongly motivated by considerations embodied in the

Axiom of Foundation, but in neither Z nor ZF can the existence or nonexistence of non-wellfounded sets be proved.

2. The Linkage Relation and the Single Axiom

Let us proceed at once to define the new linkage relation:

$$x\,L\,y \equiv_{\mathrm{df}} \exists n \forall x_0, \dots, x_n (x_n \in \cdots \in x_0 = x \rightarrow \exists m \exists y_0, \dots, y_m (x_n = y_m \in \cdots \in y_0 = y)). \tag{1}$$

Read this informally as "x links with y". For the present we will leave the definition of L in its informal state, not yet attempting to define it more precisely in terms of the primitive symbol $\in$.[1] Our immediate task is to get a good intuitive grasp of its meaning and potential usefulness. The first thing to notice is how closely L is involved with the following "descendance relations":

$$x\,D^n y \equiv_{\mathrm{df}} \exists y_0, \dots, y_n (x = y_n \in \cdots \in y_0 = y), \tag{2}$$

$$x\,D y \equiv_{\mathrm{df}} \exists n\; x\,D^n y. \tag{3}$$

Read these relations as "x is an nth-generation descendant of y" and "x is a descendant of y". Then if we note that $x\,L\,y$ is equivalent to

$$\exists n \forall z (z\,D^n x \rightarrow z\,D y),\text{[2]} \tag{4}$$

we can read $x\,L\,y$ quite simply as "there exists an n such that every nth-generation descendant of x is a descendant of y". $x\,L\,y$ can also be read more loosely as "there exists an n such that every $\in$-chain of length n descending from x links up with some $\in$-chain descending from y". Notice that every set is a 0th-generation descendant of itself: when $n = 0$, $x = x_0 \in \cdots \in x_0 = y$ reduces to $x = x_0 = y$.

(Warning: some readers have found this terminology confusing and have urged me to consider switching "descendant" to "ancestor". Anyone for whom this might be a problem is advised to consult note 3 now before proceeding further.)[3]

Now consider the following "naïve axiom":

$$\forall F \exists_1 y \forall x (x \in y \leftrightarrow Fx). \tag{5}$$

Read this as "for any property F there exists a unique set y whose members are those x's which satisfy F". (5) is inconsistent of course: take Fx to be $x \notin x$ and (5) implies $\exists y (y \in y \leftrightarrow y \notin y)$ (Russell's Paradox). But we obtain our new single axiom, now consistent, simply by incorporating the linkage relation into (5) as follows:

Axiom 1 $\quad \forall F \forall z \exists_1 y \forall x (x \in y \leftrightarrow Fxz \wedge x\,L\,z).$

Compare Axiom 1 with Zermelo's celebrated *Aussonderungsaxiom* (Axiom of Separation):

$$\forall F \forall z \exists y \forall x (x \in y \leftrightarrow Fx \wedge x \in z).$$

It is now clear that what we are doing here is taking the naïve notion of a set embodied in (5) and restricting it in the same spirit and by a very similar device that Zermelo used for this purpose but much more weakly.

Some readers may find it helpful at this point to see that xLy is equivalent to $\exists n(x \in P^n(TCy))$ (for some n, x belongs to the nth iterated power set of the transitive closure of y). Then the Axiom of Separation says that for any property F and set z the set $\{x \in z|Fx\}$ exists, and Axiom 1 says that for any relation F and set z the set $\{x \in Vz|Fxz\}$ exists, where Vz is the union of the iterated power sets of the transitive closure of z. This formulation gives the content of Axiom 1 in terms that may be more familiar to some readers, less familiar to others; for either sort of reader the official form of Axiom 1 avoids the need to define the various set operations we have just used.

As mentioned earlier, the seminal motivation behind Axiom 1 was the idea of reining in the set notion by means of the wellfoundedness notion or, equivalently, the rejection of infinitely descending $\in$-chains of the form $\cdots x_2 \in x_1 \in x_0$. The details of how this rejection works are embodied in the following:

If x links with y and no infinite $\in$-chain descends from y then no infinite $\in$-chain descends from x. (6)

If Axiom 1 defines a set y relative to a given set z then every member of y (but not necessarily y itself) links with z. (7)

If Axiom 1 defines a set y relative to a given z, and no infinite $\in$-chain descends from z, then no infinite $\in$-chain descends from y.[4] (8)

It is clear from (8) that the recursive generation of sets by Axiom 1, starting from the empty set zero, will never produce an infinitely descending $\in$-chain, (unless Axiom 1 is inconsistent!).

It is tempting to regard (8) itself as evidence for the consistency of Axiom 1. Suppose we are further tempted to put a simpler and weaker restriction $L_?$ in place of the linkage restriction in Axiom 1 to obtain a stronger axiom:

$xL_?y \equiv_{\text{df}}$ every infinite $\in$-chain descending from x contains some link which is a descendant of y.

Axiom $1_?$ $\quad \forall F \forall z \exists_1 y \forall x(x \in y \leftrightarrow Fxz \wedge x\ L_?z)$.

Then (8) goes through just as before with $L_?$ and Axiom $1_?$ substituted for L and Axiom 1, and we can conclude as before that Axiom $1_?$ will never produce an infinitely descending $\in$-chain unless it is inconsistent. But Axiom $1_?$ falls prey to a version of the Mirimanoff paradox and is, in fact, inconsistent.[5] So if (8) cannot protect Axiom $1_?$ from inconsistency then it will fail to protect Axiom 1 as well. But at least Axiom 1 does not fall into the Mirimanoff trap, either directly or in the manner of Axiom $1_?$, and we can hope, on other grounds, that it is consistent.

Note that Axiom 1 is not yet a sentence of a formal language, even if we replace the abbreviation xLz by its definition, the latter remaining in an informal state. Whether the linkage relation and Axiom 1 are to be regarded as receiving better treatment by informal or by formal axiomatic method is one of the controversial features of our project mentioned in the Introduction. A more in-depth discussion of this question will be undertaken in Section 6.

3. Z without the Axiom of Infinity

We take the following to be all the axioms of Z except the Axiom of Infinity:

Extensionality	$\forall y \forall z(\forall x(x \in y \leftrightarrow x \in z) \rightarrow y = z)$,
Pair	$\forall z_1 \forall z_2 \exists y \forall x(x \in y \leftrightarrow x = z_1 \vee x = z_2)$,
Union	$\forall z \exists y \forall x(x \in y \leftrightarrow \exists u(x \in u \wedge u \in z))$,
Powerset	$\forall z \exists y \forall x(x \in y \leftrightarrow \forall u(u \in x \rightarrow u \in z))$,
Separation	$\forall F \forall z \exists y \forall x(x \in y \leftrightarrow Fx \wedge x \in z)$,
Choice[6]	$\exists F(\forall x \exists_1 y Fxy \wedge \forall x \forall y(\exists z\, z \in x \wedge Fxy \rightarrow y \in x))$.

These axioms contain only symbols of the language of second-order logic and the primitive symbol $\in$ of set theory. The expressions occurring in Axiom 1 are not thus confined. As it turns out, however, the only value of n needed for deriving all the axioms of Z except Infinity from Axiom 1 is the value $n = 1$. So we can substitute the following in place of Axiom 1:

Axiom 2 $\quad \forall F \forall z \exists_1 y \forall x(x \in y \leftrightarrow Fxz \wedge \forall w(w \in x \rightarrow wD'z))$,

where D' is a formal version of D defined as follows:

$$xD'y \equiv_{\mathrm{df}} \forall z(y \in z \wedge \forall u \forall v(v \in z \wedge u \in v \rightarrow u \in z) \rightarrow x \in z). \tag{9}$$

$xD'y$ is equivalent to $x \in TCy$. Thus Axiom 2 is equivalent to "for any relation F and set z, $\{x \subseteq TCz | Fxz\}$ exists". Notice that (9) does not give the intended meaning of D' as the descendant relation unless for every z there exists a y which contains all the descendants of z; this is readily guaranteed by Axiom 2, as we will see presently.

If Fxz is any sentence with free variables x and z such that Fxz implies $\forall w(w \in x \rightarrow wDz)$ then $Fxz \wedge \forall w(w \in x \rightarrow wDz)$ is equivalent to Fxz and the following is a theorem by Axiom 2:

$$\forall z \exists_1 y \forall x(x \in y \leftrightarrow Fxz). \tag{10}$$

So (10) can function as a theorem scheme; to prove that a sentence of the form (10) is a theorem, simply verify that Fxz implies $\forall w(w \in x \rightarrow wD'z)$. For example, the existence of the set of descendants of z needed to give D' its proper meaning follows from the fact that $xD'z$ implies $\forall w(w \in x \rightarrow wD'z)$. Similarly, we see at once that Union, Powerset, and Separation follow from Axiom 2.[7] Extensionality comes from the following which also comes by way of the theorem scheme (10):

$$\forall z \exists_1 y \forall x(x \in y \leftrightarrow x \in z).$$

We will show later on, in Section 5, that the Pair axiom is not needed in our system S; its purposes can be equally well served by the following:

$$\forall z_1 \forall z_2 \forall z(z_1 \in z \wedge z_2 \in z \rightarrow \exists y \forall x(x \in y \leftrightarrow x = z_1 \vee x = z_2)).^{8} \tag{11}$$

(11) follows readily from the Axiom of Separation as we will see presently.

Finally, we must acknowledge that the Axiom of Choice bears no particular relationship to Axioms 1 or 2, but Choice comes at once from the following valid formula of second-order logic:

$$\forall G \exists F(\forall x \exists_1 y Fxy \wedge \forall x \forall y(\exists z Gzx \wedge Fxy \rightarrow Gyx)).^{9} \tag{12}$$

All of these axioms of Z except Separation and Choice have the same formulation whether in first- or second-order logic. For various reasons some people prefer axioms of set theory formulated in first-order logic. Then Separation, no longer a single axiom, becomes an *axiom-scheme*, generating infinitely many axioms each having the following form:

$$\forall z_1, \ldots, \forall z_n \forall z \exists y \forall x (x \in y \leftrightarrow F x z_1, \ldots, z_n \wedge x \in z). \tag{13}$$

When $n = 2$, (13) gives us just what we need to derive (11) as promised a few lines back:

$$\forall z_1 \forall z_2 \forall z \exists y \forall x (x \in y \leftrightarrow (x = z_1 \vee x = z_2) \wedge x \in z).$$

In a first-order treatment Axioms 1 and 2 are transformed from single axioms into axiom-schemes in the same way that separation is turned into (13). To derive an axiom of the form (13) from the Axiom of Separation simply instantiate F to $F x z_1, \ldots, z_n$ and then generalize on $z_1, \ldots, z_n$.

The Axiom of Choice can also be formulated in first-order logic, augmented by a very useful but much neglected device, Hilbert's ε-operator, as follows:

$$\forall y (\exists x\; x \in y \rightarrow (\varepsilon x\; x \in y) \in y)).\text{[10]} \tag{14}$$

In this new form the Axiom of Choice can be derived from a corresponding formula of first-order logic:

$$\forall y (\exists x F x y \rightarrow F (\varepsilon x F x y) y)).$$

Readers who are unfamiliar with Hilbert's ε-operator will find a brief introduction to it in note 10. For further details see Hilbert and Bernays [6], pp. 9–18.

Here then is the situation: Axiom 2 and the Axiom of Infinity are all we need as a basis for Z. While Axiom 2 could be written out on a single line using only primitive notation, the result would be rather messy. A clean and simple alternative, still in primitive notation, is to regard D' as a second primitive alongside $\in$ and treat (9) as a further axiom governing this primitive. Then all the Zermelo axioms except the Axiom of Infinity are derivable from two axioms, (9) and Axiom 2.[11]

4. The Axiom of Infinity

The standard form for the Axiom of Infinity in Z (due to Zermelo—not the current form due to von Neumann) is

Infinity $\quad \exists z (0 \in z \wedge \forall x \forall y (x \in z \wedge y S x \rightarrow y \in z)),$

where S is a successor relation defined as follows:

$$y S x \equiv_{\text{df}} \forall z (z \in y \leftrightarrow z = x).$$

Informally, Infinity follows from the existence of the set of numbers, that is,

$$\exists y \forall x (x \in y \leftrightarrow \exists n \exists x_0, \ldots, \exists x_n (x = x_n\; S, \ldots, S\; x_0 = 0)), \tag{15}$$

where 0 is defined from Axiom 1 by taking F to be $x \neq x$.

(15) follows from Axiom 1: using the strategy we learned from theorem scheme (10) we need only verify that $\exists n \exists x_0, \ldots, \exists x_n (x = x_n\; S, \ldots, S\; x_0 = 0)$ implies $x L 0$, that is, $\exists n \forall y (y D^n x \rightarrow y D 0)$.

To obtain the other axioms of Z informally from Axiom 1, first get $\forall z \exists_1 y \forall x (x \in y \leftrightarrow \forall w (w \in x \rightarrow w D z) \wedge x L z))$, notice that $\forall w (w \in x \rightarrow w D z)$ implies $x L z$

(take n to be 1) and derive these other axioms just as in Section 3, with D now substituted in place of D'.

So it appears from the point of view of informal axiomatic method that if we countenance such standard expressions as $x_1 \in \cdots \in x_n$ and $x_1 S, \ldots, Sx_n$, and such standard forms of reasoning with these expressions as we have just employed, all the axioms of Z including Infinity follow from Axiom 1.

Our next task is to devise a formal definition of the linkage relation and a formal version of Axiom 1, prove the Axiom of Infinity, and prove Axiom 2. To this end we introduce the following definitions and lemmas:

$Ex \equiv_{\mathrm{df}} \forall y\, y \notin x,$

$Nn \equiv_{\mathrm{df}} \forall j (jD'n \to Ej \vee \exists i\ jSi),$[12]

$xD^n y \equiv_{\mathrm{df}} Nn \wedge \exists F(\forall u \forall v \forall w(Fuv \wedge Fuw \to Fvw) \wedge \exists u(Eu \wedge Fux) \wedge Fny \wedge \forall i \forall j(jD'n \wedge jSi \to \exists u \exists v(Fiu \wedge Fjv \wedge u \in v))),$

$xL'y \equiv_{\mathrm{df}} \exists n(Nn \wedge \forall z(zD^n x \to zD'y)).$

$\forall F \forall z \exists_1 y \forall x(x \in y \leftrightarrow Fxz \wedge xL'z),$ Axiom 1′

$ySx \wedge nSm \wedge zD^n y \to zD^m x,$[13] Lemma 1

$\exists_1 n(Nn \wedge En) \wedge xLx \wedge (Nm \to \exists_1 n(Nn \wedge nSm)).$[14] Lemma 2

$0 \equiv_{\mathrm{df}} \iota y Ey,$

$1 \equiv_{\mathrm{df}} \iota y\ yS0,$

$V_0 \equiv_{\mathrm{df}} \iota y(\forall x(x \in y \leftrightarrow xL0)).$

$xD^1 y \to x \in y.$[15] Lemma 3

The Axiom of Infinity and Axiom 2 now follow at once from Axiom 1′.

Proof of Infinity: $0L0$ so $0 \in V_0$. Assume $x \in V_0 \wedge ySx$. Then there exists a number m such that $\forall z(zD^m x \to zD'0)$ and a number n such that nSm. So $zD^n y \to zD^m x$, $\forall z(zD^n x \to zD'0)$, $yL0$, $y \in V_0$, and $\forall x \forall y(x \in V_0 \wedge ySx \to y \in V_0)$. □

Proof of Axiom 2: $\forall w(w \in x \to wD'z)$ implies $\forall w(wD^1 x \to wD'z)$ implies $xL'z$. □

To obtain the same results in first-order logic we represent relational expressions of the form Fxy by means of ordered pairs as follows:

$$zPxy \equiv_{\mathrm{df}} \forall w(w \in z \leftrightarrow w = x \vee w = y),$$
$$z\overrightarrow{P}xy \equiv_{\mathrm{df}} \forall w(w \in z \leftrightarrow wPxx \vee wPxy),$$
$$fxy \equiv_{\mathrm{df}} \exists z(z \in f \wedge z\overrightarrow{P}xy).$$

Then if we replace the uppercase letter F by the lowercase letter f in all definitions, lemmas, axioms, and proofs, remove the initial quantifier from Axiom 1′, and read Axiom 1′ as an axiom scheme rather than a single axiom, it turns out that our entire second-order treatment will serve equally well as a first-order treatment.

A surprising feature of this transformation of the second-order proofs into first-order proofs is that the three additional definitions (for pairs, ordered pairs, and relations) make no appearances in the new proofs; the relational expression fxy

could be defined in any arbitrary way and the proofs would still go through correctly as given. But while the definitions are not needed for the proofs, they are needed to give the intended meanings to expressions in which fxy occurs, and we rely on these meanings as our only basis for believing that our axiom system will be consistent. Give fxy some different definition and risk contradiction; give it the definition we do and it is reasonable to believe that no contradiction will arise.

We now have a basis for comparing the formal and informal versions of the linkage relation. It could be argued that the informal L of (4) wears its meaning on its sleeve whereas the formal L' complicates and obscures this simple meaning. It could also be argued that the simplicity of (4) is deceptive and that its true meaning can only be fully grasped by unpacking it as we have done. Finally, when we compare the formal and informal versions of our single axiom, we observe that the informal Axiom 1 can be expressed in primitive notation together with the notation of numbered quantifiers in common usage in informal proofs. These quantifiers offer direct access to simple proofs of all the axioms of Z. The formal Axiom 1′ would be hopelessly messy if expressed in primitive notation, and formal proofs based on Axiom 1′, though not very complicated, are far more so than their informal counterparts. But the formal proofs use only standard first- or second-order logic, whose rules are familiar and explicit, unlike the informal and implicit logic which underlies the use of the numbered quantifiers.

5. Beyond Z

With all the definitions and theorems of Z now available to us, we next proceed to show to what extent the system S which we obtain informally from Axiom 1, or formally from Axiom 1′ or from Axiom 2 and an appropriate axiom of infinity, is stronger than Z. In doing so we will arrive at the particular form of the axiom of infinity we propose to conjoin with Axiom 2 to bring the formal versions based on Axiom 1′ and Axiom 2 into harmony with each other.

Perhaps the most widely remarked example of a set not available in Z is $\{N, PN, PPN, \ldots\}$, where N is the set of numbers and P is the power-set function. In S this set can be obtained informally by defining the power-set relation xPy as $\forall z(z \in x \leftrightarrow z \subseteq y)$ and verifying that the formula to the right of the double arrow in the following implies xLN:

$$\exists_1 y \forall x(x \in y \leftrightarrow \exists n \exists x_0, \ldots, x_n(x = x_n P, \ldots, P x_0 = N)). \tag{16}$$

Similarly, we can define

$$U_0 = \cup\{0, P0, PP0, \ldots\}$$

$$U_{n+1} = \cup\{U_n, PU_n, PPU_n, \ldots\}$$

to obtain additional higher level sets in S which are not available in Z. If we attempt a similar definition for a would-be set $U = \cup\{U_1, U_2, \ldots\}$, however, we encounter a failure of linkage between the sets U_{n+1} and U_n needed in analogy to (16). Even though every chain descending from U_{n+1} eventually links up with U_n, there exists no n such that every nth-generation descendant of U_{n+1} is a descendant of U_n. So it appears that all sets definable in S belong to U, but that U itself is not definable in S.

A clear picture of what sets can and cannot be defined in S is now emerging. We can sharpen the focus of this picture by examining just what sets are, in fact, generated by Axiom 1. Consider the sets $V_0, V_1, \ldots$ which can be generated from the formula

$$\forall z \exists_1 y \forall x (x \in y \leftrightarrow x L z)$$

by instantiating z with 0 as we have already done earlier to get V_0, and with V_n to get V_{n+1}. It is quite easy to show, by induction, that every set y which can be generated by Axiom 1 will belong to V_n for some n.[16] So S has as its principal model the set

$$V = \cup\{V_0, V_1, \ldots\}, \tag{17}$$

that is, the intersection of all intended models for S. This model is a Cantorian set, but it is not, of course, a set definable in S.

Next, notice that the sets $V_0, V_1, \ldots$ are just the sets $U_0, U_1, \ldots$ encountered two paragraphs back. This is readily seen from the definitions of the Us and the Vs.[17] So the universe $U = V$ is the principal model for S. And the appropriate form of an axiom of infinity needed to guarantee that the formal versions of S based on Axiom 1′ or on Axiom 2 will have the same principal model as the informal version is

Axiom 3 $\quad \forall x \exists y (x \in y \wedge \forall u \forall v (u \in y \wedge v P u \rightarrow v \in y)).$

This is readily seen by instantiating x in Axiom 3 with 0 to get U_0 and with U_n to get U_{n+1}.

We are now in a position to fulfill an earlier promise to show how (11) serves as a substitute for the Pair axiom for all practical purposes: any two sets x_1 and x_2 which we could ever encounter in practice would both belong to some set U_n, thus allowing (11) to do its work.

We now see that the sets of S rise somewhat higher in the Cantorian transfinite than do the sets of Z: the principal model for Z is the set V_1 whereas the principal model for S is the set $V = \cup\{V_0, V_1, \ldots\}$. While many occasions arise for working with sets higher than V or lower than V_1, it would be a rare event when any of the sets $V_2, V_3, \ldots$ would come into play. So V_1 is a more natural model than V, and Z is a more natural working system of sets than S. Occasions do sometimes arise, however, when it is convenient to have V_1 as a set, which it is in S but not in Z. This suggests a limited practical role for S along with its main theoretical role—the derivability of most of contemporary mathematics from a single axiom.

Neither Z nor S is a theory capable of dealing with all the transfinite sets and numbers originally envisioned somewhat vaguely by Cantor or those more precisely definable when the Axiom of Replacement is adjoined to Z to produce ZF. Of course, ZF is obtainable at once from S by adjoining Replacement to Axiom 1, informally, or to Axioms 1′ or Axioms 2 and 3, formally.

6. Informal Method

By now the reader may be convinced that the linkage notion has interesting potential for reducing the number of axioms required in basic set theory. But there are problems: on one hand informal derivations from Axiom 1, while appropriate heuristically, might readily be regarded as falling well short of genuine proofs; on the other hand Axiom 1′, while formally correct, carries perhaps too much definitional baggage to qualify as a single axiom.

The informal approach which is dictated when Axiom 1 rather than Axiom 1′ is chosen can be tightened up somewhat. First let us become clearer about what constitutes a sentence in our system S, in order to represent relations by open sentences. Sentences are to be constituted from the primitive symbol $\in$ and standard logical operations, as with the formal method, but they may also be introduced into the system by definitions of standard type, either as simple abbreviations or as definitions by induction. For example, define xD^ny inductively as $x = y$ for the numeral 0, and as $\exists z(x \in z \wedge zD^ny)$ for the numeral $n + 1$. On the basis of this definition each of the following expressions will qualify as a sentence in S: $xD^ny, xD^1y, xD^2y, \ldots$. Inductive definitions will introduce numerals, number variables, and newly defined expressions into our sentences.

With this understanding of the notion of an informal sentence we turn next to the notion of informal implication: a set of sentences $S_1, \ldots, S_n$ will be said to imply a sentence S if and only if S is true in every interpretation for which all the sentences $S_1, \ldots, S_n$ are true. An interpretation is an assignment of a meaning for $\in$ as some arbitrary 2-term relation on some arbitrary domain of objects. In interpreting a sentence S we assume that the numerals and logical operators appearing in S have their standard meanings as numbers and logical operations. Set variables occurring in S are no longer to be regarded as necessarily ranging over sets—they range over whatever is the domain of objects on which $\in$ has been interpreted. Number variables range over numbers and relation variables range over relations defined by open sentences.

From the beginning we have availed ourselves of the three different styles of variables. In the formal method the number variables have been used only as a convenient subclass of the set variables to designate certain sets we have taken as numbers. In the informal method we are now outlining, confusion will arise unless we are careful to keep the number variables and the set variables distinct from each other. To this end we now are to understand that the only sentences which contain numerals or number variables are those which are introduced by inductive definitions or abbreviations.

Summarizing, the open sentences we use to represent properties and relations are defined inductively rather than formally, and we use standard procedures of demonstration based on the notion of semantic implication instead of formal proof procedures based on syntactic rules.

What we have just offered constitutes an account of certain features of informal axiomatic method which are pertinent to our particular project. Giving at least this brief account is important for the purpose of removing a worrisome threat of circularity in the method. Many informal proofs in mathematics appeal to our notions about sets in ways which obviously should be avoided, if possible, when we are involved in the project of providing a foundation for these very notions. Now we have shifted our proof procedures away from set notions toward pure logic and arithmetic.[18] To fix the point firmly in mind, the reader is invited, as an exercise, to review the informal demonstration of the Axiom of Infinity in Section 4, to ensure that the argument can be carried out on the basis of the new account of sentence and the new account of implication, free of all taint of circular set-theoretic reasoning.

Notice, however, that circularity is not always vicious. Rather than seeing Axiom 1 as constituting a foundation for the other axioms of set theory we might simply be interested in seeing how linkage and Axiom 1 interact with these other axioms. Then we could label Axiom 1 as “Proposition 1” and investigate the interaction with complete freedom from any concerns about circularity. Nor is a loose heuristic method always undesirable. It is still the clearest way of getting an initial hold on the central ideas of our project. Though formally more correct, Axiom 2 lacks motivation except as derived from Axiom 1, and it is not apparent that formal correctness bestows a higher degree of conceptual clarity on Axiom 1′ than is enjoyed by Axiom 1.

With the informal method somewhat spelled out, looseness and potential circularity become less worrisome. But opting for semantic as opposed to syntactic proof procedures is problematic. Triple ambiguity lurks in the concept of a theorem in an axiom system: should we say (1) that every sentence implied by the axioms is a theorem, (2) that every sentence which can be shown to follow from the axioms is a theorem, or (3) that every sentence which has been shown to follow from the axioms is a theorem? Given that the proof procedures of first-order logic are complete, theorems in senses 1 and 2 are the same as each other for first-order axiom systems, and some hesitation in calling something a theorem prior to its having been demonstrated as such (sense 3) seems of little consequence. But in an informal semantic method such as we are proposing here the ambiguity between (1) and (2) is not easy to resolve comfortably. We would probably prefer to say that something is a theorem only if it can be shown to be such (sense 2), but in the absence of precise and complete proof procedures we have been forced to fall back on sense 1 and say that something is a theorem whenever it is implied by the axioms, whether or not this can ever be shown.

Absence of complete proof procedures for second-order logic subject second-order axiom systems to the same and similar ambiguities. We have already mentioned the fact, for example, that practitioners of second-order logic do not all agree on whether an axiom or principle of choice should be countenanced as a theorem of this logic. But many of the proof procedures of second-order logic are agreed upon and fully explicit, including all the second-order procedures we have used in this paper except as noted for the axiom of choice. In these regards second-order logic is considerably more stable than is the informal inductive method we have been discussing.

A surprising (and no doubt controversial) feature of our informal method is that every true sentence of elementary arithmetic is a theorem of S. To see why, first take elementary arithmetic (call this A) to be the set of first-order sentences built up from two primitives, $Sxyz$ and $Pxyz$, interpreted as $x + y = z$ and $x * y = z$, respectively. Upon giving the standard inductive definitions for S and P, it follows from the account we have given above for the notion of a sentence of S that each sentence of A is a sentence of S—this in spite of the fact that no sentence of A contains the symbol $\in$ or any set variables. But then each sentence of A which is true is true in every (irrelevant) interpretation which we might give to $\in$. So every true sentence of A is true in every interpretation of $\in$ for which the axioms of S are true. So, by our account of informal implication, every true sentence of A is implied by these axioms. Finally, by our recent decision regarding how to understand the notion of a theorem, each true sentence of A turns out to be a theorem of S.

This somewhat bizarre consequence of our treatment of informal axiomatic method suggests that by incorporating arithmetic alongside logic in the method we can enlarge our theoretical, though not necessarily our practical, resources for obtaining theorems of set theory. We know from Gödel's incompleteness theorem that there are truths of arithmetic which are not theorems of Z or ZF, but which we have just seen are informal theorems of S. We also see that the formal and informal versions of S do not share the same theorems; if we are to think of S as a single system, this must be understood in terms of the sets which are generable in S, not in terms of the theorems which are generable from the axioms.

The situation we now find ourselves in suggests the following interesting challenge: are there informal consequences of Axiom 1 which are not confined to arithmetic, which are not translatable into provable sentences of standard formal set theories, and which we would regard as new results of genuine mathematical interest if we could find them? Meeting this challenge would be a good way of establishing the legitimacy of informal set theory as an adjunct to standard formal set theory. An interesting counterchallenge from the other side would be to show that no such consequences exist.

7. Other Axioms

The fact that most of the basic axioms of set theory can be derived from only one or two axioms may be seen by some readers as at best an interesting curiosity in the absence of any garnering of significant new theorems. So we must argue that in the absence of new theorems there are other reasons for being interested in the system S. Much of the argument will hinge on what we take to be an important distinction between "basic" and "higher" axioms of set theory, our interest here being confined to the basic axioms. In particular, it will hinge on whether the Axiom of Replacement is best regarded as a basic or a higher axiom or something in between.

In his introduction to Bernays ([1], p. 23), Fraenkel points out that the Axiom of Replacement yields the Axiom of Separation, and goes on to remark, "Yet it would not be practical to drop V, for I – VII are sufficient for the bulk of set theory while VIII is required for certain purposes only" (where I – VII are the axioms of Z, V is Separation, and VIII is Replacement). It is certainly no longer true that the axioms of Z are sufficient for the bulk of set theory, but Fraenkel's remark can be updated. By now set theory has taken on a life of its own as a separately developed branch of higher mathematics; but much of the rest of mathematics continues to require only the axioms of which Fraenkel was speaking, and it is still legitimate to consider the question of the role of Replacement among these axioms. In doing so we may now ask for what purposes the bulk of mathematics (as distinguished from the bulk of set theory) requires the Axiom of Replacement. After a previous remark ([1], p. 22) to the effect that Z is inadequate for defining the set $\{N, PN, PPN, \ldots\}$, Fraenkel says, "To fill this gap in a way which yields particular sets such as just mentioned as well as the general theory of ordinal numbers and of transfinite induction, an axiom of the following type is required: [Replacement]."

The general theory of cardinal and ordinal numbers and of transfinite induction is not amenable to treatment in S. ZF is adequate for most purposes, and for most purposes for which Z is inadequate Replacement would be required. Not much of a role, then, is left for S, practically speaking. Theoretically speaking, however, it is conceivable that linkage and Axiom 1 could come to be regarded as no less basic than the concepts and axioms of Z—a view supported by the fact that all the axioms of Z are derivable from Axiom 1. In this case S could come to share the role that Fraenkel envisaged for Z, reserving Replacement for special purposes only. This point of view would be enhanced if it were felt that there is anything problematic about Replacement.

In three-quarters of a century since Replacement was introduced into set theory by Fraenkel and Skolem its case has been defended many times, most frequently by the claim that in all those years no inconsistency of Replacement with the other axioms of set theory has ever arisen. This is at best a strong empirical argument, not the sort of sound logical demonstration which would traditionally be demanded for such a central article of mathematics. To be sure Z, without Replacement, enjoys no sound demonstration of consistency either, but unlike Replacement the axioms of Z have not generated the same level or frequency of defense, or defensiveness, suggesting some uneasiness on the part of the defenders, or at least the recognition that Replacement is not altogether unproblematic compared to other axioms.

The main feature of Replacement which distinguishes it from the axioms of Z is that it provides for the generation of sets on a much higher level of the Cantorian hierarchy where the sets in question are less readily conceptualized and less basic. On the basis of the celebrated doctrine of limitation of size it can be argued that Replacement is an appropriate axiom because the size of any set it generates is limited to the size of the generating set. In the sense of cardinal size this is obviously true, but attempts to justify other axioms under the limitation of size doctrine have suggested different ways other than simple cardinality by which size is to be limited. In particular, the Powerset Axiom comes under limitation of size because the cardinality of a set it generates is greater than but still limited by the cardinality of the generating set. So Powerset and Replacement both come under the limitation of size doctrine, each in its particular way. But these two axioms working in combination with each other are what gives rise to the higher level sets of ZF which Fraenkel and others contrast with the basic sets of Z.

A full discussion of the many facets of limitation of size is given in Hallett [5]. Two of these facets fall in well with Axiom 1, which gives (1) a precisely defined manner by which sets are generated from previously generated sets, and (2) a precisely defined limit to the size (cardinality) of the sets generated.

More significant than limitation of size as a criterion for estimating the appropriateness of adopting Replacement as a basic axiom is the question of its connection with the cumulative hierarchical notion of a set. In [2] Boolos argues that the comprehension axioms of Z are all rooted in this notion, whereas Replacement is not. He shows how each of the pertinent axioms of Z can be motivated by what he calls the iterative concept of a set, unlike Replacement, for which no such treatment can be given.

Any suggestion that such concerns as these should be regarded as sufficient to hamper the free and liberal use of Replacement in applications which require it will, of course, be properly rejected. Instead, our suggestion is that Replacement be excluded from the inner circle of "basic" axioms and put on the same basis that is widely adopted for the Axiom of Choice: call attention to its use on specific occasions and use it only when necessity or convenience dictates. It is doubtful that Replacement should be regarded as less problematic than Choice, though problematic in different ways.

Another concept and another axiom which are certainly as basic as linkage and Axiom 1 are the wellfoundedness concept and the Axiom of Foundation. Perhaps because the Axiom of Foundation requires only infrequent practical use, it tends to be treated as peripheral to the other axioms of Z or ZF. But when descendancy has a prominent role, as in S, it would certainly be appropriate to include Foundation as another axiom in the system. Consider, for example, that as things stand with S it is not possible even to prove the sentence $\forall n(Nn \to 0D'n)$. This particular anomaly is readily avoided in the informal version of S, and in the formal version we might have opted for a less simple definition of number to get around this problem. But the best way to make things work out smoothly in general would simply be to adopt the Axiom of Foundation, which we now recommend unreservedly. Other axioms would also work. Two such are $\forall x 0 D x$ and $\forall x 0 D' x$. Besides following the tradition of treating Foundation as peripheral, the main reason we have refrained from adopting this or some other axiom to guarantee wellfoundedness for S was to demonstrate that most of mathematics can be derived from a single axiom. Another reason is that we were able to show in Section 2, without resorting to Foundation, that all sets definable in S are indeed wellfounded, using a meta-argument instead of a proof within S itself.

On the subject of other axioms, we noted earlier that the Axiom of Choice is not derivable either from Axiom 1 or Axiom 1′. So it would perhaps be preferable to follow tradition in treating Choice as a separate axiom of set theory, rather than trying to smuggle it in as a logical principle as we did in Section 3. Here our main reason for breaking rather than going along with the tradition was again to avoid resorting to more than a single axiom.

8. Concluding Remarks

Up to this point I have attempted to remain neutral when entering controversial territory. In this final section I will come out and express some preferences of my own. I guess my feeling here is that a person who has succeeded in reducing most of mathematics to a single axiom deserves some respect, and that other people may be interested in his opinions.

I have started out from the premise that Cantorian notions, though problematic, are inseparable in the twentieth century from mathematics itself. Convinced that twentieth-century mathematics is legitimate, I am equally convinced that basic Cantorian notions will ultimately be found legitimate as well, when sufficiently clarified. Meanwhile, we must continue to live with the uneasiness and frustration which arise from a lack of sufficient clarity in our notions. If Cantorian notions are not clear enough to justify confidence in our understanding of the model universe U of the system S, however, then it seems to me that much of the contemporary mathematical enterprise is itself at stake. I take it that all this mathematics is here to stay, is rooted in the legitimate Cantorian, and will continue to flourish. Most of this flourishing can occur by taking approach 1 expressed in the Introduction.

Some people basically in sympathy with the point of view just expressed might nonetheless be interested in pushing a little deeper into the foundations of their set notions. Here, at a minimum, the notion of wellfoundedness and a cumulative hierarchy of sets will come into play. The system S falls right in with these notions. S will generally be regarded as stopping far short of constituting the entire hierarchy, but it bites off a sufficient piece of it for doing much mathematics, and for expressing basic hierarchical notions. This point of view is clearly compatible with all of the approaches 2 – 4 of the Introduction.

When we come to approach 5 of the Introduction things change dramatically, but S still falls in well enough, at least in its formal versions based on Axiom 1′ or on Axioms 2 and 3. A striking feature of this approach is that it invites us to abandon basic notions of Cantorian truth and to entertain either CH or $-$CH, say, as appropriate axioms. My own attitude here would be that it is certainly reasonable to go ahead and assume either one so long as we don't know which is true, to see where this leads. But I would hold that one or the other *is* true. Finding out which was a high priority on Hilbert's celebrated list of turn-of-the-twentieth-century challenges, and it is a disappointment to me that I have not heard of its being on any list of new turn-of-the-twenty-first-century challenges. Apparently Cohen agrees with me here; in [3], p. 151, he remarks, "A point of view which the author feels may eventually come to be accepted is that CH is *obviously* false."

Cohen, of course, gives no formal proof or informal demonstration of the falsity of CH. What he offers instead is important insight into how it can seem unreasonable to think CH would be true, and he says, "Perhaps later generations will see the problem more clearly. . . ." In my opinion it is unfortunate that a growing trend seems to regard the continuum question as having been settled by the combination of Gödel's consistency result and Cohen's independence result: CH can't be proved or disproved, it is just another "parallel postulate" for establishing the legitimacy of "non-Euclidean set theory." But in my opinion, shared I presume by Cohen, later generations will regard this as a backward step, looking to apply a nineteenth-century idea inappropriately to a twentieth-century problem. I emphasize the continuum problem because its principal locus is in basic rather than in higher set theory, and what seems demanded are new basic ways of thinking and/or new basic axioms.

One such basic way of thinking, somewhat new, is to incorporate more of the semantic and arithmetic approaches of standard mathematical reasoning into set theory which by now has become highly formal. I have no illusions that the meager foundation of an informal method which I have outlined in Section 6 will go any distance at all toward a resolution of the continuum problem, for example, but I think it is a gesture in the right direction.

Unlike most readers of this journal and this paper, who are likely to prefer formal over informal methods, my own preference is for the informal. I believe Axiom 2 works very well but I also feel that Axiom 2 by itself is not too interesting and that it gets most of its motivation from the informal Axiom 1. I feel that Axiom $1'$ is about as direct and elegant a formalization of my basic idea as anyone is likely to come up with. Still, to me, Axiom 1 wins hands down over Axiom $1'$ in elegance and perspicuity. And I share the fairly commonly held view that a genuine axiom should be expressible in primitive notation, or at least more nearly so than is the case with Axiom $1'$.

I also believe that informal definitions of basic concepts are often preferable to their formal counterparts. As an example, consider the formal $x\,D'\,y$, needed as a starting point from which to arrive eventually at a formal definition of $x\,D^n\,y$. Informally we go directly to $x\,D^n\,y$, either by way of numbered quantifiers or by way of a simple definition by induction. $x\,D'\,y$ and $x\,Dy$ would commonly be regarded as expressing the same concept (x is a descendant of y) but I don't believe that they do. The formal approach fails to define the descendancy relation D' (or the linkage relation L') in a way which would be appropriate for all domains. Consider the domain $\{0, 1\}$, where $0 \in 1, 1 \notin 0, 0 \notin 0, 1 \notin 1$. From (3) we get $-(1D0)$ while from (9) we get the anomaly $1D'0$. Thus we see that the two definitions are not logically equivalent and, I would say, do not define the same concept, though they do define relations which are extensionally equivalent on the universe U which constitutes the model for our system S. However, besides exploring the consequences of Axiom 1, one of our central goals has been to introduce a new concept, linkage, and familiarize ourselves with its basic meaning. It is not intended that this meaning should be confined to particular domains such as U. The intended meaning is adequately expressed by means of the informal definition (3), inadequately by means of the formal definition (9).

In general, I feel that formal definitions tend to work "properly" only in combination with other formal machinery, as is the case for $x\,D'y$, which only works when the existence of enough sets is assured by considerations which go beyond the immediate concept in question. I feel that in such cases the definitions do indeed work well, but not quite well enough.

I believe the two most original ideas I have presented in this paper are the notion of linkage and my proposal for a novel way of using arithmetic in set theory. Both are rooted in familiar notions: wellfoundedness and inductive definitions, respectively. I realize, somewhat painfully, that very few ramifications of the linkage notion have been explored here—only the barest foundation for such an exploration has been laid. I hope that from this foundation, however, others will be motivated to go further. I myself am too old now to have much interest in doing so. I also realize that my somewhat novel proposal for using arithmetic in set theory is not even as much as half-baked in the meager foundation of it which is presented in Section 6. I hope that, here as well, others will be motivated to proceed where I myself am not likely to go. Finally, I would like to say that I have derived much pleasure and satisfaction from thinking long and hard about the ideas I have presented here, such as they are, and I hope that others may do so as well.

Notes

1. It is expected that some readers will be unfamiliar with the second-order formal language we will be dealing with and whose vocabulary we also use as informal shorthand for the following common expressions: $-$ (not), $\wedge$ (and), $\vee$ (or), $\rightarrow$ (if…then), $\leftrightarrow$ (if and only if), $\forall$ (for all), $\exists$ (for some), $=$ (equals). We also employ three styles of variables: $x, y, \ldots,$ ranging over sets, $m, n, \ldots$ ranging over non-negative integers, and $F, G, \ldots$ ranging over properties and relations. A property or relation is expressed by an open sentence containing free (unbound, nondummy) occurrences of one or of two variables, respectively. Examples: $-\exists y(y \in x), \forall z(z \in x \rightarrow z \in y)$ (x is empty, x is a subset of y). The vocabulary of first-order logic is the same except that the variables $F, G, \ldots$ are omitted.

2. $$\exists n \forall x_0, \ldots, x_n (x_n \in \cdots \in x_0 = x \rightarrow x_n D y) \Leftrightarrow$$
$$\exists n \forall z \forall x_0, \ldots, x_n (z = x_n \wedge x_n \in \cdots \in x_0 = x \rightarrow z D y) \Leftrightarrow$$
$$\exists n \forall z (\exists x_0, \ldots, x_n (z = x_n \in \cdots \in x_0 = x) \rightarrow z D y)$$

3. From the time of Frege and Russell numbers have often been defined in terms of an ancestral relation according to which the number 5, for example, would be a fifth-generation descendant of the number 0. In set theory, however, we think of an ascending rather than a descending hierarchy of sets, and of descending rather than ascending $\in$-chains. So we are confronted with two familiar but conflicting metaphors: one in which numbers and

sets are generated from (are descendants of) zero or the empty set, and the other in which numbers and sets rise (ascend) from zero or the empty set to higher levels. While one or the other of these metaphors will carry more weight with different readers, it seems to me on balance that the concerns of this paper are better served by going with the metaphor of sets as rising rather than descending from the empty set and thinking of x as a descendant of y when x lies on an $\in$-chain descending from y down to x. Less confusion is likely to arise from saying that x_2 is a second-generation descendant of x_0 on the descending chain $\cdots x_2 \in x_1 \in x_0$ than if we put it the other way around.

4. Demonstrations of (6)–(8): If x links with y and some infinite chain c descends from x, then there exists an n and a d such that d is the nth-generation descendant of x on c, d is a descendant of y, and some infinite $\in$-chain descends from y (thus (6)). Now let Axiom 1 define y relative to z, and let x be any element of y. Then xLz (part of (7)). So if no infinite $\in$-chain descends from z none descends from x, so none descends from y (thus (8)). Now consider the set y for which $\forall x(x \in y \leftrightarrow xL0)$, where 0 is the empty set zero, defined by taking $x \neq x$ for Fxz in Axiom 1. The sets $0, \{0\}, \{\{0\}\}, \ldots$ all belong to y. So no n exists such that every nth-generation descendant of y is a descendant of zero (i.e., is zero). So $-yL0$ (the other part of (7)).

5. The Mirimanoff paradox: the would-be set of all wellfounded sets would be wellfounded, thus a member of itself, thus not wellfounded. This paradox cannot arise from Axiom 1 which provides no way of defining a set of all wellfounded sets. Adapting the paradox to linkage, assume Axiom $1_?$ and consider the set y for which $\forall x(x \in y \leftrightarrow xL_?0)$. Then $yL_?0$, $y \in y$, 0 belongs to the chain $\cdots y \in y \in y$, $0 = y$, $y \notin y$, $y \in y \wedge y \notin y$. But with Axiom 1 and the set y for which $\forall x(x \in y \leftrightarrow xL0)$ this argument never leaves the starting gate; instead we have $-yL0$, as we saw in note 4.

6. "There exists a function F such that, for every nonempty set x, $F'x \in x$."

7. In the case of Powerset, for example, we must show that $\forall w(w \in x \rightarrow w \in z)$ implies $\forall w(w \in x \rightarrow wD'z)$, an easy exercise. That Powerset follows from Axiom 2 is also immediately evident from its formulation as the expression "$\{x \subseteq z | x \subseteq z\}$ exists".

8. A simple alternative to our perhaps too ready dismissal of the Pair axiom is to replace Axiom 2 by the following:

$$\text{Axiom } 2' = \forall F \forall z_1 \forall z_2 \exists_1 y \forall x(x \in y \leftrightarrow Fxz_1z_2 \wedge \forall w(w \in x \rightarrow wDz_1 \vee wDz_2)).$$

To get Pair from Axiom $2'$ take Fxz_1z_2 to be $x = z_1 \vee x = z_2$ which implies $\forall w(w \in x \rightarrow wDz_1 \vee wDz_2)$. Axiom 2 follows from Axiom $2'$ by instantiating both variables z_1 and z_2 to z.

9. The alleged validity of (12) in second-order logic is controversial. What renders (12) valid or invalid is whether, more generally, one accepts or rejects Choice as a valid principle of logic. If it is rejected then the Axiom of Choice may as well be rejected from set theory too, which is simply to say that if we intend to provide a basis for Choice in S we may as well do as we have done, relegating Choice to the underlying logic.

10. The Hilbert operator εx attaches itself to a sentence Fx containing free occurrences of x to produce a term εxFx which is closely analogous to the more familiar term ιxFx. ιxFx is associated with the uniqueness quantifier $\exists_1 xFx$, and ιxFx designates the unique x such that Fx whenever $\exists_1 xFx$ holds. Similarly, εxFx is associated with the quantifier

$\exists x Fx$, and $\varepsilon x Fx$ designates the chosen x such that Fx whenever $\exists x Fx$ holds. In case either $\exists_1 x Fx$ or $\exists x Fx$ does not hold, the designation of $\iota x Fx$ or $\varepsilon x Fx$ may be determined in some arbitrary manner or may simply remain undetermined. Hilbert discovered that the entirety of first-order logic can be built up from a single logical schema of remarkable simplicity: $Fx \to F(\varepsilon x Fx)$. For example, define existential quantification thus: $\exists x Fx \equiv_{\text{df}} F(\varepsilon x Fx)$ (there exists an x satisfying F if and only if the chosen x satisfies F). Then the schema $Fx \to F(\varepsilon x Fx)$ takes on the more familiar form $Fx \to \exists x Fx$, and (14) takes on the valid form

$$\forall y((\varepsilon x\ x \in y) \in y \to (\varepsilon x\ x \in y) \in y).$$

11. Whether the Axiom of Choice should be included in this claim is dubious, given that (9) and Axiom 2 have nothing to do with its derivation. Whether Choice should even have been included among the other axioms of Z is also dubious; Zermelo set theory with Choice included is nowadays often referred to as ZC rather than simply Z. Because we want to interest the reader in a system of set theory based on a single axiom which is powerful enough for deriving most of contemporary mathematics, we have elected to include Choice in our considerations.

12. Strictly speaking this treatment for N should not be regarded as giving a correct definition of a number; it would do so only where infinite descending $\in$-chains are ruled out as, for example, in the presence of the Axiom of Foundation. We have given a meta-argument above to show that such chains will never arise in our system, but this cannot be shown by a formal argument within the system itself. It would be easy to give a correct definition in place of this incorrect one, but this one works particularly smoothly for giving all the axioms of Z, after which, if one wishes, a more usual definition of a number can be given to supersede this one. Or, as we propose later on, one can simply adjoin the Axiom of Foundation to Axiom 1 and stay with the treatment given.

13. Proof of Lemma 1: Assume $ySx \wedge nSm \wedge zD^n y$. Then $Nn \wedge \exists F(\forall u \forall v \forall w(Fuv \wedge Fuw \to v = w) \wedge \exists u(Eu \wedge Fuz) \wedge Fny \wedge \forall i \forall j(jD'n \wedge jSi \to \exists u \exists v(Fiu \wedge Fjv \wedge u \in v)))$. First, $Nn, \forall j(jD'n \to Ej \vee \exists i\ jSi), jD'm \to jD'n$, so Nm. Next, $nD'n \wedge nSm$, so $\exists u \exists v(Fmu \wedge Fnv \wedge u \in v)$. So $u \in y, u = x,$ & Fmx. Finally, $jD'm \to jD'n$, so $\forall i \forall j(jD'm \wedge jSi \to \exists u \exists v(Fiu \wedge Fjv \wedge u \in v)))$. So $zD^m x$. □

14. Proof of Lemma 2: $\exists_1 n \forall x(x \in n \leftrightarrow x \neq x \wedge xLz), \exists_1 n En, \forall j(jD'n \to Ej \vee \exists i\ jSi)$, so Nn. Next, let m be the unique n such that $Nn \wedge En$ and assume $zD^m x$. Then $\exists F(\forall u \forall v \forall w(Fuv \wedge Fuw \to v = w) \wedge Fmz \wedge Fmx), z = x, zD'x$. So $zD^m x \to zD'x$ and xLx. Finally, assume Nm. Then $\exists_1 n \forall x(x \in n \leftrightarrow x = m \wedge xLm), x = m \to xLm, \exists_1 n\ nSm, \forall j(jD'm \to Ej \vee \exists i\ jSi), jD'n \to j = n \vee jD'm, j = n \to Ej \vee \exists i\ jSi, jD'n \to Ej \vee \exists i\ jSi$, so Nn. □

15. Proof of Lemma 3: Assume $xD^1 y$. Then $\exists F(\forall u \forall v \forall w(Fuv \wedge Fuw \to v = w) \wedge F\,0x \wedge F\,1y \wedge \forall i \forall j(jD'1 \wedge jSi \to \exists u \exists v(Fiu \wedge Fjv \wedge u \in v))), 1D'1 \wedge 1S0, \exists u \exists v(F\,0u \wedge F\,1v \wedge u \in v), u \in y, x = u$ and $x \in y$. □

16. More particularly, every set y which can be generated by n applications of Axiom 1 will belong to V_{2n}. For $n = 1$, the only set which can be generated by Axiom 1 is the empty set 0, for which we have in turn $0L0, 0 \in V_1, 0LV_1, 0 \in V_2$. For $n > 1$ suppose, for the induction, that all sets which can be generated by n applications of Axiom 1 are members of V_{2n}. Let y be any set which can be generated by $n+1$ applications of Axiom 1, and let x be any element of y. Then for some z in V_{2n} and some F we have: $F,\ xLz,\ xLV_{2n},\ x \in V_{2n+1},\ \forall x(x \in y \to x \in V_{2n+1}),\ yLV_{2n+1},\ y \in V_{2n+2}$.

17. $x \in U_1 \Leftrightarrow x \in 0 \vee x \in P0 \vee x \in PP0 \vee \cdots \Leftrightarrow xL0 \Leftrightarrow x \in V_1$. Similarly, $x \in U_{m+1} \Leftrightarrow x \in U_m \vee x \in PU_m \vee x \in PPU_m \vee \cdots \Leftrightarrow xLU_m \Leftrightarrow xLV_m \Leftrightarrow x \in V_{m+1}$.

18. Many people object to using arithmetic notions in establishing the foundations of set theory on grounds that set theory is the proper basis for establishing arithmetic. Probably just as many people object to the idea that arithmetic has or needs its basis in set theory, which they see as problematic where arithmetic is not. Other people agree, on one hand, that arithmetic can certainly stand on its own, and on the other hand, that a proper foundation for set theory must also stand on its own without borrowing support from arithmetic. Such preferences are often strongly held and may produce some resistance to the way we are proposing to incorporate arithmetic into our informal proof procedures.

References

[1] Bernays, P., *Axiomatic Set Theory*, North-Holland Publishing Co., Amsterdam, 1958. Zbl 0082.26301. MR 21:4912. 163

[2] Boolos, G., "The iterative conception of set," *Journal of Philosophy*, vol. 68 (1971), pp. 215–31. 164

[3] Cohen, P., *Set Theory and the Continuum Hypothesis*, W. A. Benjamin, Inc., New York, 1966. Zbl 0182.01301. MR 38:999. 166

[4] Gödel, K., "What is Cantor's continuum problem?," *American Mathematical Monthly*, vol. 54 (1947), pp. 515–25. Zbl 0038.03003. MR 9,403h. 152

[5] Hallett, M., *Cantorian Set Theory and Limitation of Size*, Oxford University Press, New York, 1984. Zbl 0656.03030. MR 86e:03003. 164

[6] Hilbert, D., and P. Bernays, *Grundlagen der Mathematik. Volume 2*, Springer, Berlin, 1939. Zbl 0020.19301. MR 42:7477. 157

Acknowledgments

Many helpful suggestions and generous contributions of time and encouragement offered by Clifton McIntosh and Elliott Mendelson are much appreciated. I am especially indebted to the anonymous referee of this journal who has rescued me from several crucial errors and provided a large number of penetrating and useful comments.

Department of Philosophy
University of Utah
1601 Sunnydale Lane
Salt Lake City UT 84108
benbon@uofu.net