The Bulletin of Symbolic Logic
Volume 13, Number 3, Sept. 2007

RELATING FIRST-ORDER SET THEORIES AND ELEMENTARY TOPOSES

STEVE AWODEY, CARSTEN BUTZ, ALEX SIMPSON, AND THOMAS STREICHER

Abstract. We show how to interpret the language of first-order set theory in an elementary topos endowed with, as extra structure, a directed structural system of inclusions (dssi). As our main result, we obtain a complete axiomatization of the intuitionistic set theory validated by all such interpretations. Since every elementary topos is equivalent to one carrying a dssi, we thus obtain a first-order set theory whose associated categories of sets are exactly the elementary toposes. In addition, we show that the full axiom of Separation is validated whenever the dssi is superdirected. This gives a uniform explanation for the known facts that cocomplete and realizability toposes provide models for Intuitionistic Zermelo–Fraenkel set theory (IZF).

§1. Introduction. The notion of elementary topos abstracts from the structure of the category of sets. The abstraction is sufficiently general that elementary toposes encompass a rich collection of other very different categories, including categories that have arisen in fields as diverse as algebraic geometry, algebraic topology, mathematical logic, and combinatorics. Nonetheless, elementary toposes retain many of the essential features of the category of sets. In particular, elementary toposes possess an *internal logic*, which is a form of higher-order type theory, see e.g., [11, 13, 9]. This logic allows one to reason with objects of the topos as if they were *abstract sets* in the sense of [12]; that is, as if they were unstructured collections of elements. Although the reasoning supported in this way is both powerful and natural, it differs in several respects from the set-theoretic reasoning available in the familiar first-order set theories, such as Zermelo–Fraenkel set theory (ZF).

A first main difference between the internal logic and ZF is:

1. Except in the special case of boolean toposes, the underlying internal logic of a topos is intuitionistic rather than classical.

Many toposes of mathematical interest are not boolean. Thus the use of intuitionistic logic is unavoidable. Moreover, as fields such as *synthetic differential geometry* and *synthetic domain theory* demonstrate, the non-validity of classical logic has mathematical applications. In these areas, intuitionistic logic offers the opportunity of working consistently with convenient but

Received November 20, 2006.

1079-8986/07/1303-0002/$2.90

classically inconsistent properties, such as the existence of nilpotent infinitesimals, or the existence of nontrivial sets over which every endofunction has a fixed point.

Although the intuitionistic internal logic of toposes is a powerful tool, there are potential applications of set-theoretic reasoning in toposes for which it is too restrictive. This is due to a second main difference between the internal logic and first-order set theories.

2. In first-order set theories, one can quantify over the class of all sets, whereas, in the internal logic of a topos, every quantifier is bounded by an object of a topos, i.e., by a set.

Sometimes, one would like to reason about mathematical structures derived from the topos that are not "small", and so cannot be considered internally at all. For example, one often considers derived categories (e.g., the category of internal *locales*) that are not themselves small categories from the viewpoint of the topos. The standard mathematical approach to handling non-small categories relative to a topos is to invoke the machinery of fibrations (or the essentially equivalent machinery of indexed categories). This paper provides the basis for an alternative approach. We show how to conservatively extend the internal logic of a topos to explicitly permit direct set-theoretic reasoning about non-small structures. To achieve this, we directly address issue 2 above, by embedding the internal logic in a first-order set theory within which one can quantify over any class, including the class of all sets (i.e., the class of all objects of the topos). In general, this extended logic should provide a useful tool for establishing properties of non-small structures (e.g., large categories), relative to a topos, using natural set-theoretic arguments. In fact, one such application of our work has already appeared in [18].

In Section 2, we present the set theory that we shall interpret over an arbitrary elementary topos (with natural numbers object), which we call Basic Intuitionistic Set Theory (BIST). Although very natural, and based on familiar looking set-theoretic axioms, there are several differences compared with standard formulations of intuitionistic set theories. Two of the differences are minor: in BIST the universe may contain non-sets (a.k.a. atoms or urelements) as well as sets, and non-well-founded sets are permitted (though not obliged to exist). The essential difference is the following.

3. BIST is a conservative extension of intuitionistic higher-order arithmetic (HAH). In particular, by Gödel's second incompleteness theorem, it cannot prove the consistency of HAH.

This property is unavoidable because we wish to faithfully embed the internal logic of the free topos (with natural numbers object) in BIST, and this logic is exactly HAH.

Property 3 means that BIST is necessarily proof-theoretically weaker than ZF. That such weakness is necessary for interpreting first-order set theory in toposes has long been recognised. The traditional account has been that the appropriate set theory is *bounded Zermelo* (*bZ*) set theory (also known as Mac Lane set theory [14]), which is ZF set theory with the axiom of Replacement removed and with Separation restricted to bounded (i.e., Δ_0) formulas. The standard results connecting bZ set theory with toposes run as follows. First, from any (ordinary first-order) model of bZ one can construct a well-pointed (hence boolean) topos whose objects are the elements of the model and whose internal logic expresses truth in the model. Conversely, given any well-pointed topos $\mathcal{E}$, certain "transitive objects" can be identified, out of which a model of bZ can be constructed. This model captures that part of the internal logic of $\mathcal{E}$ that pertains to transitive objects. See, e.g., [13] for an account of this correspondence.

This standard story is unsatisfactory in several respects. First, it applies only to well-pointed (hence boolean) toposes. Second, the set theory is only able to express properties of transitive objects in $\mathcal{E}$, potentially ignoring whole swathes of the topos. Third, with the absence of Replacement, bZ is not a particularly convenient or natural set theory to reason in, see [14] for a critique.

The goal of this announcement is to establish a more satisfactory connection between elementary toposes and the theory BIST. Unlike bZ, one of the attractive features of BIST is that it contains the full axiom of Replacement. We shall show that BIST can be interpreted over an arbitrary elementary topos (with natural numbers object) in such a way that the class of all sets in the set theory can be understood as being exactly the collection of all objects of the topos. In fact, not only does our interpretation model full Replacement, but it even validates the (intuitionistically stronger) axiom of (strong) Collection (Coll).

Some readers familiar with classical set theory may be feeling uncomfortable at this point. In bZ set theory, it is the absence of Replacement and the restriction of Separation that weakens the proof-theoretic strength of the set theory so that it is compatible with the internal logic of elementary toposes. BIST, however, has full Replacement. In classical set theory, Replacement, which is equivalent to Collection and implies full Separation, takes one beyond the the proof-theoretic strength of elementary toposes. The situation is completely different under intuitionistic logic. Intuitionistically, as has long been known, the full axioms of Replacement and Collection *are* compatible with proof-theoretically weak set theories, see, e.g., [15, 7, 1, 2]. Readers who are unfamiliar with this phenomenon can find examples illustrating the situation in the discussion at the end of Section 4.

The precise connection between BIST and elementary toposes is elaborated in Sections 3 and 4. In order to interpret unbounded quantification over the class of all sets, we have to address a fourth difference between the internal logic of toposes and first-order set theories.

4. In first-order set theories, one can compare the elements of different sets for equality, whereas, in the internal logic of a topos, one can only compare elements of the same object.

In Section 3, we consider additional structure on an elementary topos that enables the comparison of (generalized) elements of different objects. This additional structure, a *directed structural system of inclusions* (*dssi*), directly implements a well-behaved notion of subset relation between objects of a topos. Although not particularly natural from a category-theoretic point of view, the structure of a dssi turns out to be exactly what is needed to obtain an interpretation of the full language of first-order set theory in a topos, including unbounded quantification; and thus indeed resolves issue 2 above. We present this interpretation in Section 4, using a suitably defined notion of "forcing" over a dssi.

In fact, a special case of our forcing semantics for first-order set theory in toposes was first presented by Hayashi in [8], where the notion of inclusion was provided by the canonical notion of inclusion map between the transitive objects in a topos. One benefit of our more general axiomatic notion of dssi is that our logic is able to express properties of arbitrary (non-transitive) objects of the topos. More substantively, we extend Hayashi's account in three significant ways. First, as mentioned above, we show that, for any elementary topos, the forcing semantics always validates the full axiom of Collection (and hence Replacement). Thus we obtain a model of BIST plus Collection (henceforth BIST+ Coll), which is a very natural set theory in its own right. Second, we give correct conditions under which the full axiom of Separation is modelled (BIST itself supports only a restricted separation principle). Third, we obtain a completeness result (Theorem 4.2) which shows that the theory BIST+Coll axiomatizes exactly the set-theoretic properties validated by our forcing semantics. This theorem, whose proof is by no means routine, constitutes the major technical contribution of the present work. It also fulfils a longstanding wish of Saunders Mac Lane, who often expressed the desire to find a first-order set theory whose notion of set corresponds to that given by elementary toposes.

In mathematical applications of topos theory, one is often interested in "real world" toposes, such as Grothendieck and realizability toposes, defined from the "external" category of sets (which we take to be axiomatized by ZFC). It is known from previous work [6, 8, 10] that such toposes are capable of interpreting Friedman's IZF set theory, which is proof-theoretically as strong as ZF. Thus, if one is primarily interested in such real world toposes, the above account is unsatisfactory in merely detailing how to interpret the weak set theory BIST+ Coll inside them.

To address this, in parallel with the development already described, we further show how the approach discussed above adapts to model the full Separation axiom (Sep) in toposes such as cocomplete and realizability toposes. The appropriate structure we require for this task is a modification of the notion of dssi from Section 3, obtained by strengthening the directedness property to require upper bounds for arbitrary (rather than just finite) sets of objects. Given a topos with such a *superdirected structural system of inclusions* (*sdssi*), the forcing interpretation of Section 4 does indeed model the full Separation axiom. Since cocomplete toposes and realizability toposes can all be endowed with sdssi's, we thus obtain a uniform explanation of why all such toposes model IZF (the set theory BIST+ Coll+ Sep is intertranslatable with IZF). It seems that no such uniform explanation was known before.

This article is an announcement of results taken from a forthcoming paper [4], where proofs for all the results stated here can be found. That paper contains, in addition, another major component not discussed here. A second class of category-theoretic models of BIST is considered, based on the idea of axiomatizing the category-theoretic structure of the category of classes of BIST, following the lead of Joyal and Moerdijk's *Algebraic Set Theory* (*AST*) [10] and its subsequent refinements in [17, 5]. The details of the relevant category-theoretic models have been surveyed in separate articles [19, 3], and are thus not included here. Nevertheless, the class-category semantics of BIST discussed in [19, 3] is intimately connected with the forcing semantics, and, moreover, is used as a crucial element in the proof of Theorem 4.2 below [4]. We remark that there have been several further contributions to AST since the results detailed in this announcement were first obtained, many building on the approach of [4]. We refer the interested reader to the Algebraic Set Theory website: `http://www.cmu.edu/mobius/ast/`.

§2. Basic Intuitionistic Set Theory (BIST) and extensions. All first-order set theories considered in this announcement are built on top of a basic theory, BIST (Basic Intuitionistic Set Theory). The axiomatization of BIST is primarily motivated by the desire to find the most natural first-order set theory under which an arbitrary elementary topos may be considered as a category of sets. Nonetheless, BIST is also well motivated as a set theory capturing basic principles of set-theoretic reasoning in informal mathematics.

The axioms of BIST axiomatize properties of the intuitive idea of a mathematical universe consisting of mathematical "objects". The universe gives rise to notions of "class" and of "set". *Classes* are arbitrary collections of mathematical objects; whereas *sets* are collections that are, in some sense, small. The important feature of sets is that they themselves constitute mathematical objects belonging to the universe. The axioms of BIST simply require that the collection of sets be closed under various useful operations on sets, all familiar from mathematical practice. Moreover, in keeping with informal mathematical practice, we do not assume that the only mathematical objects in existence are sets.

Membership	$y \in x \to \mathsf{S}(x)$
Extensionality	$\mathsf{S}(x) \wedge \mathsf{S}(y) \wedge (\forall z.\, z \in x \leftrightarrow z \in y) \to x = y$
Indexed-Union	$\mathsf{S}(x) \wedge (\forall y \in x.\, \text{Ƨ}z.\, \phi) \to \text{Ƨ}z.\, \exists y \in x.\, \phi$
Emptyset	$\text{Ƨ}z.\, \bot$
Pairing	$\text{Ƨ}z.\, z = x \vee z = y$
Equality	$\text{Ƨ}z.\, z = x \wedge z = y$
Powerset	$\mathsf{S}(x) \to \text{Ƨ}y.\, y \subseteq x$

FIGURE 1. Axioms for BIST^{-}.

Coll $\quad \mathsf{S}(x) \wedge (\forall y \in x.\, \exists z.\, \phi) \to$

$$\exists w.\, (\mathsf{S}(w) \wedge (\forall y \in x. \exists z \in w.\, \phi) \wedge (\forall z \in w. \exists y \in x.\, \phi))$$

FIGURE 2. Collection axiom.

The set theory BIST is formulated as a theory in intuitionistic first-order logic with equality.[1] The language contains one unary predicate, S, and one binary predicate, $\in$. The formula $\mathsf{S}(x)$ expresses that x is a set. The binary predicate is, of course, set membership.

Figure 1 presents the axioms for BIST^{-}, which is BIST without the axiom of infinity. All axioms are implicitly universally quantified over their free variables. The axioms make use of the following notational devices. As is standard, we write $\forall x \in y.\, \phi$ and $\exists x \in y.\, \phi$ as abbreviations for the formulas $\forall x.\, (x \in y \to \phi)$ and $\exists x.\, (x \in y \wedge \phi)$ respectively, and we refer to the prefixes $\forall x \in y$ and $\exists x \in y$ as *bounded quantifiers*. In the presence of non-sets, we define the subset relation, $x \subseteq y$, as abbreviating

$$\mathsf{S}(x) \wedge \mathsf{S}(y) \wedge \forall z \in x.\, z \in y.$$

This is important in the formulation of the Powerset axiom. We also use the notation $\text{Ƨ}x.\, \phi$, which abbreviates

$$\exists y.\, (\mathsf{S}(y) \wedge \forall x.\, (x \in y \leftrightarrow \phi)),$$

where y is a variable not occurring free in ϕ. Thus $\text{Ƨ}x.\, \phi$ states that the class $\{x \mid \phi\}$ forms a set. Equivalently, Ƨ can be understood as the generalized quantifier "there are set-many".

Often we shall consider BIST^{-} together with the axiom of Collection, presented in Figure 2.[2] One reason for not including Collection as one of the axioms of BIST^{-} is that it seems better to formulate the results that do not

[1]As discussed in Section 1, the use of intuitionistic logic is essential for formulating a set theory interpretable in any elementary topos.

[2]Collection, in this form, is often called *Strong Collection*, because of the extra clause $\forall z \in w. \exists y \in x.\, \phi$, which is not present in the Collection axiom as usually formulated. The inclusion of the additional clause is necessary in set theories, like BIST^{-}, that do not have full Separation.

require Collection for a basic theory without it. Another is that Collection has a different character from the other axioms in asserting the existence of a set that is not uniquely characterized by the properties it is required to satisfy.

There are three main non-standard ingredients in the axioms of BIST$^-$. The first is the Indexed-Union axiom, which is taken from [2] (where it is called Union-Rep). In the presence of the other axioms, Indexed-Union combines the familiar axioms below,

Union	$\mathsf{S}(x) \wedge (\forall y \in x.\, \mathsf{S}(y)) \rightarrow \text{Ƨ}z.\, \exists y \in x.\, z \in y,$
Replacement	$\mathsf{S}(x) \wedge (\forall y \in x.\, \exists! z.\, \phi) \rightarrow \text{Ƨ}z.\, \exists y \in x.\, \phi,$

into one simple axiom, which is also in a form that is convenient to use. We emphasise that there is no restriction on the formulas ϕ allowed to appear in Indexed-Union. This means that BIST$^-$ supports the full Replacement schema above. The second non-standard feature of BIST$^-$ is the inclusion of an explicit Equality axiom. This is to permit the third non-standard feature, the absence of any Separation axiom. In the presence of the other axioms, including Equality and Indexed-Union (full Replacement is crucial), this turns out not to be a major weakness. As we shall see below, many instances of Separation are derivable in BIST$^-$.

First, we establish notation for working with BIST$^-$. As is standard, we make free use of derived constants and operations: writing $\emptyset$ for the emptyset, $\{x\}$ and $\{x, y\}$ for a singleton and pair respectively, and $x \cup y$ for the union of two sets x and y. We write δ_{xy} for the set $\{z \mid z = x \wedge z = y\}$. It follows from the Equality and Indexed-Union axioms that, for sets x and y, the intersection $x \cap y$ is a set, because $x \cap y = \bigcup_{z \in x} \bigcup_{w \in y} \delta_{zw}$.

We now study Separation in BIST$^-$. By an *instance of Separation*, we mean a formula of the form[3]

$$\phi[x, y]\text{-Sep} \quad \mathsf{S}(x) \rightarrow \text{Ƨ}y.\, (y \in x \wedge \phi),$$

which states that the subclass $\{y \in x \mid \phi\}$ of x is actually a sub*set* of x. We write Sep for the full Separation schema: $\phi[x, y]$-Sep for all ϕ. Although the full Sep schema is not derivable in BIST$^-$, many instances of it are. To see this, as in [2], we analyse the formulas ϕ for which the corresponding instances of Separation are derivable. For any formula ϕ, we write $!\phi$ to abbreviate the following special case of Separation

$$\text{Ƨ}z.\, (z = \emptyset \wedge \phi),$$

where z is not free in ϕ. We read $!\phi$ as stating that the property ϕ is *restricted*.[4] The utility of the concept is given by the lemma below, showing

[3] We write $\phi[x, y]$ to mean a formula ϕ with the free variables x and y (which may or may not occur in ϕ) distinguished. Moreover, once we have distinguished x and y, we write $\phi[t, u]$ for the formula $\phi[t/x, u/y]$. Note that ϕ *is* permitted to contain free variables other than x, y.

[4] The terminology "restricted" is sometimes used to refer to formulas in which all quantifiers are bounded. We instead use "bounded" for the latter syntactic condition.

that the notion of restrictedness exactly captures when a property can be used in an instance of Separation.

LEMMA 2.1. $\mathrm{BIST}^- \vdash (\forall y \in x.\, !\phi) \leftrightarrow \phi[x, y]$-Sep.

We next state important closure properties of restricted propositions.

LEMMA 2.2. *The following all hold in* BIST^-.

1. $!(x = y)$.
2. *If* $\mathsf{S}(x)$ *then* $!(y \in x)$.
3. *If* $!\phi$ *and* $!\psi$ *then* $!(\phi \wedge \psi)$, $!(\phi \vee \psi)$, $!(\phi \rightarrow \psi)$ *and* $!(\neg\phi)$.
4. *If* $\mathsf{S}(x)$ *and* $\forall y \in x.\, !\phi$ *then* $!(\forall y \in x.\, \phi)$ *and* $!(\exists y \in x.\, \phi)$.
5. *If* $\phi \vee \neg\phi$ *then* $!\phi$.

The following immediate corollary gives a useful class of instances of Separation that are derivable in BIST^-.

COROLLARY 2.3. *Suppose that* $\phi[x_1, \ldots, x_k]$ *is a formula containing no atomic subformula of the form* $\mathsf{S}(z)$ *and such that every quantifier is bounded and of the form* $\forall y \in x_i$ *or* $\exists y \in x_i$ *for some* $1 \leq i \leq k$. *Then*

$$\mathrm{BIST}^- \vdash \mathsf{S}(x_1) \wedge \cdots \wedge \mathsf{S}(x_k) \rightarrow !\phi.$$

At this point, it is convenient to develop further notation. Any formula $\phi[x]$ determines a class $\{x \mid \phi\}$, which is a set just if $\text{Ƨ}x.\, \phi$. Given a class $A = \{x \mid \phi\}$, we write $y \in A$ for $\phi[y]$, and we use relative quantifiers $\forall x \in A$ and $\exists x \in A$ in the obvious way.

Given two classes A and B, we write $A \times B$ for the product class:

$$\{p \mid \exists x \in A.\, \exists y \in B.\, p = (x, y)\},$$

where $(x, y) = \{\{x\}, \{x, y\}\}$ is the standard Kuratowski pairing construction. Using Indexed-Union, one can prove that if A and B are both sets then so is $A \times B$. Similarly, we write $A + B$ for the coproduct class

$$\{p \mid (\exists x \in A.\, p = (\{x\}, \emptyset)) \vee (\exists y \in B.\, p = (\emptyset, \{y\}))\}.$$

Given a set x, we write A^x for the class

$$\{f \mid \mathsf{S}(f) \wedge (\forall p \in f.\, p \in x \times A) \wedge (\forall y \in x.\, \exists! z.\, (y, z) \in f)\}$$

of all functions from x to A. By the Powerset axiom, if A is a set then so is A^x. We shall use standard notation for manipulating functions.

We next turn to the axiom of Infinity. As we are permitting non-sets in the universe, there is no reason to require the individual natural numbers themselves to be sets. Infinity is thus formulated as in Figure 3. Define

$$\mathrm{BIST} = \mathrm{BIST}^- + \mathrm{Inf}.$$

For the sake of comparison, we also include, in Figure 3, the familiar von Neumann axiom of Infinity, which does make assumptions about the nature of the elements of the assumed infinite set. It will follow from the results of Section 4 that:

$$\begin{array}{ll} \text{Inf} & \exists I.\, \exists 0 \in I.\, \exists s \in I^I.\, (\forall x \in I.\, s(x) \neq 0) \wedge \\ & \qquad\qquad (\forall x, y \in I.\, s(x) = s(y) \to x = y) \\ \text{vN-Inf} & \exists I.\, (\emptyset \in I \wedge \forall x \in I.\, \mathsf{S}(x) \wedge x \cup \{x\} \in I) \end{array}$$

FIGURE 3. Infinity axioms.

PROPOSITION 2.4. BIST+ Coll $\nvdash$ vN-Inf.

It is instructive to construct the set of natural numbers in BIST and to derive its induction principle. The axiom of Infinity gives us an infinite set I together with an element 0 and a function s. We define N to be the intersection of all subsets of I containing 0 and closed under s. By the Powerset axiom and Lemma 2.2, N is a set. This definition of the natural numbers determines N up to isomorphism.

There is a minor clumsiness inherent in the way we have formulated the Infinity axiom and derived the natural numbers from it. Since the infinite structure $(I, 0, s)$ is not uniquely characterized by the Infinity axiom, there is no definite description for N available in our first-order language. The best we can do is use the formula $Nat(N, 0, s)$:

$$0 \in N \wedge s \in N^N \wedge (\forall x \in N.\, s(x) \neq 0) \wedge (\forall x, y \in N.\, s(x) = s(y) \to x = y)$$
$$\wedge\, \forall X \in \mathcal{P}N.\, (0 \in X \wedge (\forall x.\, x \in X \to s(x) \in X)) \to X = N,$$

where $N, 0, s$ are variables, to assert that $(N, 0, s)$ forms a legitimate natural numbers structure. Henceforth, for convenience, we shall sometimes state that some property ψ, mentioning $N, 0, s$, is derivable in BIST. In doing so, what we really mean is that the formula

$$\forall N, 0, s.\, (Nat(N, 0, s) \to \psi)$$

is derivable in BIST. Thus, informally, we treat $N, 0, s$ as if they were constants added to the language and we treat $Nat(N, 0, s)$ as if it were an axiom. The reader may wonder why we do not simply add such constants and assume $Nat(N, 0, s)$ (instead of our axiom of Infinity) and hence avoid the fuss. Our reason for not doing so is that, in Section 4, we consider semantic models of the first-order language and we should like it to be a *property* of such models whether or not they validate the axiom of Infinity. This is the case with Infinity as we have formulated it, but would not be the case if it were formulated using additional constants, which would require extra *structure* on the models.

For a formula $\phi[x]$, the induction principle for ϕ is

$$\phi[x]\text{-Ind} \quad \phi[0] \wedge (\forall x \in N.\, \phi[x] \to \phi[s(x)]) \to \forall x \in N.\, \phi[x].$$

We write Ind for the full induction principle, ϕ-Ind for all formulas ϕ, and we write RInd for the schema of *Restricted Induction*:

$$\text{RInd} \quad (\forall x \in N.\, !\phi) \to \phi[x]\text{-Ind}.$$

DE	Decidable Equality	$x = y \vee \neg(x = y)$
REM	Restricted Excluded Middle	$(!\phi) \rightarrow (\phi \vee \neg\phi)$
LEM	Law of Excluded Middle	$\phi \vee \neg\phi$

FIGURE 4. Excluded middle axioms.

LEMMA 2.5. BIST $\vdash$ RInd.

As induction holds for restricted properties, by Lemma 2.1, we have:

COROLLARY 2.6. BIST+ Sep $\vdash$ Ind.

Figure 4 contains three other axioms that we shall consider adding to our theories. LEM is the full Law of the Excluded Middle, REM is its restriction to restricted formulas and DE (the axiom of Decidable Equality) its restriction to equalities. The latter two turn out to be equivalent.

LEMMA 2.7. *In* BIST$^-$, *axioms* DE *and* REM *are equivalent*.

Henceforth, we consider only REM. Of course, properties established for REM also hold *inter alia* for DE.

PROPOSITION 2.8. BIST$^-$+ LEM $\vdash$ Sep.

COROLLARY 2.9. BIST$^-$+ Sep+ REM = BIST$^-$+ LEM.

In the sequel, we shall show how to interpret the theories BIST+ Coll in any elementary topos with natural numbers object. Also, we shall interpret BIST+Coll+REM in any boolean topos with natural numbers object. From these results, we shall deduce

PROPOSITION 2.10. BIST+ Coll+ REM $\nvdash$ Con(HAH),

where Con(HAH) is the Π^0_1 formula asserting the consistency of Higher-order Heyting Arithmetic [20]. Indeed, this proposition is a consequence of the conservativity of our interpretation of BIST+Coll+REM over the internal logic of boolean toposes, see Proposition 4.6 and surrounding discussion. On the other hand,

PROPOSITION 2.11. BIST+ Ind $\vdash$ Con(HAH).

COROLLARY 2.12. *If any of the schemas* Ind, Sep *or* LEM *are added to* BIST *then* Con(HAH) *is derivable. Hence, none of these schemas is derivable in* BIST+ Coll+ REM.

Proposition 2.10 shows that BIST+ Coll is considerably weaker than ZF set theory. As well as BIST, we shall also be interested in the theory:

$$\text{IST} = \text{BIST} + \text{Sep},$$

introduced in [17]. The theory IST is intertranslatable with Friedman's Intuitionistic Zermelo–Fraenkel set theory, in its version IZF_R with Replacement rather than Collection, see [16]. Similarly, IST+Coll is intertranslatable with full IZF itself.

We end this section with a brief discussion about the relationship between BIST and other intuitionistic set theories in the literature. None of the existing literature on weak set theories interpretable in arbitrary elementary toposes includes unrestricted Replacement or Collection axioms in such theories. In having such principles, our set theories are similar to the "constructive" set theories of Myhill, Friedman and Aczel [15, 7, 1, 2]. However, because of our acceptance of the Powerset axiom, none of the set theories discussed above are "constructive" in the predicative sense intended by these authors.[5] In fact, in comparison with Aczel's CZF [1, 2], the theory BIST+ Coll represents *both* a strengthening *and* a weakening. It is a strengthening because it has the Powerset axiom, and this indeed amounts to a strengthening in terms of proof-theoretic strength. However, the full Ind schema is derivable in Aczel's CZF, but not in BIST+ Coll.

§3. **Toposes and systems of inclusions.** In this section we introduce the categories we shall use as models of BIST^- and its extensions. For us, a category $\mathcal{K}$ will always be *locally small*, i.e., the collection of objects $|\mathcal{K}|$ forms a (possibly proper) class, but the collection of morphisms $\mathcal{K}(A, B)$, between any two objects A, B, forms a set. We write **Set** for the category of sets. Of course all this needs to be understood relative to some meta-theory supporting a class/set distinction. To keep matters simple in this announcement, our meta-theory throughout is ZFC. A more careful analysis of meta-theoretic assumptions is given in the full version of the paper [4].

To fix notation, we briefly recall that an (*elementary*) *topos* is a category $\mathcal{E}$ with finite limits and with powerobjects:

Definition 3.1. A category $\mathcal{E}$ with finite limits has *powerobjects* if, for every object B there is an object $\mathcal{P}(B)$ and a mono $\ni_B \rightarrowtail \mathcal{P}(B) \times B$ such that, for every mono $R \overset{r}{\rightarrowtail} A \times B$ there exists a unique map $\chi_r \colon A \longrightarrow \mathcal{P}(B)$ fitting into a pullback diagram:

$$\begin{tikzcd} R \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \ni_B \arrow[d, tail] \\ A \times B \arrow[r, "\chi_r \times 1_B"'] & \mathcal{P}(B) \times B \end{tikzcd}$$

The intuitive reading of the above data is that objects are sets, the powerobject $\mathcal{P}(B)$ is the powerset of B, and $\ni_B$ is the membership relation.

We shall always assume that toposes come with *specified structure*, i.e., we have specified binary products $A \xleftarrow{\pi_1} A \times B \xrightarrow{\pi_2} B$, specified terminal

[5]For us, Powerset is unavoidable because we are investigating set theories associated with elementary toposes, where powerobjects are a basic ingredient of the structure.

object **1**, a specified equalizer for every parallel pair, and specified data providing the powerobject structure as above.

Any morphism $f : A \longrightarrow B$ in a topos factors (uniquely up to isomorphism) as an epi followed by a mono

$$f = A \twoheadrightarrow \mathrm{Im}(f) \rightarrowtail B.$$

Thus, given $f : A \longrightarrow B$, we can factor the composite on the left below, to obtain the morphisms on the right.

$$\ni_A \rightarrowtail \mathcal{P}(A) \times A \xrightarrow{1_{\mathcal{P}(A)} \times f} \mathcal{P}(A) \times B = \ni_A \twoheadrightarrow R_f \overset{r_f}{\rightarrowtail} \mathcal{P}(A) \times B$$

By the definition of powerobjects, we obtain $\chi_{r_f} : \mathcal{P}(A) \to \mathcal{P}(B)$. We write $\mathcal{P}f : \mathcal{P}(A) \to \mathcal{P}(B)$ for χ_{r_f}. Intuitively, the morphism $\mathcal{P}f$ represents the direct-image function determined by f. Its definition is independent of the choice of factorization. The operations $A \mapsto \mathcal{P}A$ and $f \mapsto \mathcal{P}f$ are the actions on objects and morphisms respectively of the *covariant powerobject functor*.

We wish to interpret the first-order language of Section 2 in an elementary topos $\mathcal{E}$. In fact, the topos $\mathcal{E}$ alone does not determine a canonical such interpretation. Thus the interpretation needs to be defined with reference to additional structure on $\mathcal{E}$. The required extra structure, a *directed structural system of inclusions* (*dssi*), is a collection of special maps, "inclusions", intended to implement a "subset" relation between objects of the topos. The situation is summarised by the equation:

$$\text{model of BIST}^- = \text{elementary topos } \mathcal{E} + \text{dssi on } \mathcal{E}. \tag{1}$$

In the remainder of this section, we introduce and analyse the required notion of dssi.

DEFINITION 3.2 (System of inclusions). A *system of inclusions* on a category $\mathcal{K}$ is a subcategory $\mathcal{I}$ (the *inclusion* maps, denoted $\hookrightarrow$) satisfying the four conditions below.

(si1) Every inclusion is a monomorphism in $\mathcal{K}$.

(si2) There is at most one inclusion between any two objects of $\mathcal{K}$.

(si3) For every mono $P \overset{m}{\rightarrowtail} A$ in $\mathcal{K}$ there exists an inclusion $A_m \hookrightarrow A$ that is isomorphic to m (in the slice category $\mathcal{K}/A$).

(si4) Given a commuting diagram, with i, j inclusions,

$$\begin{tikzcd} A' \arrow[r, hook, "i"] & A \\ A'' \arrow[u, "m"] \arrow[ur, hook, "j"'] & \end{tikzcd} \tag{2}$$

then m (which is necessarily a mono) is an inclusion.

We shall always assume that systems of inclusions come with a specified means of finding $A_m \hookrightarrow A$ from m in fulfilling (si3). By (si3), every object of $\mathcal{K}$ is an object of $\mathcal{I}$, hence every identity morphism in $\mathcal{K}$ is an inclusion. By (si2), the objects of $\mathcal{I}$ are preordered by inclusions. We write $A \equiv B$ if $A \hookrightarrow B \hookrightarrow A$. If $A \overset{i}{\hookrightarrow} B$ then $A \equiv B$ iff i is an isomorphism, in which case i^{-1} is the inclusion from B to A. We do not assume that inclusions form a partial order since there is no gain in convenience by doing so.

When working with an elementary topos $\mathcal{E}$ with a specified system of inclusions $\mathcal{I}$, we always take the image factorization of a morphism $A \xrightarrow{f} B$ in $\mathcal{E}$ to be of the form

$$A \xrightarrow{f} B = A \xrightarrow{e_f} \twoheadrightarrow \mathrm{Im}(f) \overset{i_f}{\hookrightarrow} B,$$

i.e., an epi followed by an *inclusion*, using (si3) to obtain such an image.

DEFINITION 3.3 (Directed system of inclusions). A system of inclusions $\mathcal{I}$ on a category $\mathcal{K}$ (with at least one object) is said to be *directed* if the induced preorder on $\mathcal{I}$ is directed (i.e., if, for any pair of objects A, B, there exists a specified object C_{AB} with $A \hookrightarrow C_{AB} \hookleftarrow B$).

DEFINITION 3.4 (Structural system of inclusions). A system of inclusions $\mathcal{I}$ on an elementary topos $\mathcal{E}$ is said to be *structural* if it satisfies the conditions below relating inclusions to the specified structure on $\mathcal{E}$.

(ssi1) For any parallel pair $A \underset{g}{\overset{f}{\rightrightarrows}} B$, the specified equalizer $E \rightarrowtail A$ is an inclusion.

(ssi2) For all inclusions $A' \overset{i}{\hookrightarrow} A$ and $B' \overset{j}{\hookrightarrow} B$, the specified product $A' \times B' \overset{i \times j}{\rightarrowtail} A \times B$ is an inclusion.

(ssi3) For every object A, the membership mono $\ni_A \rightarrowtail \mathcal{P}(A) \times A$ is an inclusion.

(ssi4) For every inclusion $A' \overset{i}{\hookrightarrow} A$, the direct-image map $\mathcal{P}A' \overset{\mathcal{P}i}{\rightarrowtail} \mathcal{P}A$ is an inclusion.

The structure we shall require to interpret the first-order language of Section 2 is a *directed structural system of inclusions* (henceforth *dssi*).

We make some observations concerning the existence of dssis. First, we observe that not every topos can have a dssi placed upon it. For a simple counterexample, consider the full subcategory of **Set** whose objects are the cardinals. This is a topos, as it is equivalent to **Set** itself. However, it can have no system of inclusions placed upon it. Indeed, if there were a system of inclusions, then, by condition (si3) of Definition 3.2, each of the two morphisms $1 \longrightarrow 2$ would have to be an inclusion, thus violating condition (si2). Since subset inclusions give a (partially-ordered) dssi on **Set**, the existence of a dssi is not preserved under equivalence of categories. In fact, every topos is equivalent to one carrying a dssi.

$$
\begin{array}{lll}
X \Vdash_\rho \mathsf{S}(x) & \text{iff} & \text{there exists } B \text{ with } I_x \hookrightarrow \mathcal{P}B \\
X \Vdash_\rho x = y & \text{iff} & \text{there exists } B \text{ with } A_x \xhookrightarrow{i} B \xhookleftarrow{j} A_y \\
 & & \quad \text{such that } i \circ \rho_x = j \circ \rho_y \\
X \Vdash_\rho x \in y & \text{iff} & \text{there exist inclusions } I_x \xhookrightarrow{i} B \text{ and } I_y \xhookrightarrow{j} \mathcal{P}B \text{ such} \\
 & & \text{that } X \xrightarrow{\langle j \circ e_y,\, i \circ e_x \rangle} \mathcal{P}B \times B \text{ factors through } \ni_B \\
X \Vdash_\rho \bot & \text{iff} & X \text{ is an initial object} \\
X \Vdash_\rho \phi \wedge \psi & \text{iff} & X \Vdash_\rho \phi \text{ and } X \Vdash_\rho \psi \\
X \Vdash_\rho \phi \vee \psi & \text{iff} & \text{there exist jointly epic } Y \xrightarrow{s} X \text{ and } Z \xrightarrow{t} X \\
 & & \text{such that } Y \Vdash_{\rho \circ s} \phi \text{ and } Z \Vdash_{\rho \circ t} \psi \\
X \Vdash_\rho \phi \rightarrow \psi & \text{iff} & \text{for all } Y \xrightarrow{t} X,\ Y \Vdash_{\rho \circ t} \phi \text{ implies } Y \Vdash_{\rho \circ t} \psi \\
X \Vdash_\rho \forall x.\, \phi & \text{iff} & \text{for all } Y \xrightarrow{t} X \text{ and } Y \xrightarrow{a} A,\ Y \Vdash_{(\rho \circ t)[a/x]} \phi \\
X \Vdash_\rho \exists x.\, \phi & \text{iff} & \text{there exists an epi } Y \xrightarrow{t} X \text{ and map } Y \xrightarrow{a} A \\
 & & \text{such that } Y \Vdash_{(\rho \circ t)[a/x]} \phi
\end{array}
$$

FIGURE 5. The forcing relation.

THEOREM 3.5. *Given a topos $\mathcal{E}$, there exists an equivalent category $\mathcal{E}'$ carrying a dssi $\mathcal{I}'$ relative to specified topos structure on $\mathcal{E}'$.*

To end this section, we discuss the extra structure we shall require to interpret IST and other set theories with full Separation.

DEFINITION 3.6 (Superdirected system of inclusions). A system of inclusions $\mathcal{I}$ on $\mathcal{K}$ is said to be *superdirected* if, for every set $\mathcal{A}$ of objects of $\mathcal{K}$, there exists an object B that is an upper bound for $\mathcal{A}$ in $\mathcal{I}$.

The structure we shall require to interpret set theories with full Separation is a *superdirected structural system of inclusions* (henceforth *sdssi*).

PROPOSITION 3.7. *If $\mathcal{E}$ is a small topos with an sdssi then, for every object A, it holds that $A \equiv \mathbf{1}$, hence every object is isomorphic to $\mathbf{1}$.*

Thus sdssi's are only interesting on locally small toposes whose objects form a proper class. The final two results of this section show that sdssi's are available on important classes of toposes from the literature.

THEOREM 3.8. *For any cocomplete topos $\mathcal{E}$, there is an equivalent category $\mathcal{E}'$ carrying an sdssi $\mathcal{I}'$ relative to specified topos structure on $\mathcal{E}'$.*

In particular, any Grothendieck topos can (up to equivalence) be endowed with an sdssi.

THEOREM 3.9. *For any realizability topos $\mathcal{E}$, there is an equivalent category $\mathcal{E}'$ carrying an sdssi $\mathcal{I}'$ relative to specified topos structure on $\mathcal{E}'$.*

§4. Interpreting set theory in a topos with inclusions. In this section we present the interpretation of the first-order language of Section 2 in an arbitrary elementary topos with dssi. This interpretation always validates the axioms of $\mathrm{BIST}^{-}+$ Coll. In addition, the axiom of Infinity (hence $\mathrm{BIST}+$ Coll) is validated if and only if the topos has a natural numbers object. Also, Restricted Excluded Middle is validated if and only if the topos is boolean. Furthermore, all the theories covered above are complete with respect to validity in the appropriate class of models. Finally, in the case that the dssi is superdirected, full Separation, hence the theory IST, is validated.

Let $\mathcal{E}$ be an arbitrary elementary topos with dssi $\mathcal{I}$. The interpretation of the first-order language is similar to the well-known Kripke–Joyal semantics of the Mitchell–Bénabou language, cf. [13], but with two main differences. First, we have to interpret the untyped relations $x = y$, $\mathsf{S}(x)$ and $x \in y$. Second, we have to interpret unbounded quantification. To address these issues, we make essential use of the inclusion structure on $\mathcal{E}$. In doing so, we closely follow Hayashi [8], who interpreted the ordinary language of first-order set theory using the canonical inclusions between so-called transitive objects in $\mathcal{E}$. The difference in our case is that we work with an arbitrary dssi on $\mathcal{E}$. See Section 1 for further comparison.

We interpret a formula $\phi(x_1, \ldots, x_k)$ (i.e., with at most $x_1, \ldots, x_k$ free) relative to the following data: an object X of $\mathcal{E}$, a "world"; and an "X-environment" ρ mapping each free variable $x \in \{x_1, \ldots, x_k\}$ to a morphism $X \xrightarrow{\rho_x} A_x$ in $\mathcal{E}$. We write $X \Vdash_\rho \phi$ for the associated "forcing" relation, which is defined inductively in Figure 5. In the definition, we use the notation $X \xrightarrow{e_x} I_x \xhookrightarrow{i_x} A_x$ for the epi-inclusion factorization of ρ_x. Also, given $Y \xrightarrow{t} X$, we write $\rho \circ t$ for the Y-environment mapping x to $\rho_x \circ t$. Similarly, given morphisms $A_x \xrightarrow{b_x} B_x$, for each free variable x, we write $b \circ \rho$ for the X-environment mapping x to $b_x \circ \rho_x$. Finally, given a variable $x \notin \{x_1, \ldots, x_k\}$, and a morphism $a\colon X \longrightarrow A_x$, we write $\rho[a/x]$ for the environment that agrees with ρ on $\{x_1, \ldots, x_k\}$, and which also maps x to a.

The next lemma gives a direct formulation of the derived forcing conditions for the various set-theoretic abbreviations used in Section 2.

LEMMA 4.1. *If* $I_z \xhookrightarrow{k} \mathcal{P}C$ *then*

$X \Vdash_\rho \forall x \in z.\,\phi$ *iff for all* $Y \xrightarrow{t'} X$ *and* $Y \xrightarrow{s'} C$, *if* $Y \xrightarrow{(k \circ e_z \circ t',\, s')} \mathcal{P}C \times C$ *factors through* $\ni_C$ *then* $Y \Vdash_{(\rho\circ t')[s'/x]} \phi$,

$X \Vdash_\rho \exists x \in z.\,\phi$ *iff there exists an epi* $Y \xrightarrow{t} X$ *and map* $Y \xrightarrow{s} C$ *such that* $Y \xrightarrow{(k \circ e_z \circ t,\, s)} \mathcal{P}C \times C$ *factors through* $\ni_C$ *and* $Y \Vdash_{(\rho\circ t)[s/x]} \phi$,

$X \Vdash_\rho x \subseteq y$ *iff there exists* B *such that* $I_x \xhookrightarrow{i} \mathcal{P}B \xhookleftarrow{j} I_y$ *and* $(i \circ e_x, j \circ e_y)\colon X \longrightarrow \mathcal{P}B \times \mathcal{P}B$ *factors through* $\subseteq_B \hookrightarrow \mathcal{P}B \times \mathcal{P}B$,

$X \Vdash_\rho ⅋x.\,\phi$ *iff* *there exist objects B and $R \hookrightarrow X \times B$ such that, for all A and maps $Y \xrightarrow{t} X$ and $Y \xrightarrow{s} A$, $Y \Vdash_{(\rho\circ t)[s/x]} \phi$ iff $\mathrm{Im}(p) \hookrightarrow R$, where $p = \langle t, s\rangle\colon Y \longrightarrow X \times A$,*

$X \Vdash_\rho !\phi$ *iff* *the family $\{Y \mid Y \xhookrightarrow{i} X$ and $Y \Vdash_{\rho\circ i} \phi\}$ has a greatest element under inclusion.*

For a sentence ϕ, we write $(\mathcal{E},\mathcal{I}) \models \phi$ to mean that, for all worlds X, it holds that $X \Vdash \phi$ (equivalently that $\mathbf{1} \models \phi$). Similarly, for a theory (i.e., set of sentences $\mathcal{T}$), we write $(\mathcal{E},\mathcal{I}) \models \mathcal{T}$ to mean that $(\mathcal{E},\mathcal{I}) \models \phi$, for all $\phi \in \mathcal{T}$. The theorem below is the main result of this announcement.

THEOREM 4.2 (Soundness and completeness). *For any theory $\mathcal{T}$ and sentence ϕ, the following are equivalent.*

1. $\mathrm{BIST}^- + \mathrm{Coll} + \mathcal{T} \vdash \phi$.
2. $(\mathcal{E},\mathcal{I}) \models \phi$, *for every topos $\mathcal{E}$ and dssi $\mathcal{I}$ satisfying $(\mathcal{E},\mathcal{I}) \models \mathcal{T}$.*

Soundness can be proved in the expected way by unwinding the forcing semantics and checking the validity of the axioms one by one. The details are surprisingly involved. The proof of completeness makes essential use of an alternative category-theoretic account of models of BIST, using an appropriately axiomatized notion of "category of classes", adapting earlier work on *Algebraic Set Theory* [10, 17, 5]. For the (lengthy) details, see [4].

The following two propositions can be used in combination with Theorem 4.2 to obtain sound and complete classes of models for extensions of the theory $\mathrm{BIST}^- + \mathrm{Coll}$ with Inf and/or REM.

PROPOSITION 4.3. $(\mathcal{E},\mathcal{I}) \models \mathrm{Inf}$ *iff $\mathcal{E}$ has a natural numbers object.*

PROPOSITION 4.4. $(\mathcal{E},\mathcal{I}) \models \mathrm{REM}$ *iff $\mathcal{E}$ is a boolean topos.*

An interesting aspect of Proposition 4.4 is that the underlying logic of the first-order set theories that we associate with boolean toposes is not in general classical. Such set theories always satisfy the restricted law of excluded middle REM, but not in general the full law LEM. Such "semiclassical" set theories have appeared elsewhere in the literature on intuitionistic set theories, see e.g., [16]. Here, as in [8], we find them arising naturally as a consequence of our forcing semantics.

The next result states that, in the presence of a superdirected system of inclusions, the full Separation schema is validated.

PROPOSITION 4.5. *If $\mathcal{I}$ is an sdssi on $\mathcal{E}$ then* $(\mathcal{E},\mathcal{I}) \models \mathrm{Sep}$.

Since BIST+ Sep+ Coll interprets IZF, we now, by Theorems 3.8 and 3.9, have the promised uniform explanation for why all Grothendieck and realizability toposes provide models of IZF.

In contrast to the characterizations of Inf and REM, Proposition 4.5 only establishes a sufficient condition for the validity of full Separation.

Indeed, there is no reason to expect BIST$^-$+ Coll+ Sep to be a complete axiomatization of the valid sentences with respect to toposes with sdssi's.

We next consider a further important aspect about the forcing semantics of the first-order language, its conservativity over the internal logic of $\mathcal{E}$. In order to fully express this, using the tools of the present section, one would need to add constants to the first-order language for the global points in $\mathcal{E}$, interpret these in the evident way in the forcing semantics, and give a laborious translation of the typed internal language of $\mathcal{E}$ into first-order set theory augmented with the constants. In principle, all this is routine. In practice, it would be tedious. Rather than pursuing this line any further, we instead refer the reader to the full paper [4], where the tools of categorical logic are used to express the desired conservativity property in more natural terms. At this point, we simply remark on one important consequence of the general conservativity result.

PROPOSITION 4.6. *Suppose $\mathcal{E}$ has a natural numbers object. Then for any first-order sentence ϕ in the language of arithmetic, $\mathcal{E} \models \phi$ in the internal logic of $\mathcal{E}$ if and only if $(\mathcal{E}, \mathcal{I}) \models \phi$ in the forcing semantics (using the natural translation of ϕ in each case).*

Proposition 2.10 follows as a consequence of the above result, by an application of Gödel's second incompleteness theorem.

We end this announcement with simple applications of the soundness theorem to obtain non-derivability results. Let A be any set. For each ordinal α, we construct the von-Neumann hierarchy $V_\alpha(A)$ relative to A as a set of atoms in the standard way, *viz*:

$$
\begin{aligned}
V_{\alpha+1}(A) &= A + \mathcal{P}(V_\alpha(A)), \\
V_\lambda(A) &= \bigcup_{\alpha<\lambda} V_\alpha(A) \qquad\qquad \lambda \text{ a limit ordinal.}
\end{aligned}
$$

Note that $V_0 = \emptyset$, and $\alpha \leq \beta$ implies $V_\alpha(A) \subseteq V_\beta(A)$. We write $V(A)$ for the unbounded hierarchy $\bigcup_\alpha V_\alpha(A)$.

For a limit ordinal $\lambda > 0$, we define the category $\mathbf{V}_\lambda(A)$ to have subsets $X \subseteq V_\alpha(A)$, for any $\alpha < \lambda$, as objects, and arbitrary functions as morphisms. It is readily checked that $\mathbf{V}_\lambda(A)$ is a boolean topos. Moreover, subset inclusions provide a dssi on $\mathbf{V}_\lambda(A)$ relative to the naturally given topos structure. In the propositions below, we omit explicit mention of the inclusion maps, which are always taken to be subset inclusions.

PROPOSITION 4.7.

1. $\mathbf{V}_\lambda(A) \models$ Inf *if and only if* $\lambda > \omega$ *or* $|A| \geq \aleph_0$.
2. $\mathbf{V}_\lambda(A) \models$ vN-Inf *if and only if* $\lambda > \omega$.

In particular, $\mathbf{V}_\omega(\mathbb{N}) \models$ Inf but $\mathbf{V}_\omega(\mathbb{N}) \not\models$ vN-Inf. Proposition 2.4 follows as an immediate consequence. In fact, more generally:

COROLLARY 4.8. BIST+ Coll+ REM $\nvdash$ vN-Inf.

By Proposition 4.7, we have that $\mathbf{V}_{\omega+\omega}(\emptyset) \models$ vN-Inf. Hence, $\mathbf{V}_{\omega+\omega}(\emptyset)$ is a model of BIST+ Coll+ REM+ vN-Inf. Examples such as this may run contrary to the expectations of readers familiar with classical set theory, where, in order to model Replacement and Collection, it is necessary to consider cumulative hierarchies $V_\lambda(A)$ with λ a strongly inaccessible cardinal. The difference in our setting is that our forcing semantics builds Collection directly into its interpretation of the existential quantifier. The "price" one pays for this is that the underlying logic of the set theory is intuitionistic. In consequence, the standard arguments using Replacement that take one outside of $V_\lambda(A)$ for λ non-inaccessible, are not reproducible. For example, attempts to construct the union of the chain $N, \mathcal{P}(N), \mathcal{P}^2(N), \ldots$ founder at it being impossible to define this chain inside the set theory, as the model $\mathbf{V}_{\omega+\omega}(\emptyset)$ demonstrates. Indeed, although $\mathbf{V}_{\omega+\omega}(\emptyset)$ is a model of BIST+Coll+REM+vN-Inf, it does not model Ind (thus LEM and Sep are also invalidated).

This example raises the question of what Replacement is useful for in BIST, since such classical constructions based on Replacement are unavailable. In fact, Replacement *is* a useful principle, but not for the usual reasons. It rather turns out to be essential for the development of any mathematics that involves an interplay between small and large structures.[6] This brings us full circle to our main envisaged application of BIST as a natural language for developing the mathematics of large structures relative to the internal logic of elementary toposes. We hope that this language will find interesting mathematical applications.

Acknowledgements. This work began in summer 2000 following discussions between Awodey and Simpson at the Category Theory 2000 conference in Como, Italy. We have been grateful for several opportunities to present this work since October 2001, in particular at: PSSL, Copenhagen, 2002 (Simpson and Butz); Meeting on Mathematische Logik, Oberwolfach, 2002 (Streicher); Logic Colloquium, Germany, 2002 (Butz); Ramifications of Category Theory, Florence, 2003 (Awodey); MFPS, Pittsburgh, 2004 (Awodey); Annual ASL Meeting, Stanford, 2005 (Awodey), and in graduate courses at RIMS, Kyoto University, autumn 2002 (Simpson) and CMU, Pittsburgh, spring 2003 (Awodey). We have benefited from discussions with Peter Aczel, Nicola Gambino, Saunders Mac Lane, Kentaro Sato and Dana Scott. Simpson's research was supported by an EPSRC Advanced Research Fellowship (2001–7), and a visiting professorship at RIMS, Kyoto University (2002–2003).

REFERENCES

[1] P. Aczel, *The type theoretic interpretation of constructive set theory*, ***Logic colloquium '77***, North Holland, 1978, pp. 55–66.

[6] For an example illustrating this point, consider formalizing the statement and proof of Freyd's Adjoint Functor Theorem for locally small categories in BIST.

[2] P. ACZEL and M. RATHJEN, *Notes on constructive set theory*, Technical Report 40, 2000/2001, Mittag–Leffler Institutes, Sweden, 2001.

[3] S. AWODEY, *An outline of algebraic set theory*, in preparation, 2007.

[4] S. AWODEY, C. BUTZ, A.K. SIMPSON, and T. STREICHER, *Relating first-order set theories, toposes and categories of classes*, in preparation, 2007.

[5] C. BUTZ, *Bernays–Gödel type theory*, ***Journal of Pure and Applied Algebra***, vol. 178 (2003), pp. 1–23.

[6] M.P. FOURMAN, *Sheaf models for set theory*, ***Journal of Pure and Applied Algebra***, vol. 19 (1980), pp. 91–101.

[7] H. FRIEDMAN, *Set theoretic foundations for constructive analysis*, ***Annals of Mathematics***, vol. 105 (1977), pp. 1–28.

[8] S. HAYASHI, *On set theories in toposes*, ***Logic Symposia, Hakone 1979, 1980***, Lecture Notes in Mathematics, vol. 891, Springer, 1981, pp. 23–29.

[9] P. T. JOHNSTONE, ***Sketches of an elephant: a topos theory compendium***, vol. 1–2, Oxford Logic Guides, no. 43–44, Oxford University Press, 2002.

[10] A. JOYAL and I. MOERDIJK, ***Algebraic set theory***, LMS Lecture Notes 220, Cambridge University Press, 1995.

[11] J. LAMBEK and P.J. SCOTT, ***Introduction to higher order categorical logic***, Cambridge University Press, 1986.

[12] F. W. LAWVERE and R. ROSEBRUGH, ***Sets for mathematics***, Cambridge University Press, 2003.

[13] S. MAC LANE and I. MOERDIJK, ***Sheaves in geometry and logic: a first introduction to topos theory***, Universitext, Springer Verlag, 1992.

[14] A. R. D. MATHIAS, *The strength of Mac Lane set theory*, ***Annals of Pure and Applied Logic***, vol. 110 (2001), pp. 107–234.

[15] J. MYHILL, *Constructive set theory*, ***The Journal of Symbolic Logic***, vol. 40 (1975), pp. 347–382.

[16] A. ŠČEDROV, *Intuitionistic set theory*, ***Harvey Friedman's research on the foundations of mathematics***, Elsevier Science Publishers, 1985, pp. 257–284.

[17] A. K. SIMPSON, *Elementary axioms for categories of classes*, ***Proceedings of the 14th annual IEEE symposium on logic in computer science***, 1999, pp. 77–85.

[18] ———, *Computational adequacy for recursive types in models of intuitionistic set theory*, ***Annals of Pure and Applied Logic***, vol. 130 (2004), pp. 207–275.

[19] ———, ***Constructive set theories and their category-theoretic models. From sets and types to topology and analysis***, Oxford University Press, 2005.

[20] A. S. TROELSTRA and D. VAN DALEN, ***Constructivism in mathematics***, vol. I–II, North Holland, Amsterdam, 1988.

DEPARTMENT OF PHILOSOPHY, CMU
PITTSBURGH, USA
E-mail: awodey@cmu.edu

IT, UNIVERSITY OF COPENHAGEN, DENMARK
E-mail: butz@it-c.dk

LFCS, SCHOOL OF INFORMATICS
UNIVERSITY OF EDINBURGH, UK
E-mail: Alex.Simpson@ed.ac.uk

FACHBEREICH MATHEMATIK
TU DARMSTADT, GERMANY
E-mail: streicher@mathematik.tu-darmstadt.de