Relating first-order set theories, toposes and categories of classes

Steve Awodey [1]

Department of Philosophy, Carnegie Mellon University, Pittsburgh, USA

Carsten Butz [2]

IT University, Copenhagen, Denmark

Alex Simpson [3]

LFCS, School of Informatics, University of Edinburgh, UK

Thomas Streicher [4]

Fachbereich Mathematik, Technische Universität Darmstadt, Germany

---

DRAFT OF 25 October 2007

THIS IS A DRAFT OF AN UNFISHED PAPER. IT HAS BEEN PLACED ON-LINE IN SUPPORT OF OUR ACCOMPANYING BSL ANNOUNCEMENT (PUBLISHED SEPTEMBER 2007).

THE CURRENT STATUS OF THIS PAPER IS AS FOLLOWS. §12, MISSING. ALL OTHER SECTIONS: MATHEMATICAL CONTENTS IN PLACE, BUT MUCH TIDYING, INSERTING OF REFERENCES, ETC., REMAINS TO BE DONE.

Contents

[1] Steve's thanks.
[2] Carsten's thanks.
[3] Research supported by an EPSRC Advanced Research Fellowship (2001–), and a visiting professorship at RIMS, Kyoto University (2002–2003).
[4] Thomas' thanks.

1 Introduction

The notion of elementary topos abstracts from the structure of the category of sets. retaining many of its essential features. Nonetheless, elementary toposes include 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; see [[REFS]] for general overviews.

Not only are elementary toposes generalized categories of sets, but it is also possible to view them as categories of generalized sets. Indeed, elementary toposes possess an *internal logic*, which is a form of higher-order type theory. see e.g. [8,**?**,9,6], and which allows one to reason with objects of the topos as if they were *abstract sets* in the sense of [**?**]; that is, as if they were collections of elements. The reasoning supported by the internal logic is both natural and powerful, but 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. The use of intuitionistic logic is thus an inevitable feature of internal reasoning in toposes. Furthermore, as fields such as *synthetic differential geometry* [**?**] and *synthetic domain theory* [**?**] demonstrate, the non-validity of classical logic is a strength rather than a weakness of the internal logic. In these areas, intuitionistic logic offers the opportuntity of working consistently with useful but 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 classes, such as 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, the category of *locales* relative to a topos is frequently considered as the natural home for doing topology in a topos [6] [[CHECK REF]]. Although locally small, the category of locales is not a small category (from the viewpoint of the topos), and there is therefore no way of quantifying over all locales directly within the internal logic itself. Similarly, recent approaches to synthetic domain theory work with a derived category of *predomains* relative to a topos, which is also locally small, but not necessarily small [**?**,**?**].

The standard 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). In this paper we provide the foundations for an alternative more elementary 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 which does allow quantification over classes, including the class of all sets (i.e. all objects of the topos). We believe that this extended logic will provide a useful tool for establishing properties of non-small structures (e.g., large categories), relative to a topos, using straightforward set-theoretic arguments. In fact, one such application of our work has already appeared [**?**].

In Part I of the paper, we present the set theory that we shall interpret over an aribitrary 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, such as Friedman's IZF [[REFS]]. Two of the differences are minor: in BIST the universe may contain non-sets as well as sets, and non-well-founded sets are permitted (though not obliged to exist). These differences are inessential conveniences, adopted to make the connections established in this paper more natural. (Arguably, they also make BIST closer to mathematical practice, see [**?**].) 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 BIST to be compatible with the internal logic of any elementary topos (with natural numbers object), and in the free such topos the internal logic is exactly HAH.

Property (3) means that BIST is necessarily proof-theoretically weaker than IZF (which has the same proof-theoretic strength as ZF). That such weakness is necessary has long been recognised. The traditional account has been that the appropriate set theory is *bounded Zermelo (bZ)* set theory,[5] which is ZF set theory with the axiom of Replacement removed and with Separation restricted to bounded (i.e. Δ_0) formulas [[REFS]]. 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 boolean topos whose objects are the elements of the model and whose internal logic expresses truth in the model. Conversely, given any well-pointed (hence boolean) 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 [[REFS]] for detailed accounts of this correspondence.

This standard story is unsatisfactory in several respects. First, it applies only to well-pointed (hence boolean) toposes. Second, by only expressing properties of transitive objects in $\mathcal{E}$, whole swathes of such a topos may be ignored by the set theory. Third, with the absence of Replacment, bZ is neither a particularly convenient nor natural set theory to reason in, see [**?**] for a critique.

We argue that the set theory BIST introduced in Section 2 provides a much more satisfactory connection with elementary toposes. We have already stated that we shall interpret this set theory over an arbitrary elementary topos (with natural numbers object). In fact, we shall do this 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. Thus any elementary topos is (equivalent to) a category of sets compatible with the set theory BIST. Moreover, we believe that BIST is a rather natural theory in terms of the set-theoretic reasoning it supports. In particular, one of its attractive features is that it contains the full axiom of Replacement. In fact, not only do we model Replacement, but we also show that every topos validates the stronger axiom of Collection (Coll). Some readers familiar with classical (but not intuitionistic) set theory may be feeling uncomfortable at this point. In classical set theory, Replacement is equivalent to Collection and implies full Separation, thus taking one beyond the the proof-theoretic strength of elementary toposes. The situation is completely different under intuitionistic logic. Indeed, it has long been known from work of Friedman and others that, over an intuitionistic base logic, the full axioms of Replacement and Collection are compatible with proof-theoretically weak set theories [[REFS]]. A reader eager to see examples illustrating this weakness is referred to the discussion at the end of Section 4.

[5] Also known as "Mac Lane" set theory [**?**].

The precise connection between BIST and elementary toposes is elaborated in Part II of this paper. In order to interpret quantification over classes, we have to address a fourth difference between the internal logic of toposes and first-order set theories.

(4) In first-order set theories (such as BIST), 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, which 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. In particular, a dssi on a topos induces a finite union operation on objects, using which (generalized) elements of different objects can be compared for equality.

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.

We mention, at this point, that a similar forcing semantics for first-order set theory in toposes was previously introduced by Hayashi in [?]. In Hayashi's case, the notion of inclusion was provided by the canonical notion of inclusion map between the transitive objects in a topos, and his interpretation of first-order set theory was thus only able to express properties of transitive objects. In contrast, because we use the general axiomatic notion of dssi on a topos, all objects of the topos are included in our interpretation. Furthermore, we considerably extend Hayashi's results 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 showing that the theory BIST+ Coll axiomatizes exactly the set-theoretic properties validated by our forcing semantics. That such a completeness result holds is by no means routine, and its proof is one of the main contributions of the present paper.

The proof of completeness for the forcing interpretation involves a lengthy detour through an axiomatic theory of "categories of classes", which is also of interest in its own right. This is the topic of Part III of the paper.

The idea behind Part III is to consider a second type of category-theoretic model for first-order set theories. Because such set theories permit quantification over classes, rather than merely considering categories of sets, it is natural to instead take categories of classes as the models since this allows the quantifiers of the set theory to be interpreted using the quantifiers in the internal logic of the categories. This idea was first proposed and developed in the pioneering book on *Algebraic Set Theory* by Joyal and Moerdijk [7], in which they gave an axiomatic account of categories of classes, imposing sufficient structure for these to model Friedman's IZF set theory. Their axiomatic structure was later refined by the third author, who obtained a corresponding completeness result for IZF [11]. See also [3] for related work.

In this paper we are interested in axiomatizing appropriate structure on a category of classes suitable for modelling the set theory BIST of Part I. We introduce this in two stages. In Section 5, we present the notion of a category with *basic class structure*, which axiomatizes those properties of the category of classes that are compatible with a very weak (predicative) constructive set theory. Although the study of such predicative set theories is outside the scope of the present paper (cf. [**?**,**?**]), the notion of basic class structure nonetheless serves the purpose of identifying the basic category-theoretic structure of categories of classes. Second, in Section 6, we consider the additional properties that we need to axiomatize a *category of classes*, intended to correspond to the structure of the category of classes in the set theory BIST. Such categories of classes provide the main vehicle for our investigations throughout the remainder of Part III.

The precise connection between BIST and categories of classes is elaborated in Section 7. Any category of classes $\mathcal{C}$ contains a *universal object* U, and we show how this is perceived as a set-theoretic universe by the internal logic of $\mathcal{C}$. Indeed, such universes always validate the axioms of BIST. Thus BIST is sound with respect to universes in categories of classes. In fact, BIST is also complete for such interpretations. The proof is by construction of a simple syntactic category, following [11].

The goal of Section 8 is to show that every elementary topos embeds as the full subcategory of sets within some category of classes. Since categories of classes model BIST, this justifies our earlier assertion that, for any elementary topos, the collection of objects of the topos (more precisely, of an equivalent topos, see below) can be seen as the class of all sets in a model of BIST. In order to obtain the embedding result, we again require a dssi (in the sense of Section 3) on the topos. The category of classes is then obtained by a form of "ideal completion", analogous to the ideal completion of a partial order.

The construction of Section 8 gives rise to a second interpretation of the theory BIST+ Coll over an elementary topos (with dssi), since this theory is modelled by the universal object in the category of ideals. In the short Section 9, we show that the new interpretation in ideals coincides with the old interpetation given by the forcing semantics of Section 4. Thus the soundness of BIST+ Coll in the ideal completion of a topos, provides a second proof of the soundness of the theory BIST+ Coll with respect to the forcing interpretation of Section 4. Furthermore, the completeness of the forcing semantics is thereby reduced to the completeness of BIST+ Coll with respect to categories of ideals.

In Section 10, we finally prove this missing completeness result. The approach is to reduce the known completeness of BIST+ Coll with respect to arbitrary categories of classes (satisfying an appropriate Collection axiom), from Section 7, to an analogous result for categories of ideals. To this end, we show that any categories of classes satisfying Collection has a suitably "conservative" embedding into a category of ideals. The proof of this result fully exploits the elementary nature of our axiomatization of categories of classes, making use of the closure of categories of classes under filtered colimits and other general model-theoretic constructions from categorical logic.

Parts I–III described above form the main body of the paper. However, there is a second thread within them, the discussion of which we have postponed till now. It is known that many naturally occurring toposes, which are defined over the external category of sets (which we take to be axiomatized by ZFC) are able to model Friedman's IZF set theory, which is proof-theoretically as strong as ZFC. For example, all cocomplete toposes (and hence all Grothendieck toposes) enjoy this property; see Fourman [4] and Hayashi [**?**] for two different accounts of this. Similarly, all realizability toposes [**?**,**?**] also model IZF, as follows, for example, from McCarty's realizability interpretation of IZF [**?**]. Thus, if one is primarily interested in such "real world" toposes, then the account above is unsatisfactory in merely detailing how to interpret the weak set theory BIST inside them.

To address this, in parallel with the development already described, we further show how the approach described above adapts to model the full Separation axiom (Sep) in toposes such as cocomplete and realizability toposes. (The set theory BIST+ Coll+ Sep is interinterpretable with IZF.) The appropriate structure we require for this task is a modification of the notion of dssi from Section 3, extended 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)*, we show that 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. To our knowledge, no such uniform explanation was known before.

We also show that the construction of the category of ideals, of Section 8, adapts in the presence of an sdssi. Indeed, given an sdssi on a topos, we define the full subcategory of *superideals* within the category of ideals. We show that this is again a category of classes, which, in addition, satisfies the Separation axiom of [7,11]. In particular, the category of superideals is a category with *class(ic)* structure in the sense of [11], and models both BIST + Coll + Sep and IZF. We therefore obtain a uniform embedding of both cocomplete and realizability toposes in categories with class structure *à la* [11]. We mention that one application of these embeddings has already appeared in Section 15 of [**?**].

Finally, in Part IV of the paper, we fulfil some technical obligations postponed from earlier. In Section 11, we show that every elementary topos is equivalent to a topos carrying a dssi. Thus the forcing interpretation and construction of the category of ideals can indeed be defined for any topos, as claimed above. We also show that every cocomplete topos (again up to equivalence) can be endowed with an sdssi. Similarly, in Section 12, we show that every realizability topos is also equivalent to one carrying an sdssi. In doing so, we establish that every object in a realizability topos occurs (up to isomorphism) somewhere within the cumulative hierarchy of McCarty's realizability interpretation of IZF. Thus the difference between realizability toposes and McCarty's realizability interpretation of set theory turns out to be purely presentational rather than substantive.

PART I — FIRST-ORDER SET THEORIES

2 Basic Intuitionistic Set Theory (BIST) and extensions

All first-order set theories considered in this paper 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. It is from this latter viewpoint that we introduce the theory.

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.

Membership	$y \in x \;\rightarrow\; \mathsf{S}(x)$
Extensionality	$\mathsf{S}(x) \,\wedge\, \mathsf{S}(y) \,\wedge\, (\forall z.\, z \in x \leftrightarrow z \in y) \;\rightarrow\; x = y$
Indexed-Union	$\mathsf{S}(x) \,\wedge\, (\forall y \in x.\, \reflectbox{S} z.\, \phi) \;\rightarrow\; \reflectbox{S} z.\, \exists y \in x.\, \phi$
Emptyset	$\reflectbox{$\mathsf{S}$} z.\, \perp$
Pairing	$\reflectbox{$\mathsf{S}$} z.\, z = x \,\vee\, z = y$
Equality	$\reflectbox{$\mathsf{S}$} z.\, z = x \,\wedge\, z = y$
Powerset	$\mathsf{S}(x) \;\rightarrow\; \reflectbox{S} y.\, y \subseteq x$

Fig. 1. Axioms for BIST$^-$

$$\text{Coll} \qquad \begin{array}{l} \mathsf{S}(x) \,\wedge\, (\forall y \in x.\, \exists z.\, \phi) \;\rightarrow \\ \qquad\qquad \exists w.\, (\, \mathsf{S}(w) \,\wedge\, (\forall y \in x. \exists z \in w.\, \phi) \wedge\, (\forall z \in w. \exists y \in x.\, \phi)\,) \end{array}$$

Fig. 2. Collection axiom

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.

The set theory BIST is formulated as a theory in intuitionistic first-order logic with equality. [6] 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 $\forall x.\, (x \in y \rightarrow \phi)$ 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, it is appropriate to 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

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

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 a generalized quantifier, reading $\text{Ƨ}x.\,\phi$ as "there are set-many x satisfying ϕ".

Often we shall consider BIST$^-$ together with the axiom of Collection, presented in Figure 2.[7] One reason for not including Collection as one of the axioms of BIST$^-$ is that it seems better to formulate the many results that do not 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 demonstrate, 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 (defined using a combination of Pairing and Indexed-Union). We write δ_{xy} for the set $\{z \mid z=x\wedge z=y\}$ (which is a set by the Equality axiom). It follows from the Equality and Indexed-Union axioms that, for sets x and y, the intersection $x\cap y$ is a set, because $x\cap y=\bigcup_{z\in x}\bigcup_{w\in y}\delta_{zw}$.

[7] Coll, 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.

We now study Separation in BIST$^-$. By an *instance of Separation*, we mean a formula of the form [8]

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

which states that the subclass $\{y \in x \mid \phi\}$ of x is actually a sub*set* of x. We now analyse the instances of Separation that are derivable in BIST$^-$.

Following [2], the development hinges on identifying when a formula ϕ expresses a property of a restricted kind that is possible to use in instances of Separation. 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.* [9] Note that, trivially, $(\phi \leftrightarrow \psi) \;\rightarrow\; (!\phi \leftrightarrow !\psi)$. The utility of the concept is given by the lemma below, showing that the notion of restrictedness exactly captures when a property can be used in an instance of Separation.

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

PROOF. We reason in BIST$^-$. Suppose that, for all $y \in x$, $!\phi$, and also $\mathsf{S}(x)$. We must show that $\text{Ƨ}y.\,(y \in x \;\wedge\; \phi)$. For each $y \in x$, we have $\text{Ƨ}z.\, z = \emptyset \wedge \phi$. Hence, by Replacement, $\text{Ƨ}z.\, z = y \wedge \phi$. Thus by Indexed-Union, $\text{Ƨ}z.\,(\exists y \in x.\, z = y \wedge \phi)$. I.e. $\text{Ƨ}y.\,(y \in x \;\wedge\; \phi)$ as required.

Conversely, suppose that $\phi[x, y]$-Sep holds. Take any $y_0 \in x$. By Membership, x is a set hence $\text{Ƨ}z.\,(\exists y \in x.\, z = y \wedge \phi)$. Write w for this set. Then $w \cap \{y_0\}$ is a set. For any $z \in w \cap \{y_0\}$ there exists a unique v such that $v = \emptyset$. Therefore, by Replacement, $\{v \mid v = \emptyset \wedge \exists z.\, z \in w \cap \{y_0\}\}$ is a set. In other words, $\{v \mid v = \emptyset \wedge \phi[x, y_0]\}$ is a set, i.e. $!\phi[x, y_0]$. Thus indeed, $\forall y \in x.\, !\phi$. $\square$

We next establish important closure properties of restricted propositions.

Lemma 2.2 *The following all hold in* BIST$^-$*.*

(1) $!(x = y)$.

[8] 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.

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

(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* $!\phi$.

PROOF. We reason in BIST$^-$.

(1) Using Equality, $\{v \mid v = x \wedge v = y\}$ is a set, call it w. For every $v \in w$ there exists a unique u with $u = \emptyset$. So, by Replacement, $\{z = \emptyset \mid \exists v.\, v \in w\}$ is a set, i.e. $\text{Ƨ}z.\,(z = \emptyset \,\wedge\, x = y)$ as required.

(4) Suppose $\mathsf{S}(x)$ and, for all $y \in x$, $!\phi$, i.e. $\text{Ƨ}z.\,(z = \emptyset \,\wedge\, \phi)$. That $!(\exists y \in x.\, \phi)$ holds follows from Indexed-Union, because $\text{Ƨ}z.\, \exists y \in x.\,(z = \emptyset \,\wedge\, \phi)$, hence $\text{Ƨ}z.\,(z = \emptyset \,\wedge\, \exists y \in x.\, \phi)$. To show that $!(\forall y \in x.\, \phi)$, consider the set $w = \{y \in x \mid \phi\}$, which is a set by Lemma 2.1. By (1) above and Lemma 2.1, $\{z \in \{\emptyset\} \mid w = x\}$ is a set. But $w = x$ iff $\forall y \in x.\, \phi$. Hence indeed $\text{Ƨ}z.\,(z = \emptyset \,\wedge\, \forall y \in x.\, \phi)$, i.e. $!(\forall y \in x.\, \phi)$.

(2) We have
$$y \in x \;\leftrightarrow\; \exists z \in x.\, z = y\,.$$
Thus, if $\mathsf{S}(x)$, then we obtain $!(y \in x)$ by combining (1) and (4) above.

(3) Suppose $!\phi$ and $!\psi$. We show that $!(\phi \rightarrow \psi)$, which is the most interesting case. For this, we have
$$(\phi \rightarrow \psi) \;\leftrightarrow\; (\forall z \in \{z \mid z = \emptyset \,\wedge\, \phi\}.\, \psi)\,.$$
But $\{z \mid z = \emptyset \,\wedge\, \phi\}$ is a set because $!\phi$. Also $!\psi$. Thus $!(\phi \rightarrow \psi)$ by (4).

(5) Suppose $\phi \vee \neg\phi$. Then, (for all $x \in \{\emptyset\}$) there exists a unique set y satisfying
$$(y = \emptyset \,\wedge\, \phi) \,\vee\, (y = \{\emptyset\} \,\wedge\, \neg\phi)\,.$$
So, by Replacement, $w = \{y \mid (y = \emptyset \,\wedge\, \phi) \,\vee\, (y = \{\emptyset\} \,\wedge\, \neg\phi)\}$ is a set. By (1) and Lemma 2.1, $\{y \mid y \in w \,\wedge\, y = \emptyset\}$ is a set. But
$$(y = \emptyset \,\wedge\, \phi) \;\leftrightarrow\; (y \in w \,\wedge\, y = \emptyset)\,.$$
So indeed $\text{Ƨ}y.\,(y = \emptyset \,\wedge\, \phi)$.

□

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

R$\mathcal{S}$	Restricted Sethood	$!\mathsf{S}(x)$
R$\exists$	Restricted $\exists$	$(\forall x.\, !\phi) \;\rightarrow !(\exists x.\, \phi)$
R$\forall$	Restricted $\forall$	$(\forall x.\, !\phi) \;\rightarrow !(\forall x.\, \phi)$

Fig. 3. Axioms on restricted properties

In order to obtain further instances of Separation, it is necessary to augment BIST^{-} with further axioms. In this connection, we study the axioms in Figure 3. The point of the first lemma is that the result holds without the assumption $\mathsf{S}(x)$.

Lemma 2.4 BIST^{-}+ R$\mathcal{S}$ $\vdash$ Ƨ$y.\, y \in x$.

PROOF. We reason in BIST^{-}+ R$\mathcal{S}$. Consider the set $\{x\}$. By R$\mathcal{S}$, we have a set $u = \{x' \in \{x\} \mid \mathsf{S}(x')\}$. Clearly, for all $x' \in u$, $\mathsf{S}(x')$. So $w = \bigcup u$ is a set, i.e. Ƨ$y.\, y \in w$. But $y \in w \;\leftrightarrow\; y \in x$. Thus indeed Ƨ$y.\, y \in x$. □

Corollary 2.5 *The following all hold in* BIST^{-}+ R$\mathcal{S}$*.*

(1) $!\mathsf{S}(x)$*;*
(2) $!(y \in x)$*;*
(3) if $\forall y \in x.\, !\phi$ *then* $!(\forall y \in x.\, \phi)$ *and* $!(\exists y \in x.\, \phi)$.

PROOF. Statement 1 is immediate. Statements 2 and 3 follow easily from Lemma 2.2.(2)&(4), because $y \in x$ holds if and only if y is a member of the collection of all elements of x, which is a set by Lemma 2.4. □

We say that a formula is *bounded* if all quantifiers occurring in it are bounded, and we write bSep for the schema of *bounded Separation*, namely $\phi[x, y]$-Sep for all bounded ϕ. By combining Lemmas 2.1, 2.2 and Corollary 2.5, it is clear that bounded Separation is derivable in BIST^{-}+ R$\mathcal{S}$. Moreover, as R$\mathcal{S}$ is itself an instance of bounded Separation, we obtain:

Corollary 2.6 BIST^{-}+ bSep $=$ BIST^{-}+ R$\mathcal{S}$*.*

We write Sep for the full Separation schema: $\phi[x, y]$-Sep for all ϕ. Obviously, this is equivalent to the schema $!\phi$ for all ϕ. To obtain Sep from bounded Separation, it suffices for restricted properties to be closed under arbitrary quantification. In fact, as the next lemma shows, closure under existential quantification is alone sufficient. This will prove useful in Section 4 for verifying Sep in models.

Lemma 2.7 $\mathrm{BIST}^{-} + \mathrm{R}\exists \vdash \mathrm{R}\forall$. [10]

PROOF. Assume $\mathrm{R}\exists$. Suppose that $\forall x.\, !\phi$. We show below that

$$(\forall x.\, \phi) \quad \leftrightarrow \quad \forall p \in \mathcal{P}(\{\emptyset\}).\, (\exists x.\, (\phi \rightarrow \emptyset \in p)) \rightarrow \emptyset \in p\,. \tag{1}$$

It then follows that $!(\forall x.\, \phi)$, because the right-hand formula is restricted by Lemma 2.2 and $\mathrm{R}\exists$.

For the left-to-right implication of (1), suppose $\forall x.\, \phi$, and suppose that $p \in \mathcal{P}(\{\emptyset\})$ satisfies $\exists x.\, (\phi[x] \rightarrow \emptyset \in p)$. Then there is some x_0 such that $\phi[x_0] \rightarrow \emptyset \in p$. But $\phi[x_0]$ because $\forall x.\, \phi$. Thus indeed $\emptyset \in p$.

For the converse, suppose that the right-hand side of (1) holds. We must show that $\forall x.\, \phi$. Take any x_0. Define $p_0 = \{\emptyset \mid \phi[x_0]\}$. Then p_0 is a set because $!\phi[x_0]$. Thus $p_0 \in \mathcal{P}(\{\emptyset\})$. Hence, by the assumption, we have $(\exists x.\, (\phi[x] \rightarrow \emptyset \in p_0)) \rightarrow \emptyset \in p_0$. But, by the definition of p_0, we have $\phi[x_0] \rightarrow \emptyset \in p_0$. So $\emptyset \in p_0$. Hence, again by the definition of p_0, we have $\phi[x_0]$ as required. $\square$

The above proof was inspired by the derivation of universal quantification from existential quantification in [11].

Corollary 2.8 $\mathrm{BIST}^{-} + \mathrm{Sep} = \mathrm{BIST}^{-} + \mathrm{R}\mathcal{S} + \mathrm{R}\exists$.

PROOF. That $\mathrm{BIST}^{-} + \mathrm{Sep}$ validates $\mathrm{R}\mathcal{S}$ and $\mathrm{R}\exists$ is immediate. For the converse, we have that $\mathrm{R}\exists$ implies $\mathrm{R}\forall$, by Lemma 2.7. Thus, we can derive $!\phi$, for any formula ϕ, by induction on its structure, using the closure conditions of Lemma 2.2, Corollary 2.5, $\mathrm{R}\exists$ and $\mathrm{R}\forall$. $\square$

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 $\text{Ƨ}x.\, \phi$. We write U for the class $\{x \mid x = x\}$, and $\mathcal{S}$ for the class $\{x \mid \mathsf{S}(x)\}$. 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:

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

[10] Here, $\mathrm{R}\exists$ and $\mathrm{R}\forall$ are the full schemas.

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

Fig. 4. Infinity axioms

where $(x, y) = \{\{x\}, \{x, y\}\}$ is the standard Kuratowski pairing construction. [11] Using Indexed-Union, one can prove that if A and B are both sets then so is $A \times B$ [REF]. 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 4. Define

$$\text{BIST} \;=\; \text{BIST}^{-} + \text{Inf}\,.$$

For the sake of comparison, we also include, in Figure 4, the familiar von Neumann axiom of Infinity, which does make assumptions about the nature of the elements of the assumed infinite set. We shall show in Section 4 that:

Proposition 2.9 $\text{BIST} + \text{Coll} \;\not\vdash\; \text{vN-Inf}\,.$

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

[11] See [REF] for a proof that Kuratowski pairing works intuitionistically.

use the formula $Nat(N, 0, s)$:

$$\begin{aligned}&0 \in N \ \wedge\ s \in N^N \ \wedge\ (\forall x \in N.\, s(x) \neq 0) \ \wedge\ (\forall x, y \in N.\, s(x) = s(y) \ \rightarrow\ x = y) \\ &\qquad \wedge\ \forall X \in \mathcal{P}N.\, (0 \in X \ \wedge\ \forall x \in X.\, s(x) \in X) \ \rightarrow\ X = N \ ,\end{aligned}$$

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

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

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. (Indeed this is common practice in the formulation of weak intuitionistic set theories, see e.g. [[REFS]].) Our reason for not doing so is that, in Parts II–III, we shall consider various 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

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

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

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

Lemma 2.10 BIST $\vdash$ RInd.

PROOF. Reasoning in BIST, suppose, for all $x \in N$, $!\phi[x]$. Then, by Lemma 2.1, the class $X = \{x \mid x \in N \wedge \phi[x]\}$ is a subset of N. Thus the induction property holds by the definition of N from I as the smallest subset containing 0 and closed under s. $\Box$

Thus induction holds for restricted properties.

Corollary 2.11 BIST+ Sep $\vdash$ 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	$\phi \vee \neg\phi$

Fig. 5. Excluded middle axioms

PROOF. Immediate from Lemma 2.10. $\Box$

On the other hand:

Proposition 2.12 BIST+ Ind $\vdash$ vN-Inf .

PROOF. One proves the following statement by induction.

$$\begin{aligned}&\forall n \in N.\, \exists! f_n \in \mathcal{S}^{\{x \in N \mid x \leq n\}}.\\ &\quad \forall x \in \{x \in N \mid x \leq n\}.\ (x = 0 \rightarrow f_n(x) = \emptyset) \wedge \\ &\qquad\qquad (x > 0 \rightarrow f_n(x) = f_n(x-1) \cup \{f_n(x-1)\})\,,\end{aligned}$$

making use of standard arithmetic operations and relations. Then a set satisfying vN-Inf is constructed as the union of the images of all f_n, using Indexed-Union. $\Box$

Corollary 2.13 BIST+ Coll $\nvdash$ Ind .

PROOF. Immediate from Propositions 2.9 and 2.12. $\Box$

Figure 5 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 2.14 *In* BIST$^-$*, axioms* DE *and* REM *are equivalent.*

PROOF. REM implies DE because equalities are restricted. Conversely, working in BIST$^-$, suppose $!\phi$. Thus $w = \{z \mid z = \emptyset \wedge \phi\}$ is a set. So, by DE, either $w = \{\emptyset\}$ or $w \neq \{\emptyset\}$. In the first case ϕ holds. In the second case $\neg\phi$ holds. Thus indeed $\phi \vee \neg\phi$. $\Box$

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

Proposition 2.15 $\mathrm{BIST}^{-} + \mathrm{LEM} \vdash \mathrm{Sep}$.

PROOF. By Lemma 2.2.5, $\mathrm{BIST}^{-} + \mathrm{LEM} \vdash !\phi$, for any ϕ. Sep then follows by Lemma 2.1. □

Corollary 2.16 $\mathrm{BIST}^{-} + \mathrm{Sep} + \mathrm{REM} = \mathrm{BIST}^{-} + \mathrm{LEM}$.

PROOF. Immediate from Proposition 2.15 and Lemma 2.1. □

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 2.17 $\mathrm{BIST} + \mathrm{Coll} + \mathrm{REM} \nvdash \mathrm{Con(HAH)}$,

where Con(HAH) is the Π^0_1 formula asserting the consistency of Higher-order Heyting Arithmetic [[REF]]. 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 4.10 and surrounding discussion. On the other hand,

Proposition 2.18 $\mathrm{BIST} + \mathrm{Ind} \vdash \mathrm{Con(HAH)}$.

PROOF. One proves the following statement by induction.

$$\begin{aligned}
&\forall n \in N.\, \exists! f_n \in \mathcal{S}^{\{x \in N \mid x \leq n\}}. \\
&\qquad \forall x \in \{x \in N \mid x \leq n\}.\ (x = 0 \ \rightarrow\ f_n(x) = N) \wedge \\
&\qquad\qquad\qquad (x > 0 \ \rightarrow\ f_n(x) = \mathcal{P}(f_n(x-1)))\,,
\end{aligned}$$

where $\mathcal{P}$ is the powerset operation. Define the set $V_{\omega+\omega}$ to be the union of the images of all f_n. In the usual way, $V_{\omega+\omega}$ is a non-trivial internal model of higher-order arithmetic, where the arithmetic modelled is intuitionistic because BIST is an intuitionistic theory. □

Corollary 2.19 *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.

PROOF. By Proposition 2.15 and Corollary 2.11, we have, in BIST, the implications LEM $\implies$ Sep $\implies$ Ind. Thus, by Proposition 2.18, each schema implies Con(HAH); whence, by Proposition 2.17, none is derivable in BIST+ Coll+ REM. □

Note that, in each case, the restriction of the schema to restricted properties *is* derivable.

Proposition 2.17 shows that BIST+ Coll is considerably weaker than ZF set theory. As well as BIST, we shall also be interested in the theory:

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

introduced in [11]. IST is closely related to Friedman's Intuitionistic Zermelo-Fraenkel set theory IZF [12]. On the one hand, by adding a set-induction principle and the axiom $\forall x.\, \mathsf{S}(x)$, one obtains IZF in its version with Replacement rather than Collection. (For all theories considered in this paper, we take Replacement as basic, and explicitly mention Collection when assumed.) Thus IZF is an extension of IST. Further, by relativizing quantifiers to an appropriately defined class of well-founded hereditary sets in IST, it is straightforward to interpret IZF in IST. These translations show that IST and IZF are of equivalent proof-theoretic strength. Similarly, IST+ Coll and IZF+ Coll have equivalent strength. It is known, [[REF]], that IZF+Coll, and hence IST+Coll, proves the same Π^0_2-sentences as classical ZF. It is also known, [[REF]], that IZF, and hence IST, is strictly weaker than ZF, with regard to Π^0_2-sentences. It is an open question whether IZF, and hence IST, proves the same Π^0_1 sentences as ZF.

We end this section with a brief discussion about the relationship between BIST and other intuitionistic set theories in the literature. To the best of our knowledge, none of the existing literature on weak set theories interpretable in elementary toposes [[REFS]] considers set theories with unrestricted Replacement or Collection axioms. In having such principles, our set theories are similar to the "constructive" set theories of Myhill, Friedman and Aczel [10,5,1,2]. However, because of our acceptance of the Powerset axiom, none of the set theories presented in this section are "constructive" in the sense of these authors. [12] In fact, in comparison with Aczel's CZF [1,2], the theory IST+ 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. On the other hand, Aczel's CZF has the full Ind schema, obtained as a consequence of a general set-induction principle. In contrast, for us, the full Ind schema is ruled out by Proposition 2.18.

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

PART II — TOPOSES

3 Toposes and systems of inclusions

In this section we introduce the categories we shall use as models of BIST$^-$ and the other theories. In this part of the paper, 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. For us, the default meta-theory will be BIST itself, although we shall work with it informally. Occasionally, it will be convenient to use stronger meta-theories, e.g. ZFC. We highlight whenever this is so.

We briefly recall the definition of elementary topos. 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:

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

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 object $\mathbf{1}$, a specified equalizer for every parallel pair, and specified data providing the powerobject structure as above.

Any morphism $f \colon A \longrightarrow B$ 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 $f \colon A \longrightarrow B$, we can factor the composite on the left below, to obtain the morphisms on the right.

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

Using the defining property of powerobjects, we obtain $\chi_{r_f} \colon \mathcal{P}(A) \to \mathcal{P}(B)$. We write $\mathcal{P}f \colon \mathcal{P}(A) \to \mathcal{P}(B)$ for χ_{r_f}. This morphism is intuitively 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*.

The main goal in this part of the paper, is to interpret the first-order language of Section 2 in an elementary topos $\mathcal{E}$. Moreover, we shall show that such interpretations always model the theory BIST$^-$.

There are many inequivalent ways of interpreting the first-order language in any given topos $\mathcal{E}$. Thus the interpretation needs to be defined with reference to additional structure on $\mathcal{E}$. The required 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. Indeed, the reader should henceforth bear in mind the equation:

$$\text{model of BIST}^- \;=\; \text{elementary topos } \mathcal{E} \;+\; \text{dssi on } \mathcal{E}\,. \tag{2}$$

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 $\mathcal{K}/A$.
(si4) Given a commuting diagram

$$\begin{tikzcd} A' \arrow[r, hook, "i"] & A \\ A'' \arrow[u, "m"] \arrow[ur, hook, "j"'] & \end{tikzcd} \tag{3}$$

with i, j inclusions, 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.

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

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

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

We say that $\mathcal{I}$ is a *partially-ordered system of inclusions* when the preorder on $\mathcal{I}$ is a partial order (i.e. when $A \equiv B$ implies $A = B$).

Proposition 3.3 (McLarty (REF)) *The following are equivalent.*

(1) $\mathcal{I}$ is a partially-ordered system of inclusions on $\mathcal{K}$.
(2) $\mathcal{I}$ is a subcategory of $\mathcal{K}$ satisfying (si1), (si2) *and also:*
(si3!) *For every mono $P \overset{m}{\rightarrowtail} A$ in $\mathcal{K}$ there exists a unique inclusion $A_m \hookrightarrow A$ that is isomorphic to m in $\mathcal{K}/A$.*

PROOF. $1 \implies 2$ is trivial. For the converse, we need to show that (si1), (si2) and (si3!) together imply: (i) that inclusions form a partial order, and (ii) that (si4) holds.

For (i), given inclusions $A \overset{i}{\hookrightarrow} B \overset{j}{\hookrightarrow} A$, we have $j = i^{-1}$, so j is isomorphic to 1_A in $\mathcal{K}/A$. Also, as 1_A is an identity, it is an inclusion/ Thus, by the uniqueness part of (si3!), $j = 1_A$, hence $A = B$.

For (ii), Suppose we have i, j, m as in diagram (3). Let $A'_m \overset{k}{\hookrightarrow} A'$ be the unique inclusion isomorphic to m in $\mathcal{K}/A'$. Then $i \circ k \colon A'_m \hookrightarrow A$ is isomorphic to $j \colon A'' \hookrightarrow A$ in $\mathcal{K}/A$. Hence, by the uniqueness part of (si3!), $A'_m = A''$ and $i \circ k = j = i \circ m$. Thus, as i is a mono, we have $k = m$, i.e. m is indeed an inclusion. $\square$

Given a (preordered) system of inclusions on a small category $\mathcal{K}$, there is a straightforward construction of a partially-ordered system of inclusions on a category $\mathcal{K}/\equiv$, whose objects are equivalence classes of objects of $\mathcal{K}$ under $\equiv$. Moreover, the evident quotient functor $Q \colon \mathcal{K} \to \mathcal{K}/\equiv$ is full, faithful, surjective on objects and preserves and reflects inclusions. This might suggest that there is little to choose between the preordered and partially-ordered definitions. Also, the motivating intuition that inclusion maps represent subset inclusions might encourage one to prefer the partial order version. However, the preordered notion is the more general and useful one when working in a weak meta-theory (such as BIST). It is more useful because many constructions of systems of inclusions, e.g. those in Part IV, naturally form preorders in the first instance. It is more general because, for a locally small category $\mathcal{K}$, additional assumptions on the meta-theory are required to construct the category $\mathcal{K}/\equiv$ above.[13] Moreover, even when $\mathcal{K}/\equiv$ does exist, the quotient functor $Q\colon \mathcal{K} \to \mathcal{K}/\equiv$ is, in general, only a weak equivalence.[14] Because of these issues, we henceforth work with the preordered notion of system of inclusions.

Definition 3.4 (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 objects A, B, there exists an object C_{AB} with $A \hookrightarrow C_{AB} \hookleftarrow B$).

Again, we shall always assume that a directed system of inclusions comes with a means of selecting an upper bound C_{AB} given A and B. This selection mechanism is not required to satisfy any additional coherence properties.

Proposition 3.5 *Suppose $\mathcal{I}$ is a directed system of inclusions on an elementary topos $\mathcal{E}$. Then:*

(1) The preorder $\mathcal{I}$ has finite joins. We write $\emptyset$ for a selected least element (the "empty set"), and $A \cup B$ for a selected binary join (the "union" of A and B).
(2) An object A of $\mathcal{E}$ is initial if and only if $A \equiv \emptyset$.
(3) The preorder $\mathcal{I}$ has binary meets. We write $A \cap B$ for a selected binary meet (the "intersection" of A and B).
(4) The square below is both a pullback and a pushout in $\mathcal{E}$.

$$\begin{tikzcd}
A \cap B \arrow[r, hook] \arrow[d, hook] & A \arrow[d, hook] \\
B \arrow[r, hook] & A \cup B
\end{tikzcd}$$

PROOF. First we construct $A \cup B$. Let C be such that $A \overset{i}{\hookrightarrow} C \overset{j}{\hookleftarrow} B$. We obtain the map $A + B \xrightarrow{[i,j]} C$. Define $A \cup_C B$ to be the object in the

[13] Because $\mathcal{K}$ is only locally small, the equivalence classes of objects under $\equiv$ may be proper classes, and there is no reason for a class of all equivalence classes to exist.
[14] A *weak equivalence* is a functor $F\colon \mathcal{K}_1 \to \mathcal{K}_2$ that is full, faithful and essentially surjective on objects, i.e. for every object $Y \in |\mathcal{K}_2|$ there exists $X \in |\mathcal{K}_1|$ with $FX \cong Y$. An *equivalence* requires, in addition, a functor $G\colon \mathcal{K}_2 \to \mathcal{K}_1$ such that GF and FG are naturally isomorphic to the identity functors on $\mathcal{K}_1$ and $\mathcal{K}_2$ respectively. Only in the presence of global choice is every weak equivalence an equivalence.

image factorization

$$A+B \xrightarrow{[i,j]} C \quad = \quad A+B \twoheadrightarrow A \cup_C B \hookrightarrow C\,.$$

Now suppose $C \xhookrightarrow{k} D$, so $A \xhookrightarrow{k \circ i} D \xhookleftarrow{k \circ j} B$. Define $A \cup_D B$ as above. But

$$\begin{aligned} A+B \xrightarrow{[k\circ i, k\circ j]} D &= A+B \xrightarrow{[i,j]} C \xhookrightarrow{k} D \\ &= A+B \twoheadrightarrow A \cup_C B \hookrightarrow C \hookrightarrow D \\ &= A+B \twoheadrightarrow A \cup_C B \hookrightarrow D\,. \end{aligned}$$

So, by the uniqueness of image factorization, the inclusions $A \cup_C B \hookrightarrow D$ and $A \cup_D B \hookrightarrow D$ are isomorphic in $\mathcal{K}/D$. So, by (si4), $A \cup_C B \equiv A \cup_D B$.

To define $A \cup B$, let C_{AB} be the specified object with $A \hookrightarrow C_{AB} \hookleftarrow B$. Define $A \cup B = A \cup_{C_{AB}} B$. To show this is a join, let C be such that $A \hookrightarrow C \hookleftarrow B$. By directedness, there exists D with $C_{AB} \hookrightarrow D \hookleftarrow C$. By the above, $A \cup B = A \cup_{C_{AB}} B \equiv A \cup_D B \equiv A \cup_C B$. Thus, $A \cup B \equiv A \cup_C B \hookrightarrow C$. So indeed $A \cup B \hookrightarrow C$.

We next show that for any two initial objects $\mathbf{0}, \mathbf{0}'$ of $\mathcal{E}$, we have $\mathbf{0} \equiv \mathbf{0}'$. By the above, we have an epi $\mathbf{0} + \mathbf{0}' \twoheadrightarrow \mathbf{0} \cup \mathbf{0}'$. But $\mathbf{0} + \mathbf{0}'$ is initial and, in any elementary topos, any image of an initial object is initial. Hence $\mathbf{0} \cup \mathbf{0}'$ is initial. Thus the inclusion $\mathbf{0} \hookrightarrow \mathbf{0} \cup \mathbf{0}'$ is an isomorphism and hence $\mathbf{0} \equiv \mathbf{0} \cup \mathbf{0}'$. Similarly, $\mathbf{0}' \equiv \mathbf{0} \cup \mathbf{0}'$. Thus indeed $\mathbf{0} \equiv \mathbf{0}'$.

Define $\emptyset$ to be a selected initial object. We must show that $\emptyset \hookrightarrow A$, for any A. Indeed, in a topos, the unique map $\emptyset \longrightarrow A$ is mono. Hence, by (si3), there exists an inclusion $\mathbf{0} \hookrightarrow A$ from some initial object $\mathbf{0}$. By the above, $\emptyset \equiv \mathbf{0} \hookrightarrow A$. Thus indeed $\emptyset \hookrightarrow A$. We have now proved (1) and (2)

To define $A \cap B$, construct the pullback below.

$$\begin{tikzcd} P \arrow[r, tail, "m"] \arrow[d, tail, "n"'] \arrow[dr, phantom, "\lrcorner", very near start] & A \arrow[d, hook, "i"] \\ B \arrow[r, hook, "j"'] & A \cup B \end{tikzcd}$$

Both m and n are mono because they are pullbacks of monos. Using (si3), define $A \cap B \xhookrightarrow{k} A$ to be the inclusion representative of m. Thus we have an isomorphism $A \cap B \xrightarrow{p} P$ with $m \circ p = k$. Then $i \circ k = j \circ n \circ p$, so, by (si4), $n \circ p$ is an inclusion $A \cap B \hookrightarrow B$. Moreover, as p is an isomorphism, we have the pullback square below.

$$\begin{tikzcd}
A \cap B \arrow[r, hook, "k"] \arrow[d, hook, "n \circ p"'] \arrow[dr, phantom, "\lrcorner", very near start] & A \arrow[d, hook, "i"] \\
B \arrow[r, hook, "j"'] & A \cup B
\end{tikzcd}$$

To see that $A \cap B$ is the meet of A and B, suppose that $A \hookleftarrow C \hookrightarrow B$. By (si2), this is a cone for the diagram $A \hookrightarrow A \cup B \hookleftarrow B$. The pullback above then gives a morphism $C \longrightarrow A \cap B$, which is an inclusion by (si4). This completes the proof of (3).

To prove (4), it remains only to show that the pullback is a pushout. But this holds because $A + B \xrightarrow{[i,j]} A \cup B$ is epi, by the definition of $A \cup B$, and, in a topos, any pullback of a jointly epic pair of monos is also a pushout. $\square$

Corollary 3.6 *Given a directed system of inclusions on an elementary topos, a (necessarily commuting) square of inclusions*

$$\begin{tikzcd}
A \arrow[r, hook] \arrow[d, hook] & B \arrow[d, hook] \\
C \arrow[r, hook] & D
\end{tikzcd}$$

is a pullback if and only if $A \equiv B \cap C$.

PROOF. By Proposition 3.5.1, $B \cup C \hookrightarrow D$. Using this, both implications follow easily from Proposition 3.5.4. $\square$

One of the motivations for condidering directed systems of inclusions is to be able to compare elements of different objects for equality. For objects A, B of $\mathcal{E}$, the relation $=_{A,B} \hookrightarrow A \times B$ is defined as the inclusion representative of the subobject obtained by pairing the inclusions $A \cap B \hookrightarrow A$ and $A \cap B \hookrightarrow B$. For any C with $A \overset{i}{\hookrightarrow} C \overset{j}{\hookleftarrow} B$, it holds in the internal logic of $\mathcal{E}$ that

$$x =_{A,B} y \quad \leftrightarrow \quad i(x) = j(y) \ .$$

The following lemma states that the relations $=_{A,B}$ form what might be called a *heterogeneous equality relation.*

Lemma 3.7 *For objects A, B, C, the following hold internally in $\mathcal{E}$:*

(1) $x =_{A,A} y$ *if and only if* $x = y$
(2) $x =_{A,B} y$ *implies* $y =_{B,A} x$.
(3) If $x =_{A,B} y$ *and* $y =_{B,C} z$ *then* $x =_{A,C} z$.

PROOF. Straightforward. $\square$

Definition 3.8 (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 $A \overset{f}{\underset{g}{\rightrightarrows}} B$, the specified equalizer $E \rightarrowtail A$ is an inclusion.
(ssi2) For all inclusions $A' \xhookrightarrow{i} A$ and $B' \xhookrightarrow{j} B$, the specified product $A' \times B' \xrightarrow{i \times j} A \times B$ 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 $A' \xhookrightarrow{i} A$, the direct-image map $\mathcal{P}A' \xrightarrow{\mathcal{P}i} \mathcal{P}A$ 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*). The lemma below is helpful for constructing dssi's.

Lemma 3.9 *Let $\mathcal{I}$ be a directed system of inclusions, on an elementary topos $\mathcal{E}$, satisfying property* (ssi4)*. Then it is possible to respecify the topos structure on $\mathcal{E}$ so that $\mathcal{I}$ is a dssi with respect to the new structure.*

PROOF. For (ssi1), given a parallel pair $f, g \colon A \longrightarrow B$, let $e \colon E \rightarrowtail A$ be the equalizer originally specified. The newly specified equalizer is simply defined to be the specified inclusion $A_e \hookrightarrow A$ representing e.

For (ssi2), we specify a new product $A \times' B$ using Kuratowski pairing. Recall from [REF], that Kuratowski pairing gives a monic natural transformation

$$\mathrm{kpr}_X \;=\; X \times X \xrightarrow{(x,y) \mapsto \{\{x\},\{x,y\}\}} \mathcal{P}^2 X\,.$$

Thus, for any A, B, using the inclusions $A \xhookrightarrow{i} A \cup B \xhookleftarrow{j} B$, we have a mono

$$m_{AB} \;=\; A \times B \xrightarrow{i \times j} (A \cup B) \times (A \cup B) \xrightarrow{\mathrm{kpr}_{A \cup B}} \mathcal{P}^2(A \cup B)\,.$$

Define $A \times' B$ to be the domain of the inclusion $A \times' B \overset{p_{AB}}{\hookrightarrow} \mathcal{P}^2(A \cup B)$ that represents the mono m_{AB}. Thus we have a unique isomorphism

$$A \times' B \xrightarrow{i_{AB}} A \times B$$

such that $p_{AB} = m_{AB} \circ i_{AB}$. The projections $A \xleftarrow{\pi'_1} A \times' B \xrightarrow{\pi'_2} B$ are defined by $\pi'_i = \pi_i \circ i_{AB}$. This is a product diagram because i_{AB} is an iso. One easily verifies that, given $f\colon A \longrightarrow A'$ and $g\colon B \longrightarrow B'$, then the product morphism $f \times' g\colon A \times' B \longrightarrow A' \times' B'$ is the unique morphism satisfying $i_{A'B'} \circ (f \times' g) = (f \times g) \circ i_{AB}$.

We now show that (ssi2) holds. Suppose then that f, g are inclusions. We must show that $f \times' g$ is an inclusion. As f, g are inclusions, we have an inclusion $A \cup B \overset{k}{\hookrightarrow} A' \cup B'$. Thus the diagram below commutes.

$$\begin{tikzcd}
A \times' B \arrow[r, "i_{AB}"] \arrow[d, "f \times' g"'] & A \times B \arrow[r, "i \times j"] \arrow[d, "f \times g"'] & (A \cup B) \times (A \cup B) \arrow[r, "\mathrm{kpr}_{A \cup B}"] \arrow[d, "k \times k"] & \mathcal{P}^2(A \cup B) \arrow[d, "\mathcal{P}^2(k)"] \\
A' \times' B' \arrow[r, "i_{A'B'}"] & A' \times B' \arrow[r, "i' \times j'"] & (A' \cup B') \times (A' \cup B') \arrow[r, "\mathrm{kpr}_{A' \cup B'}"] & \mathcal{P}^2(A' \cup B')
\end{tikzcd}$$

(The middle square commutes because f, g, i, j, i', j', k are all inclusions; the right-hand square by the naturality of kpr.) The diagram expresses the equation $\mathcal{P}^2(k) \circ p_{AB} = p_{A'B'} \circ (f \times' g)$. But p_{AB} and $p_{A'B'}$ are inclusions. Moreover, by (ssi4), $\mathcal{P}^2(k)$ is also an inclusion. Thus $f \times' g$ is indeed an inclusion, by (si4).

Finally, we need to respecify the powerobject structure on $\mathcal{E}$ consistently with the new product $A \times' B$, and check that (ssi4) remains true. In fact, the object $\mathcal{P}(A)$ remains unchanged. The membership mono $\ni'_A \hookrightarrow \mathcal{P}(A) \times' A$ is defined as the inclusion representative of the mono

$$\ni_A \rightarrowtail \mathcal{P}(A) \times A \overset{i^{-1}_{\mathcal{P}(A)A}}{\rightarrowtail} \mathcal{P}(A) \times' A\,.$$

We have thus satisfied (ssi3). Moreover, one readily checks that, with this redefinition, the action of the covariant powerobject functor remains unaffected. Thus (ssi4) still holds. $\square$

We make some basic observations concerning the existence of dssis. First, we observe that not every topos can have a dssi placed upon it. For a simple counterexample, using ZFC as the meta-theory, 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**, we see that the existence of a dssi is not preserved under equivalence of categories. Nevertheless, every topos is equivalent to one carrying a dssi.

Theorem 3.10 (BIST+ REM) *Given a topos $\mathcal{E}$, there exists an equivalent category $\mathcal{E}'$ carrying a dssi $\mathcal{I}'$ relative to specified topos structure on $\mathcal{E}'$.*

By showing that here is no loss in generality in working with toposes carrying sdsi's, this theorem is essential for placing the various constructions in Parts II–III of the paper that rely on the presence of systems of inclusions in context. Nevertheless, in spite of its importance, we postpone the proof of the theorem, which is rather technical, to Part IV. The reader will have also noticed that Theorem 3.10, as stated, assumes Restricted Excluded Middle in the meta-theory. In Part IV, we shall obtain a sharper version, which merely relies on BIST as the meta-theory. Again the precise formulation of this is somewhat technical, see Proposition 11.14 for details.

We next establish some basic properties of an elementary topos $\mathcal{E}$ with a dssi $\mathcal{I}$. These properties will be useful in Sections 4 and 8.

Proposition 3.11 *Let $\mathcal{I}$ be a dssi on an elementary topos $\mathcal{E}$. Then:*

(1) $(A \times B) \cap (A' \times B') \;\equiv\; (A \cap A') \times (B \cap B')$;
(2) $(\mathcal{P}A) \cap (\mathcal{P}B) \;\equiv\; \mathcal{P}(A \cap B)$.

PROOF. For 1, the square below is a pullback, because, by Proposition 3.5.4 and (ssi2), it is a product of pullback squares.

$$\begin{tikzcd}
(A \cap A') \times (B \cap B') \arrow[r, hook] \arrow[d, hook] & A \times B \arrow[d, hook] \\
A' \times B' \arrow[r, hook] & (A \cup A') \times (B \cup B')
\end{tikzcd}$$

Thus, by Corollary 3.6, $(A \times B) \cap (A' \times B') \equiv (A \cap A') \times (B \cap B')$.

Similarly, for 2, the square below is a pullback, by Proposition 3.5.4 and (ssi4), because the covariant powerobject functor preserves pullbacks of monos.

$$\begin{tikzcd}
\mathcal{P}(A\cap B) \arrow[r, hook] \arrow[d, hook] & \mathcal{P}A \arrow[d, hook] \\
\mathcal{P}B \arrow[r, hook] & \mathcal{P}(A\cup B)
\end{tikzcd}$$

Again, by Corollary 3.6, $(\mathcal{P}A)\cap(\mathcal{P}B)\equiv\mathcal{P}(A\cap B)$. $\square$

In an elementary topos with dssi, a map $\langle s,t\rangle\colon X\longrightarrow\mathcal{P}A\times A$ factors through $\ni_A\hookrightarrow\mathcal{P}A\times A$ if and only if $\mathrm{Im}\,\langle s,t\rangle\hookrightarrow\ni_A$. Furthermore, if $i\colon A\hookrightarrow B$ then $\mathcal{P}i\times i\colon\mathcal{P}A\times A\hookrightarrow\mathcal{P}B\times B$ and hence $\ni_A\hookrightarrow\ni_B$.

Proposition 3.12 *Let $\mathcal{I}$ be an dssi on an elementary topos $\mathcal{E}$. Suppose that $\langle s,t\rangle\colon X\longrightarrow\mathcal{P}A\times A$ factors through $\ni_A\hookrightarrow\mathcal{P}A\times A$ and that* $\mathrm{Im}(s)\hookrightarrow\mathcal{P}B$*. Then* $\mathrm{Im}(t)\hookrightarrow A\cap B$ *and* $\mathrm{Im}\,\langle s,t\rangle\hookrightarrow\ni_{A\cap B}$.

PROOF. Given any $X\xrightarrow{\langle s,t\rangle}\mathcal{P}A\times A$ where $\mathrm{Im}(s)\hookrightarrow\mathcal{P}B$, trivially also $\mathrm{Im}(s)\hookrightarrow\mathcal{P}A$. So, by Proposition 3.11(2), $\mathrm{Im}(s)\hookrightarrow\mathcal{P}(A\cap B)$. Thus, $\langle s,t\rangle$ is given by the bottom-left composite below

$$\begin{tikzcd}
X \arrow[r, dotted] \arrow[dd, "{\langle e_s,t\rangle}"'] & \ni_{A\cap B} \arrow[rr, hook] \arrow[d, hook] \arrow[drr, phantom, "\lrcorner", very near start] & & \ni_A \arrow[dd, hook] \\
& \mathcal{P}(A\cap B)\times(A\cap B) \arrow[d, hook] & & \\
\mathrm{Im}(s)\times A \arrow[r, hook] & \mathcal{P}(A\cap B)\times A \arrow[rr, hook] & & \mathcal{P}A\times A
\end{tikzcd}$$

The right-hand rectangle is a pullback, because the inclusion $\mathcal{P}(A\cap B)\times A\hookrightarrow\mathcal{P}A\times A$ is obtained as $\mathcal{P}(i)\times 1_A$, where $i\colon A\cap B\hookrightarrow B$. Now suppose $\langle s,t\rangle$ factors through $\ni_A\hookrightarrow\mathcal{P}A\times A$. Then, by the pullback property, $\langle s,t\rangle$ factors via a map $X\longrightarrow\ni_{A\cap B}$. So indeed $\mathrm{Im}\,\langle s,t\rangle\hookrightarrow\ni_{A\cap B}$. Moreover, because the left-hand rectangle above commutes, also $\mathrm{Im}(t)\hookrightarrow A\cap B$. $\square$

Proposition 3.13 *If $A\hookrightarrow\mathcal{P}B$ then the collection $\{C\mid A\hookrightarrow\mathcal{P}C\}$ has a least element under the $\hookrightarrow$ relation.*

PROOF. Suppose that $A \hookrightarrow \mathcal{P}B$. Then the inclusion map is characteristic for a relation $R \hookrightarrow B \times A$ and define $\bigcup A$ to be the image factorization of $R \hookrightarrow B \times A \xrightarrow{\pi_1} B$. It is easily checked that $C = (\bigcup A) \hookrightarrow B$ is the required least element. $\square$

It will be useful in Section 8 to have a definition of coproduct that interacts well with the inclusion structure on $\mathcal{E}$. Define:

$$\begin{aligned} A + B \;=\; \{(X, Y) \colon \mathcal{P}A \times \mathcal{P}B \mid & ((\exists x \colon A.\, X = \{x\}) \wedge Y = \emptyset) \\ & \vee\; (X = \emptyset \wedge (\exists y \colon B.\, Y = \{y\}))\}\;. \end{aligned}$$

The injections are given by the maps

$$\begin{aligned} x &\mapsto (\{x\}, \emptyset) \colon A \longrightarrow A + B \\ y &\mapsto (\emptyset, \{y\}) \colon B \longrightarrow A + B\;. \end{aligned}$$

It is routine to verify that this indeed defines a coproduct.

Proposition 3.14 *The coproduct defined above enjoys the properties below.*

(1) If $A' \hookrightarrow A$ and $B' \hookrightarrow B$ then $A' + B' \hookrightarrow A + B$.
(2) If $C \hookrightarrow A + B$ then $C \equiv A' + B'$ for some $A' \hookrightarrow A$ and $B' \hookrightarrow B$.
(3) $(A + B) \cap (A' + B') \equiv (A \cap A') + (B \cap B')$.
(4) $(A + B) \cup (A' + B') \equiv (A \cup A') + (B \cup B')$.

PROOF. We just verify statement 2. Suppose $C \hookrightarrow A + B$. Define A', B' by pullback as below.

$$\begin{tikzcd}
A' \arrow[r] \arrow[d, hook] \arrow[dr, phantom, "\lrcorner", very near start] & C \arrow[d, hook] & B' \arrow[l] \arrow[d, hook] \arrow[dl, phantom, "\llcorner", very near start] \\
A \arrow[r, "\mathrm{inl}"] & A + B & B \arrow[l, "\mathrm{inr}"']
\end{tikzcd}$$

By statement 1, there is an inclusion $A' + B' \hookrightarrow A + B$. By the stability of coproducts in $\mathcal{E}$, the top edge in the diagram above is also a coproduct diagram. Thus the inclusions $C \hookrightarrow A + B$ and $A' + B' \hookrightarrow A + B$ factor through each other. Thus indeed $A' + B' \equiv C$. $\square$

We end this section with a discussion of the extra structure we shall require to interpret IST and other set theories with full Separation.

Definition 3.15 (Superdirected system of inclusions) A system of inclusions $\mathcal{I}$ on a category $\mathcal{K}$ 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 3.16 *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 $\mathbf{1}$.*

PROOF. As $\mathcal{E}$ is small, it has a set of objects. Hence, because $\mathcal{I}$ is superdirected, $\mathcal{I}$ has a greatest element U. Then $\mathcal{P}U \hookrightarrow U$, so $\mathcal{P}U \rightarrowtail U$. One can now mimic Russell's paradox in U to derive the inconsistency of the internal logic of $\mathcal{E}$. Thus every morphism in $\mathcal{E}$ is an isomorphism, and hence $A \equiv U$ for every object A, including $\mathbf{1}$. The result follows. $\square$

Thus sdssis are only interesting on locally small toposes whose objects form a proper class.

Proposition 3.17 *Suppose that $\mathcal{I}$ is an sdssi on a topos $\mathcal{E}$. Consider the statements below.*

(1) $\mathcal{E}$ is cocomplete.[15]
(2) The preorder $\mathcal{I}$ has small joins.
(3) For every object A, the subobject lattice is a complete Heyting algebra.

Then $1 \implies 2 \implies 3$.

PROOF. That $1 \implies 2$ follows by a straightforward generalization of the proof of Proposition 3.5.1. For the proof of $2 \implies 3$, assume that $\mathcal{I}$ has small joins. Let $\{P_i \overset{m_i}{\rightarrowtail} A\}_{i \in I}$ be any small family of monos. Consider the corresponding family of inclusions $\{A_{m_i} \hookrightarrow A\}_{i \in I}$. Then, using 2, we obtain $(\bigcup_{i \in I} A_{m_i}) \hookrightarrow A$, which represents the join of $\{P_i \overset{m_i}{\rightarrowtail} A\}_{i \in I}$ in the subobject lattice. $\square$

It is easy to see that the implication $2 \implies 3$ cannot be reversed. For a counterexample, take any non-trivial full subcategory of **Set**, e.g. in ZFC, the category of all finite sets, with inclusion maps given by subset inclusions. We

[15] A category $\mathcal{K}$ is *cocomplete* if every small diagram has a colimit. As it has coequalizers, a topos is cocomplete if and only if it has small coproducts.

do not, at present, know any example of a topos carrying an sdssi with small joins that is not cocomplete.

Cocompleteness is an in important condition in relation to the existence of sdssi's, as it is a sufficient condition for obtaining an analogue of Theorem 3.10.

Theorem 3.18 *For any cocomplete topos $\mathcal{E}$, there is an equivalent category $\mathcal{E}'$ carrying an sdssi $\mathcal{I}'$ relative to specified topos structure on $\mathcal{E}'$.*

However, as the next result shows, cocompleteness is not a necessary condition for the existence of an sdssi. Our proof uses ZFC as the meta-theory.

Theorem 3.19 (ZFC) *For any realizability topos $\mathcal{E}$, there is an equivalent category $\mathcal{E}'$ carrying an sdssi $\mathcal{I}'$ relative to specified topos structure on $\mathcal{E}'$.*

The proofs of Theorems 3.18 and 3.19 will be given in Sections 11 and 12 respectively.

4 Interpreting set theory in a topos with inclusions

In this section we give an interpretation of the first-order language of Section 2 in an arbitrary elementary topos with dssi. We show that this interpretation validates the axioms of BIST^{-}+ Coll. Moreover, the axiom of Infinity (hence BIST + Coll) is validated if (and only if) the topos has a natural numbers object. We exploit these general soundness results to establish the various non-derivability claims of Section 2. We also state a corresponding completeness result, which will be proved in Part III.

For the entirety of this section, 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 [[REFS]], but with two main differences. First, we have to interpret the untyped relations $\mathsf{S}(x)$, $x = y$ 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 [**?**], 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

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

Fig. 6. The forcing relation

write $X \Vdash_\rho \phi$ for the associated "forcing" relation, which is defined inductively in Figure 6. 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.

It is immediate from the definition of the forcing relation that the relation $X \Vdash_\rho \phi$ depends only on the value of ρ on variables that appear free in ϕ. The next few lemmas establish other straightforward properties.

Lemma 4.1 *For any $Y \xrightarrow{t} X$, if $X \Vdash_\rho \phi$ then $Y \Vdash_{\rho \circ t} \phi$.*

PROOF. An easy induction on the structure of ϕ. $\square$

Lemma 4.2 *For any finite jointly epic family $Y_1 \xrightarrow{t_1} X, \ldots, Y_k \xrightarrow{t_k} X$, if $Y_1 \Vdash_{\rho \circ t_1} \phi$ and $\ldots$ and $Y_k \Vdash_{\rho \circ t_k} \phi$ then $X \Vdash_\rho \phi$.*

PROOF. We first make the following observation. For each variable $x \in \mathrm{dom}(\rho)$, the map $\rho_x \circ t_i \colon Y_i \longrightarrow A_x$ factors as $Y_i \xrightarrow{e'_{x,i}} I'_{x,i} \hookrightarrow A_x$. Thus there is a commuting diagram:

$$\begin{tikzcd}
Y_1 + \cdots + Y_k \arrow[r, two heads, "e'_{x,1} + \cdots + e'_{x,k}"] \arrow[d, "t_1 + \cdots + t_k"'] & I'_{x,1} + \cdots + I'_{x,k} \arrow[r, two heads] & I'_{x,1} \cup \cdots \cup I'_{x,k} \arrow[d, hook] \\
X \arrow[rr, two heads, "e_x"'] & & I_x \arrow[r, hook, "i_x"'] & A_x
\end{tikzcd}$$

where the left edge is epi because the t_i are jointly epi. Thus, by the uniqueness of image factorizations, we have $I_x \equiv I'_{x,1} \cup \cdots \cup I'_{x,k}$.

The proof now proceeds by induction on the structure of ϕ. We give one case, to illustrate the style of argument.

If ϕ is $x \in y$ then, by assumption, for each i, there exists B'_i, with $I'_{x,i} \xhookrightarrow{i'_i} B'_i$ and $I'_{y,i} \xhookrightarrow{j'_i} \mathcal{P}B'_i$, such that $\langle j'_i \circ e'_{y,i}, i'_i \circ e'_{x,i}\rangle$ factors through $\ni_{B'_i}$. Thus $I_x \equiv I'_{x,1} \cup \cdots \cup I'_{x,k} \hookrightarrow B'_1 \cup \cdots \cup B'_k$ and $I_y \equiv I'_{y,1} \cup \cdots \cup I'_{y,k} \hookrightarrow \mathcal{P}B'_1 \cup \cdots \cup \mathcal{P}B'_k \hookrightarrow \mathcal{P}(B'_1 \cup \cdots \cup B'_k)$. So, defining $B = B'_1 \cup \cdots \cup B'_k$, we have $i \colon I_x \hookrightarrow B$ and $j \colon I_y \hookrightarrow \mathcal{P}B$. We must show that $\langle j \circ e_y, i \circ e_x\rangle \colon X \longrightarrow \mathcal{P}B \times B$ factors through $\ni_B$. Reasoning internally in $\mathcal{E}$, take any $a \colon X$. We must show that $i(e_x(a)) \in j(e_y(a))$. As the t_i are jointly epi, there exist i and $b \colon Y_i$ such that $a = t_i(b)$. By assumption, $i'_i(e'_{x,i}(b)) \in j'_i(e'_{y,i}(b))$ (where $i'_i(e'_{x,i}(b)) \colon B'_i$ and $j'_i(e'_{y,i}(b)) \colon \mathcal{P}B'_i$). By the definition of $e'_{x,i}$ and $e'_{y,i}$, the inclusion $B'_i \hookrightarrow B$ maps $i'_i(e'_{x,i}(b))$ to $i(e_x(t_i(b))) = i(e_x(a))$, and the inclusion $\mathcal{P}B'_i \hookrightarrow \mathcal{P}B$ maps $j'_i(e'_{y,i}(b))$ to $j(e_y(t_i(b))) = j(e_y(a))$. Also $\ni_{B'_i} \hookrightarrow \ni_B$, by the remarks above Proposition 3.12. So indeed $i(e_x(a)) \in j(e_y(a))$. $\square$

Lemma 4.3 *Given inclusions $A_x \xhookrightarrow{i_x} B_x$, for all free x in ϕ, it holds that $X \Vdash_\rho \phi$ if and only if $X \Vdash_{i \circ \rho} \phi$.*

PROOF. Straightforward induction on the structure of ϕ. $\square$

The lemma below establishes a convenient property of the forcing semantics of the membership relation.

Lemma 4.4 *If $X \Vdash_\rho x \in y$ and $I_y \xhookrightarrow{j} \mathcal{P}B$ then there exists an inclusion $I_x \xhookrightarrow{i} B$ such that $\langle j \circ e_y,\, i \circ e_x\rangle \colon X \longrightarrow \mathcal{P}B \times B$ factors through $\ni_B$.*

PROOF. Suppose $X \Vdash_\rho x \in y$. Then there exist $i' \colon I_x \hookrightarrow A$ and $j' \colon I_y \hookrightarrow \mathcal{P}A$ such that $\langle j' \circ e_y,\, i' \circ e_x\rangle \colon X \longrightarrow \mathcal{P}A \times A$ factors through $\ni_A$. Suppose also $I_y \xhookrightarrow{j} \mathcal{P}B$. Then, by Proposition 3.12, $I_x \equiv \mathrm{Im}(i' \circ e_x) \xhookrightarrow{j} A \cap B$ and $\mathrm{Im}\langle j' \circ e_y,\, i' \circ e_x\rangle \hookrightarrow \ni_{A\cap B}$. However, $\mathrm{Im}\langle j' \circ e_y,\, i' \circ e_x\rangle \equiv \mathrm{Im}\langle e_y,\, e_x\rangle \equiv \mathrm{Im}\langle j \circ e_y,\, i \circ e_x\rangle$. So $\mathrm{Im}\langle j \circ e_y,\, i \circ e_x\rangle \hookrightarrow \ni_{A\cap B} \hookrightarrow \ni_B$. Thus indeed, $\langle j \circ e_y,\, i \circ e_x\rangle$ factors through $\ni_B$. $\square$

The next lemma gives a direct formulation of the derived forcing conditions for the various abbreviations introduced into the set theoretic language.

Lemma 4.5 *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 $Y' \Vdash_{(\rho \circ t')[s'/x]} \phi$

iff $Y \Vdash_{(\rho\circ t)[s/x]} \phi$, *where* $Y = \{(c, a) \colon C \times X \mid c \in k(e_z(a))\}$
and $s \colon Y \longrightarrow C$ *and* $t \colon Y \longrightarrow X$ *are the projections.*

$X \Vdash_\rho \exists x \in z.\, \phi$ *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*
$\subseteq_B \hookrightarrow \mathcal{P}B \times \mathcal{P}B$

$X \Vdash_\rho \text{Ɂ}x.\, \phi$ *iff there exist objects* B *and* $R \hookrightarrow X \times B$ *such that,*
for all objects 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.*

PROOF. We include two cases: the characterization of $X \Vdash_\rho \text{Ɂ}x.\, \phi$, for which we give the proof in detail (this is the most intricate case), and the characterization of $X \Vdash_\rho !\phi$, for which we outline the argument.

Suppose $X \Vdash_\rho \text{Ɂ}x.\, \phi$, i.e. $X \Vdash_\rho \exists y.\, (\mathsf{S}(y) \wedge \forall x.\, (x \in y \leftrightarrow \phi))$, where y is not free in ϕ. Then there exist $t\colon Y \twoheadrightarrow X$ and $s\colon Y \longrightarrow A$ such that $Y \Vdash_{(\rho\circ t)[s/y]} \mathsf{S}(y)$ and

$$Y \Vdash_{(\rho\circ t)[s/y]} \forall x.\,(x \in y \leftrightarrow \phi)\,. \tag{4}$$

By Lemma 4.3, one can, without loss of generality, assume that s is also epi. Thus there exists B such that $i\colon A \hookrightarrow \mathcal{P}B$. Define

$$R = \{(a,b)\colon X \times B \mid \exists c\colon Y.\, t(c) = a \,\wedge\, b \in i(s(c))\}\,.$$

Take any maps $Y' \xrightarrow{t'} X$ and $Y' \xrightarrow{s'} A'$. We must show that $Y' \Vdash_{(\rho\circ t')[s'/x]} \phi$ iff $\mathrm{Im}\,\langle t', s'\rangle \hookrightarrow R$. Moreover, by Lemma 4.3, we can, without loss of generality, assume that s' is epi.

First, note that, for any commuting diagram:

$$\begin{tikzcd} Y'' \arrow[r, "r"] \arrow[d, "r'"'] & Y \arrow[d, "t"] \\ Y' \arrow[r, "t'"'] & X \end{tikzcd} \tag{5}$$

in which r' is epi, we have

$$\begin{array}{llll} Y' \Vdash_{(\rho\circ t')[s'/x]} \phi & \text{iff } Y'' \Vdash_{(\rho\circ t'\circ r')[s'\circ r'/x]} \phi & & \text{(by Lemmas 4.1 and 4.2)} \\ & \text{iff } Y'' \Vdash_{(\rho\circ t\circ r)[s'\circ r'/x]} \phi & & \\ & \text{iff } Y'' \Vdash_{(\rho\circ t\circ r)[s\circ r/y,\, s'\circ r'/x]} \phi & & \text{(as } y \text{ is not free in } \phi) \\ & \text{iff } Y'' \Vdash_{(\rho\circ t\circ r)[s\circ r/y,\, s'\circ r'/x]} x \in y & & \text{(by (4) above).} \end{array}$$

To show that $Y' \Vdash_{(\rho\circ t')[s'/x]} \phi$ implies $\mathrm{Im}\,\langle t', s'\rangle \hookrightarrow R$, construct Y'', r, r', as in Diagram (5), by taking the pullback of t along t'. Suppose $Y' \Vdash_{(\rho\circ t')[s'/x]} \phi$. By the above equivalences, $Y'' \Vdash_{(\rho\circ t\circ r)[s\circ r/y,\, s'\circ r'/x]} x \in y$. Since there are inclusions $\mathrm{Im}(s\circ r) \hookrightarrow A \hookrightarrow \mathcal{P}B$, it follows from Lemma 4.4 that $\mathrm{Im}(s'\circ r') \hookrightarrow B$ and $\mathrm{Im}\,\langle s\circ r,\, s'\circ r'\rangle \hookrightarrow\, \ni_B$. However, $A' \equiv \mathrm{Im}(s'\circ r')$ because s' and r' are epi, so we have $j\colon A' \hookrightarrow B$, and hence $\mathrm{Im}\,\langle t'\circ r', s'\circ r'\rangle \hookrightarrow X \times B$. We show that this inclusion factors through the subobject R. Reasoning internally in $\mathcal{E}$, take any $d\colon Y''$. Define $c = r(d)$. Then $t(c) = t'(r'(d))$. Above, we saw that $\mathrm{Im}\,\langle s\circ r,\, s'\circ r'\rangle \hookrightarrow\, \ni_B$, so $\langle i\circ s\circ r,\, j\circ s'\circ r'\rangle$ factors through $\ni_B$, hence $j(s'(r'(d))) \in_B i(s(r(d))) = i(s(c))$. This establishes that $\mathrm{Im}\,\langle t'\circ r', s'\circ r'\rangle \hookrightarrow R$. It follows that $\mathrm{Im}\,\langle t', s'\rangle \hookrightarrow R$, because r' is epi.

Conversely, suppose that $\mathrm{Im}\,\langle t', s'\rangle \hookrightarrow R$. As $R \hookrightarrow X \times B$ and $A' \equiv \mathrm{Im}(s')$, there exists $j\colon A' \hookrightarrow B$. Define Y'' by

$$Y'' \;=\; \{(c', c)\colon Y' \times Y \mid t(c) = t'(c') \;\wedge\; j(s'(c')) \in i(s(c))\},$$

and write $r'\colon Y'' \longrightarrow Y'$ and $r\colon Y'' \longrightarrow Y$ for the two projections. Trivially $t' \circ r' = t \circ r$. Also, because $\mathrm{Im}\,\langle t', s'\rangle \hookrightarrow R$, it follows from the definition of R that r' is epi. It is immediate from the definition of Y'' that $Y'' \Vdash_{(\rho \circ t \circ r)[s \circ r/y,\, s' \circ r'/x]} x \in y$. Hence, by the equivalences below Diagram 5, indeed $Y' \Vdash_{(\rho \circ t')[s'/x]} \phi$.

To prove the right-to-left implication of the characterization of $X \Vdash_\rho \mathsf{S}x.\,\phi$, suppose there exist B and $R \hookrightarrow X \times B$ with the properties in the statement of the lemma. We must show that $X \Vdash_\rho \exists y.\,(\mathsf{S}(y) \wedge \forall x.\,(x \in y \leftrightarrow \phi))$, where y is not free in ϕ. Define $r\colon X \longrightarrow \mathcal{P}B$ by

$$r(x) \;=\; \{y\colon B \mid (x, y) \in R\}.$$

we show that $X \Vdash_{\rho[r/y]} \mathsf{S}(y) \wedge \forall x.\,(x \in y \leftrightarrow \phi)$. Trivially, $X \Vdash_{\rho[r/y]} \mathsf{S}(y)$. To show that $X \Vdash_{\rho[r/y]} \forall x.\,(x \in y \leftrightarrow \phi)$, consider any $t\colon Y \longrightarrow X$ and $s\colon Y \longrightarrow A$. We must show that $X \Vdash_{(\rho[r/y] \circ t)[s/x]} x \in y$ iff $X \Vdash_{(\rho[r/y] \circ t)[s/x]} \phi$ By Lemma 4.3, we can assume that s is epi. Also, $X \Vdash_{(\rho[r/y] \circ t)[s/x]} \phi$ iff $X \Vdash_{(\rho \circ t)[s/x]} \phi$ (because y is not free in ϕ), iff $\mathrm{Im}\,\langle t, s\rangle \hookrightarrow R$ (by the main assumption). It thus suffices to show that $\mathrm{Im}\,\langle t, s\rangle \hookrightarrow R$ iff $X \Vdash_{(\rho[r/y] \circ t)[s/x]} x \in y$. For the left-to-right implication, suppose that $\mathrm{Im}\,\langle t, s\rangle \hookrightarrow R$. As $R \hookrightarrow X \times B$ and s is epi, we have $A \equiv \mathrm{Im}(s) \hookrightarrow B$. Also, by the definition of r, it holds that $\mathrm{Im}(r \circ t) \hookrightarrow \mathcal{P}B$ and $\langle r \circ t, s\rangle$ factors through $\ni_B \hookrightarrow \mathcal{P}B \times B$. Thus indeed $X \Vdash_{(\rho \circ t)[r \circ t/y,\, s/x]} x \in y$. Conversely, suppose $X \Vdash_{(\rho \circ t)[r \circ t/y,\, s/x]} x \in y$. As $\mathrm{Im}(r \circ t) \hookrightarrow \mathcal{P}B$ and s is epi, it follows from Lemma 4.4 that $i\colon A \equiv \mathrm{Im}(s) \hookrightarrow B$ and $\langle r \circ t, i \circ s\rangle$ factors through $\ni_B \hookrightarrow \mathcal{P}B \times B$. By the definition of r, it follows that $\langle t, i \circ s\rangle$ factors through $R \hookrightarrow X \times B$, i.e. $\mathrm{Im}\,\langle t, i \circ s\rangle \hookrightarrow R$. Thus indeed, $\mathrm{Im}\,\langle t, s\rangle \equiv \mathrm{Im}\,\langle t, i \circ s\rangle \hookrightarrow R$.

We now turn to the characterization of $X \Vdash_\rho {!\phi}$. First, an auxiliary remark. For any X, ρ, it is easily shown that $X \Vdash_\rho x = \emptyset$ iff $I_x \hookrightarrow \{\emptyset\}$, where we write $\{\emptyset\}$ for the object $\mathcal{P}\emptyset$ of $\mathcal{E}$.

Now suppose that $X \Vdash_\rho {!\phi}$, in other words that $X \Vdash_\rho \mathsf{S}z.\,(z = \emptyset \wedge \phi)$, where z is not free in ϕ. Thus there exists $R \hookrightarrow X \times B$ such that, for all $Y \xrightarrow{t} X$ and $Y \xrightarrow{s} A$, it holds that $Y \Vdash_{(\rho \circ t)[s/z]} z = \emptyset \wedge \phi$ iff $\mathrm{Im}\,\langle t, s\rangle \hookrightarrow R$. Using the remark above, one shows that $R \equiv R \cap (X \times \{\emptyset\})$. The we can define $i_0\colon Y_0 \hookrightarrow X$ by

$$Y_0 \;=\; \{a\colon X \mid (a, \emptyset) \in R\}.$$

We show that (i) $Y_0 \Vdash_{\rho \circ i_0} \phi$, and (ii) for any $i\colon Y \hookrightarrow X$ such that $Y \Vdash_{\rho \circ i} \phi$, it holds that $Y \hookrightarrow Y_0$. Property (i) holds because $\mathrm{Im}\,\langle i_0, \emptyset\rangle \hookrightarrow R$. For property (ii), suppose $i\colon Y \hookrightarrow X$ is such that $Y \Vdash_{\rho \circ i} \phi$, Then, by the earlier remark, $Y \Vdash_{(\rho \circ i)[\emptyset/z]} z = \emptyset \wedge \phi$, Hence $\mathrm{Im}\,\langle i, \emptyset\rangle \hookrightarrow R$. Thus $Y \hookrightarrow Y_0$ by the definition of Y_0.

Conversely, suppose there exists $Y_0 \hookrightarrow X$ such that (i) and (ii) above hold. We must show that $X \Vdash_\rho \text{Ƨ}z.\,(z = \emptyset \wedge \phi)$. Defining $R = Y_0 \times \{\emptyset\}$, we show that, for all $Z \xrightarrow{t} X$ and $Z \xrightarrow{s} A$, it holds that $Z \Vdash_{(\rho \circ t)[s/z]} z = \emptyset \wedge \phi$ iff $\mathrm{Im}\,\langle t, s\rangle \hookrightarrow R$. Take any $Z \xrightarrow{t} X$ and $Z \xrightarrow{s} A$, and let $Z \twoheadrightarrow Y \xhookrightarrow{i} X$ be the image factorization of t. Suppose $Z \Vdash_{(\rho \circ t)[s/z]} z = \emptyset \wedge \phi$. Then $\mathrm{Im}(s) \hookrightarrow \{\emptyset\}$, by the earlier remark, and $Y \Vdash_{\rho \circ i} \phi$, by Lemma 4.2. Thus $Y \hookrightarrow Y_0$, by (ii). So indeed, $\mathrm{Im}\,\langle t, s\rangle \hookrightarrow Y_0 \times \{\emptyset\} = R$. Conversely, suppose that $\mathrm{Im}\,\langle t, s\rangle \hookrightarrow Y_0 \times \{\emptyset\}$. Then $Z \Vdash_{(\rho \circ t)[s/z]} z = \emptyset$, by the earlier remark. Also, $Y \hookrightarrow Y_0$, so $Y \Vdash_{(\rho \circ i)[s/z]} \phi$, by (i), hence $Z \Vdash_{(\rho \circ t)[s/z]} \phi$, by Lemma 4.1. Thus indeed, $Z \Vdash_{(\rho \circ t)[s/z]} z = \emptyset \wedge \phi$. $\square$

For a sentence ϕ, we write $(\mathcal{E}, \mathcal{I}) \models \phi$ to mean that, for all worlds X, it holds that $X \Vdash \phi$ (by Lemma 4.1, it is enough 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 next theorem, is our main result about the forcing semantics.

Theorem 4.6 (Soundness and completeness for forcing semantics) *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 all toposes* $\mathcal{E}$ *and dssi* $\mathcal{I}$ *satisfying* $(\mathcal{E}, \mathcal{I}) \models \mathcal{T}$.

In this section, we give the proof of the soundness direction, (1) $\implies$ (2), of Theorem 4.6, and explore some of its consequences. The proof of completeness, which makes essential use of the technology of categories with class structure introduced in Part III of the paper, will eventually be given in Section 10.

PROOF of Theorem 4.6 (Soundness). The proof is in two parts. The first part is to verify that the forcing semantics soundly models the intuitionistic entailment relation. This part is completely routine, and we omit it entirely. The second part is to verify that the forcing interpretation validates the axioms of BIST^{-}+ Coll. The verification of these axioms makes extensive use of Lemma 4.5. Indeed, much of the hard work has already been done in the proof of this lemma. Here, we just give a detailed verification of the Collection axiom, which is arguably the most interesting case. The other cases are omitted.

To verify Coll, suppose we have X and ρ such that $X \Vdash_\rho \mathsf{S}(x)$ and $X \Vdash_\rho \forall y \in x.\,\exists z.\,\phi$. We must show that $X \Vdash_\rho \exists w.\,(\mathsf{S}(w) \wedge (\forall y \in x. \exists z \in w.\,\phi) \wedge (\forall z \in w. \exists y \in x.\,\phi)\,)$.

Because $X \Vdash_\rho \mathsf{S}(x)$, we have B such that $I_x \hookrightarrow \mathcal{P}B$. Define

$$Y = \{(b,a)\colon B \times X \mid b \in e_x(a)\}.$$

Let $s\colon Y \longrightarrow B$ and $t\colon Y \longrightarrow X$ be the projections. By Lemma 4.5, $Y \Vdash_{(\rho\circ t)[s/y]} \exists z.\, \phi$. So there exist $r\colon Z \twoheadrightarrow Y$ and $u\colon Z \longrightarrow A_z$ such that

$$Z \Vdash_{(\rho\circ t\circ r)[s\circ r/y,\, u/z]} \phi\,. \tag{6}$$

Define $A_w = \mathcal{P}A_z$ and $\rho_w\colon X \longrightarrow A_w$ by

$$\rho_w(a) = \{u(c) \mid c\colon Z \text{ and } t(r(c)) = a\}.$$

Henceforth, we work relative to the environment $\rho[\rho_w/w]$, for which we continue to write ρ. Using Lemma 4.5, we verify

$$X \Vdash_\rho (\forall y \in x. \exists z \in w.\, \phi) \wedge (\forall z \in w. \exists y \in x.\, \phi).$$

For the left-hand conjunct, we must show that $Y \Vdash_{(\rho\circ t)[s/y]} \exists z \in w.\, \phi$. Note that $I_w \hookrightarrow A_w = \mathcal{P}A_z$. Also, by (6), we have $r\colon Z \twoheadrightarrow Y$ and $u\colon Z \longrightarrow A_z$ such that $Z \Vdash_{(\rho\circ t\circ r)[s\circ r/y,\, u/z]} \phi$. We must show that $(\rho_w \circ t \circ r, u)$ factors through $\ni_{A_z}$. But this is immediate from the definition of ρ_w. For the right-hand conjunct, consider

$$Z' = \{(d,a)\colon A_z \times X \mid d \in \rho_w(a)\},$$

together with its projections $s'\colon Z' \longrightarrow A_z$ and $t'\colon Z' \longrightarrow X$. We must show that $Z' \Vdash_{(\rho\circ t')[s'/z]} \exists y \in x.\, \phi$. Define:

$$Y' = \{(s(r(c)),\, u(c),\, t(r(c)))\colon B \times A_z \times X \mid c\colon Z\}.$$

By the definition of ρ_w, if $(b,d,a)\colon Y'$ then $d \in \rho_w(a)$. Accordingly, there are projections $u'\colon Y' \longrightarrow B$ and $r'\colon Y' \longrightarrow Z'$. Reasoning internally in $\mathcal{E}$, we show that r' is epi. Suppose that $(d,a)\colon Z'$, i.e. $d \in \rho_w(a)$. Then $d = u(c)$ for some $c\colon Z$ such that $t(r(c)) = a$. So $(s(r(c)),\, d,\, a)\colon Y'$ is such that $r'(s(r(c)),\, d,\, a) = (d,a)$. Hence r' is indeed epi. By the definition of Y, if $(b,d,a)\colon Y'$ then $b \in e_x(a)$. Thus $(e_x \circ t' \circ r', u')$ factors through $\ni_{\mathcal{P}B}$. It remains to show that $Y' \Vdash_{(\rho\circ t'\circ r')[s'\circ r'/z,\, u'/y]} \phi$. For this, consider the morphism $\tau\colon Z \longrightarrow Y'$ defined by $\tau(c) = (s(r(c)), u(c), t(r(c)))$. Then $t'\circ r'\circ\tau = t\circ r$ and $s'\circ r'\circ\tau = u$ and $u'\circ\tau = s\circ r$. So, by (6), it holds that $Z \Vdash_{((\rho\circ t'\circ r')[s'\circ r'/z,\, u'/y])\circ\tau} \phi$. It is immediate from the definition of Y' that τ is epi. Hence, by Lemma 4.2, we have that $Y' \Vdash_{(\rho\circ t'\circ r')[s'\circ r'/z,\, u'/y]} \phi$ as required. This completes the verification of Collection. $\square$

The single case presented in the proof above should be sufficient to convey a flavour of the direct proof of soundness to the reader. The main reason for not giving a more comprehensive proof is that we shall anyway obtain a second proof of the soundness direction of Theorem 4.6 in Part III, which, although very indirect, is in many ways less brutal than a direct proof. (See Section 9 for the culmination of this proof.)

The next two propositions can be used in combination with Theorem 4.6 to obtain sound and complete classes of models for extensions of BIST^{-}+ Coll with Inf and/or REM.

Proposition 4.7 $(\mathcal{E},\mathcal{I}) \models \mathrm{Inf}$ *if and only if* $\mathcal{E}$ *has a natural numbers object.*

PROOF. We outline the proof of the more interesting (left-to-right) direction. Suppose that $(\mathcal{E},\mathcal{I}) \models \mathrm{Inf}$, i.e.

$$\begin{aligned}(\mathcal{E},\mathcal{I}) \models \exists I.\, \exists 0 \in I.\, \exists s \in I^I.\, &(\forall x \in I.\, s(x) \neq 0) \,\wedge \\ &(\forall x, y \in I.\, s(x) = s(y) \;\rightarrow\; x = y)\,.\end{aligned}$$

By stripping off all three existential quantifiers together, there exists an epi $X \twoheadrightarrow \mathbf{1}$, with maps $\rho_I \colon X \longrightarrow \mathcal{P}B$, $\rho_0 \colon X \longrightarrow B$ and $\rho_s \colon X \longrightarrow \mathcal{P}(B{\times}B)$, satisfying, internally in $\mathcal{E}$, for all $a\colon X$,

$$\begin{aligned}&\rho_0(a) \in_B \rho_I(a)\\ &\forall x\colon B.\, x \in_B \rho_I(a).\, \exists! y\colon B.\, (x,y) \in \rho_s(a)\\ &\forall x, y\colon B.\, x \in_B \rho_I(a) \;\wedge\; (x,y) \in \rho_s(a) \;\rightarrow\; y \in \rho_s(a),\end{aligned}$$

and such that:

$$X \Vdash_\rho (\forall x \in I.\, s(x) \neq 0) \;\wedge\; (\forall x, y \in I.\, s(x) = s(y) \;\rightarrow\; x = y) \tag{7}$$

Define $I_X \hookrightarrow X \times B$ as the relation represented by ρ_I. The above data determines morphisms $0_X \colon X \longrightarrow I_X$ and $s_X \colon I_X \longrightarrow I_X$. Moreover, by unwinding the meaning of (7), it holds that s_X is mono and has disjoint image from 0_X, i.e. that $[0_X, s_X] \colon X + I_X \longrightarrow I_X$ is mono. Now define I to be the exponential ${I_X}^X$ in $\mathcal{E}$. We have a point $0 \colon \mathbf{1} \longrightarrow I$ and morphism $s \colon I \longrightarrow I$ defined by

$$\begin{aligned}0 \;&=\; (a \mapsto 0_X(a))\\ s(f) \;&=\; (a \mapsto s_X(f(a))).\end{aligned}$$

Trivially, s is mono. Also, 0 and s have disjoint image, because 0_X and s_X do and the map $X \longrightarrow \mathbf{1}$ is an epi. We thus have a mono $[0, s] \colon \mathbf{1}{+}I \longrightarrow I$ in $\mathcal{E}$. It is a standard result that a natural numbers object in $\mathcal{E}$ can be constructed from such a mono. [[REFERENCE???]] $\square$

Proposition 4.8 $(\mathcal{E},\mathcal{I}) \models \mathrm{REM}$ *if and only if* $\mathcal{E}$ *is a boolean topos.*

PROOF. Suppose $\mathcal{E}$ is a boolean topos. Take any ϕ, X and ρ such that $X \Vdash_\rho !\phi$. We must show that $X \Vdash_\rho \phi \vee \neg\phi$. Let $i\colon Y \hookrightarrow X$ be the greatest subobject included in X such that $Y \Vdash_{\rho\circ i} \phi$, which exists by Lemma 4.5. Let $j\colon Z \hookrightarrow X$ be the complement of Y, which exists because $\mathcal{E}$ is boolean. As i, j are jointly epic and $Y \Vdash_{\rho\circ i} \phi$, it suffices to show that $Z \Vdash_{\rho\circ j} \neg\phi$. Accordingly, suppose $t\colon W \longrightarrow Z$ is such that $W \Vdash_{\rho\circ j\circ t} \phi$. Factoring $t\colon W \longrightarrow Z$ as $W \twoheadrightarrow Z' \xhookrightarrow{j'} Z$, we have, by Lemma 4.2, that $Z' \Vdash_{\rho\circ j\circ j'} \phi$. But then $Z' \hookrightarrow Y$ by the characterizing property of Y. Since also $Z' \hookrightarrow Z$ and Z is the complement of Y, we have that Z' is an initial object. Thus $Z' \Vdash_{\rho\circ j\circ j'} \perp$, and so $W \Vdash_{\rho\circ j\circ t} \perp$, by Lemma 4.1. This shows that indeed $Z \Vdash_{\rho\circ j} \neg\phi$.

Conversely, suppose that $(\mathcal{E}, \mathcal{I}) \models \mathrm{REM}$. Then

$$(\mathcal{E}, \mathcal{I}) \models \ \forall p.\ p \subseteq \{\emptyset\} \ \rightarrow\ (p = \emptyset \ \vee\ p = \{\emptyset\})\ ,$$

since this is a straightforward consequence of REM in BIST$^-$. The forcing semantics of the above sentence unwinds straightforwardly to obtain the following consequence in the Kripke-Joyal semantics of the internal logic of $\mathcal{E}$,

$$\mathcal{E} \models\ \forall p\colon \mathcal{P}\{\emptyset\}.\ p = \emptyset \ \vee\ p = \{\emptyset\}\ .$$

It is routine (and standard) that the property above is valid in the internal logic of $\mathcal{E}$ if and only if $\mathcal{E}$ is boolean. $\square$

We remark that the proposition above has the following perhaps surprising consequence. The underlying logic of the first-order set theories that we associate with boolean toposes is *not* 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. [[REFS]]. Here we find them arising naturally as a consequence of our forcing semantics.

At this point, we pause to discuss the meta-theory for the above results. We mention first that our proof of the completeness direction of Theorem 4.6, which appears in Part III of the paper, will use ZFC as its meta-theory. However, none of the proofs we have thus far given in the present section requires such a strong meta-theory. In fact all are formalizable in BIST itself in the following sense. If $\mathcal{E}$ is a small topos then all proofs are directly formalizable in BIST. However, when $\mathcal{E}$ is only a locally small topos, a difficulty arises. In such a case, although the forcing semantics can be formalized for any *fixed* formula ϕ, the inductive definition of the forcing semantics for *all* formulas ϕ cannot be internalized in BIST. Hence, for a locally small topos $\mathcal{E}$, the various soundess results are only formalizable in the following *schematic* sense: for any formula ϕ with a real-world proof in the relevant theory, the set-theoretic formula formalizing the statement $(\mathcal{E},\mathcal{I}) \models \phi$ is provable in BIST.[16] This situation cannot be improved upon, because, in BIST, the category **Set** is a non-trivial locally small topos with natural numbers object and dssi, and so, if the full soundness result were directly formalizable, then BIST would be able to prove its own consistency, contradicting Gödel's second incompleteness theorem.

One consequence of the above discussion is worth mentioning. Because, from the viewpoint of BIST, the category **Set** is a non-trivial locally small topos with natural numbers object and dssi, the schematic soundness result above unwinds to yield a translation of BIST+ Coll into BIST.[17] Thus the theory BIST exhibits the interesting property of being able to interpret Collection using Replacement.

Our next goal is to establish that, in the presence of a superdirected system of inclusions, the full Separation schema is validated by the forcing semantics. Thus, by Theorems 3.18 and 3.19, there is a useful collection of toposes modelling the full Separation schema. However, this result is only available if we strengthen our meta-theory by adding both Separation and Collection to BIST.

Proposition 4.9 (IST+ Coll) *If $\mathcal{I}$ is an sdssi on $\mathcal{E}$ then* $(\mathcal{E},\mathcal{I}) \models \mathrm{Sep}$.

(Because toposes with sdssi's are not small, the above result holds in the meta-theory IST+ Coll only in the schematic sense discussed above.)

PROOF. By Corollary 2.8, it suffices to verify R$\mathcal{S}$ and R$\exists$. We use the characterization of the forcing conditions for restrictedness established by Lemma 4.5.

First, we show that $(\mathcal{E},\mathcal{I}) \models \mathrm{R}\mathcal{S}$, i.e. for all X, ρ, it holds that $X \Vdash_\rho !\phi$. We must show that the family

$$\mathcal{Y} \;=\; \{Y_i \mid Y_i \xhookrightarrow{\;i\;} X \text{ and } Y \Vdash_{\rho\circ i} \phi\}$$

has a greatest element. Because $\mathcal{E}$ is locally small, the hom-set $\mathcal{E}(\mathbf{1}, \mathcal{P}X)$ determines a canonical set of inclusions into X in which each subobject of X is represented exactly once. Henceforth, we understand the inclusions in the definition of $\mathcal{Y}$ above as being restricted to this canonical set. Because the meta-theory has full Separation, the family $\mathcal{Y}$ is itself a set. For every

[16] Moreover, this schematic soundness result *should* itself be provable in a weak arithmetic such as PRA.

[17] It might be interesting to describe this translation explicitly. However, this lies outside the scope of the present paper.

$Y_i \in \mathcal{Y}$, there exists B such that $\mathrm{Im}(\rho \circ i) \hookrightarrow \mathcal{P}B$. By Collection in the meta-theory, there exists a set $\mathcal{B}$ such that, for every Y_i there exists $B \in \mathcal{B}$ with $\mathrm{Im}(\rho \circ i) \hookrightarrow \mathcal{P}B$. Using superdirectedness, there exists an upper bound C for $\mathcal{B}$ in $\mathcal{I}$. Thus, for all Y_i, we have $Y_i \hookrightarrow \mathcal{P}C$. Define $Y = X \cap \mathcal{P}C$. This is clearly the required greatest element of $\mathcal{Y}$.

To show that $(\mathcal{E}, \mathcal{I}) \models \mathrm{R}\exists$, suppose that $X \Vdash_\rho \forall x. !\phi$. We must show that $X \Vdash_\rho !(\exists x. \phi)$, i.e. that the family

$$\mathcal{Y} \;=\; \{Y_i \mid Y_i \xhookrightarrow{i} X \text{ and } Y \Vdash_{\rho \circ i} \exists x. \phi\}$$

has a greatest element. As above, we restrict the inclusions i in the definition of $\mathcal{Y}$ to the canonical ones, and, by full Separation in the meta-theory, $\mathcal{Y}$ is a set. For each $Y_i \in \mathcal{Y}$, there exist $Y_i \overset{r}{\twoheadleftarrow} Z \xrightarrow{a} A$ such that $Z \Vdash_{(\rho \circ i \circ r)[a/x]} \phi$. By Collection in the meta-theory, there is a set $\mathcal{Z}$ whose elements represent data of the form $X \xhookleftarrow{j} Y_j \overset{r}{\twoheadleftarrow} Z \xrightarrow{a} A$ for which $Z \Vdash_{(\rho \circ j \circ r)[a/x]} \phi$, such that, for every $Y_i \in \mathcal{Y}$, there exists data as above in $\mathcal{Z}$ with $Y_i = Y_j$. By superdirectedness, the set $\{A \mid (X \xhookleftarrow{j} Y_j \overset{r}{\twoheadleftarrow} Z \xrightarrow{a} A) \in \mathcal{Z}\}$ has an upper bound B in $\mathcal{I}$. Consider the projections $X \xleftarrow{t} X \times B \xrightarrow{s} B$. Because $X \Vdash_\rho \forall x. !\phi$, we have $X \times B \Vdash_{(\rho \circ t)[s/x]} !\phi$. Therefore, the family

$$\{R \mid R \xhookrightarrow{h} X \times B \text{ and } R \Vdash_{(\rho \circ t \circ h)[s \circ h/x]} \phi\}$$

has a greatest element, $S \xhookrightarrow{k} X \times B$. Let $S \overset{e}{\twoheadrightarrow} Y \xhookrightarrow{j} X$ be the image factorization of $t \circ k$. We show that Y is the required greatest element of $\mathcal{Y}$. Because $S \overset{e}{\twoheadrightarrow} Y$ and $S \Vdash_{(\rho \circ t \circ k)[s \circ k/x]} \phi$, i.e. $S \Vdash_{(\rho \circ j \circ e)[s \circ k/x]} \phi$, we indeed have that $Y \Vdash_{\rho \circ j} \exists x. \phi$. Now consider and $Y_i \in \mathcal{Y}$. We must show that $Y_i \hookrightarrow Y$. By the definitions of $\mathcal{Z}$ and B, there exists $Y_i \overset{r}{\twoheadleftarrow} Z \xrightarrow{a} A$ with $A \hookrightarrow B$ such that $Z \Vdash_{(\rho \circ i \circ r)[a/x]} \phi$. Defining b to be the composite $Z \xrightarrow{a} A \hookrightarrow B$, we have, by Lemma 4.3, that $Z \Vdash_{(\rho \circ i \circ r)[b/x]} \phi$. Let $X \overset{q}{\twoheadrightarrow} R \xhookrightarrow{h} X \times B$ be the image factorization of $\langle i \circ r, b \rangle$. Then $Z \Vdash_{(\rho \circ t \circ h \circ q)[s \circ h \circ q/x]} \phi$. So, by Lemma 4.2, $R \Vdash_{(\rho \circ t \circ h)[s \circ h/x]} \phi$. Thus $R \xhookrightarrow{l} S$, by the definition of S. Then:

$$i \circ r \;=\; t \circ h \circ q \;=\; t \circ k \circ l \circ q \;=\; j \circ e \circ l \circ q.$$

But i, j are inclusions and r an epi, so

$$Y_i \equiv \mathrm{Im}(i \circ r) \equiv \mathrm{Im}(j \circ e \circ l \circ q) \equiv \mathrm{Im}(e \circ l \circ q) \hookrightarrow Y.$$

Thus indeed $Y_i \hookrightarrow Y$. $\square$

We make one comment on the above proof. Curiously, it is not at all straightforward to directly verify the validity of the schema R$\forall$ from Figure 3 using an intuitionistic metatheory such as IST+ Coll. (The verification is easy in a classical metatheory.) Thus Lemma 2.7, on which Corollary 2.8 depends, is extremely helpful in permitting the simple proof above.

In contrast to the characterizations of Inf and REM, Proposition 4.9 only establishes a sufficient condition for the validity of full Separation. Indeed, there appears to be no reason for superdirectedness to be a necessary condition for Sep to hold. Similarly, there is no reason for BIST^{-}+ Coll+ Sep to be complete axiomatization of the valid sentences with respect to toposes with sdssi's. It would be interesting to have mathematical confirmation of these expectations.

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 is tedious. Rather than pursuing this line any further, we instead refer the reader to Section 9 in Part III, 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 4.10 *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).*

PROOF (outline). This essentially follows from the forcing semantics of the formula $Nat(N, 0, s)$ from Section 2, which characterizes $N \colon \mathbf{1} \longrightarrow \mathcal{P}A$, for some A, as classifying the natural numbers object of $\mathcal{E}$. Given this, the forcing interpretation of bounded quantifiers in Lemma 4.5 means that they are interpreted identically to quantifiers in the internal logic of $\mathcal{E}$. $\Box$

Again, there is a more natural formulation of the above result using the tools of categorical logic in Section 9. As a consequence, we obtain the postponed proof of Proposition 2.17.

PROOF of Proposition 2.17. Let $\mathcal{E}$ be the free topos with natural numbers object. By Theorem 3.10 there is an equivalent category $\mathcal{E}'$ carrying a dssi $\mathcal{I}'$. By Gödel's second incompleteness theorem, the Π^0_1 sentence Con(HAH) is not validated by the internal logic of $\mathcal{E}$. see e.g. [[REF]], and hence not by

$\mathcal{E}'$ either. Therefore, by Proposition 4.10, Con(HAH) is not validated by the forcing semantics in $\mathcal{E}'$. It now follows from the soundness results above that BIST+ Coll+ REM $\nvdash$ Con(HAH). $\square$

We end this section with further applications of the soundness theorem to obtain non-derivability results, for which we take ZFC as the meta-theory. 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*

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

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 4.11 $\mathbf{V}_\omega(\mathbb{N}) \models$ Inf*, but* $\mathbf{V}_\omega(\mathbb{N}) \not\models$ vN-Inf*.*

PROOF (outline). One can straightforwardly check the following general equivalences. The category $\mathbf{V}_\lambda(A)$ has a natural numbers object if and only if $\lambda > \omega$ or $|A| \geq \aleph_0$. Hence, by Proposition 4.7, $\mathbf{V}_\lambda(A) \models$ Inf if and only if $\lambda > \omega$ or $|A| \geq \aleph_0$. Also $\mathbf{V}_\lambda(A) \models$ vN-Inf if and only if $\lambda > \omega$ (because for $\lambda = \omega$, all sets in $V_\alpha(A)$ have finite rank and so cannot model vN-Inf). In particular, $\mathbf{V}_\omega(\mathbb{N}) \models$ Inf but $\mathbf{V}_\omega(\mathbb{N}) \not\models$ vN-Inf. $\square$

Proposition 2.9 follows as an immediate consequence. In fact, more generally:

Corollary 4.12 BIST+ Coll+ REM $\nvdash$ vN-Inf .

By the proof of Proposition 4.11, 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 the standard model theory of 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, the argument in the proof of Proposition 2.18, which attempts to construct the union of the chain $N, \mathcal{P}(N), \mathcal{P}^2(N), \ldots$, is not validated by the forcing semantics of $\mathbf{V}_{\omega+\omega}(\emptyset)$. 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). More specifically, consideration of this model shows that it is impossible to define the sequence $N, \mathcal{P}(N), \mathcal{P}^2(N), \ldots$ inside the theory BIST+ Coll+ REM+ vN-Inf. Given that the existence of such a sequence is the quintessential example of an application of Replacement in ZF set theory, some readers may wonder whether Collection and Replacement are of any practical use in BIST if they cannot be applied to obtain such standard consequences. In fact, these principles *are* highly useful in BIST for performing any form of reasoning relating small and large structures, for example the development of the theory of locally small categories. Since one of our main motivations for the present work is the development of a language for reasoning about large structures relative to any elementary topos (see Section 1 for further discussion), it is a major advantage of our approach that Replacement and Collection are validated.

We end the section with the remark that the full hierarchy $\mathbf{V}(\emptyset)$ models full Separation, by Proposition 4.9. Hence, by Corollary 2.16, the category $\mathbf{V}(\emptyset)$ is a model of the theory BIST+ Coll+ LEM. In fact, making use of Collection in ZFC to unwind the forcing semantics, it is straightforward to show that the forcing semantics in $\mathbf{V}(\emptyset)$ simply expresses meta-theoretic truth in ZFC.

PART III — CATEGORIES OF CLASSES

5 Basic class structure

In the previous sections we have shown how to interpret the language of first-order set theory in any elementary topos endowed with a system of inclusions, where the system of inclusions is used to interpret the unbounded quantifiers. There is an alternative more algebraic approach to modelling quantification over classes, namely to consider categories in which the objects themselves represent classes rather than sets. Within such categories, the "unbounded" quantifiers become *de facto* bounded, and can thus be handled using the standard machinery of categorical logic. The axiomatic basis for such an approach was developed by Joyal and Moerdijk in their book on *algebraic set theory* [7], and further refined in [11,**?**]. In this part of the paper, we adapt this approach to obtain categories of classes appropriate for modelling the set theory BIST and its variants.

Following the approach of algebraic set theory, we axiomatize properties of a collection of "small maps" within an ambient category of classes. The idea is that arbitrary maps represent functions (i.e. functional relations) between classes, and small maps are the functions with "small" fibres. The basic notion of small map determines natural notions of smallness for other concepts.

Definition 5.1 (Small object) An object A is called *small* if the unique map $A \longrightarrow 1$ is small.

Definition 5.2 (Small subobject) A subobject $A \rightarrowtail C$ is called *small* if A is a small object.

Definition 5.3 (Small relation) A relation $R \rightarrowtail C \times D$ is called *small* if its second projection $R \rightarrowtail C \times D \longrightarrow D$ is a small map.

Note that the definition of small relation is orientation-dependent. The orientation is chosen so that a morphism $f\colon A \longrightarrow B$ is small if and only if its graph $\langle 1, f\rangle\colon A \rightarrowtail A \times B$ is a small relation. (This is opposite to the orientation of "small relations" in [11].)

In this section, we define a notion of *basic class structure* on a category, adequate to ensure that the category behaves like the category of classes for a very weak first-order set theory (cf. [[REFS]]). This notion provides the basis for considering strengthened notions of class structure in subsequent sections.

Before axiomatizing the required properties of small maps, we need to place basic requirements on the ambient category. A *positive Heyting category* is a category $\mathcal{C}$ satisfying the following conditions:

- **(C1)** $\mathcal{C}$ is *regular*: i.e. it has finite limits; the kernel pair [18] k_1, k_2 of every arrow $f : A \longrightarrow B$ has a coequalizer $q : B \longrightarrow C$; and regular epimorphisms are stable under pullback.
- **(C2)** $\mathcal{C}$ has *finite coproducts*, and these are disjoint and stable under pullbacks.
- **(C3)** $\mathcal{C}$ has *dual images*, i.e. for every arrow $f : C \longrightarrow D$, the inverse image map $f^{-1} : \mathrm{Sub}(D) \to \mathrm{Sub}(C)$ has a right adjoint $\forall_f : \mathrm{Sub}(C) \to \mathrm{Sub}(D)$ (considering f^{-1} as a functor between posets).

Condition (C1) implies that every morphism $f\colon A \longrightarrow B$ in $\mathcal{C}$ factors (uniquely up to isomorphism) as a regular epi followed by a mono

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

[18] The *kernel pair* of $f\colon A \longrightarrow B$ is the span $k_1, k_2\colon P \longrightarrow A$ that forms the pullback of f along itself.

(N.B., it is not necessarily the case that every epi is regular in $\mathcal{C}$.) Moreover, such image factorizations are stable under pullback. Further, for every arrow $f : C \longrightarrow D$, the inverse image map, $f^{-1} : \mathrm{Sub}(D) \to \mathrm{Sub}(C)$ has a left adjoint, $\exists_f : \mathrm{Sub}(C) \to \mathrm{Sub}(D)$.

One reason for focusing on positive Heyting categories is the following standard proposition, cf. [[ELEPHANT???]].

Proposition 5.4 *In every positive Heyting category, each partial order* $\mathrm{Sub}(C)$ *of subobjects of* C *is a Heyting algebra. For every arrow* $f\colon C \longrightarrow D$*, the inverse image functor* $f^{-1} : \mathrm{Sub}(D) \to \mathrm{Sub}(C)$ *has both right and left adjoints* $\forall_f$ *and* $\exists_f$ *satisfying the "Beck-Chevally condition" of stability under pullbacks. In particular,* $\mathcal{C}$ *models intuitionistic, first-order logic with equality.*

By a *system of small maps* on a positive Heyting category $\mathcal{C}$ we mean a collection of arrows $\mathcal{S}$ of $\mathcal{C}$ satisfying the following conditions:

(S1) $\mathcal{S} \hookrightarrow \mathcal{C}$ is a subcategory with the same objects as $\mathcal{C}$. Thus every identity map $1_C : C \longrightarrow C$ is small, and the composite $g \circ f : A \longrightarrow C$ of any two small maps $f : A \longrightarrow B$ and $g : B \longrightarrow C$ is again small.
(S2) The pullback of a small map along any map is small. Thus in an arbitrary pullback diagram,

$$\begin{tikzcd}
C' \arrow[r] \arrow[d, "f'"'] & C \arrow[d, "f"] \\
D' \arrow[r] & D
\end{tikzcd}$$

if f is small then so is f'.

Proposition 5.5 *Given (S1) and (S2), the following are equivalent.*

(1) Every diagonal $\Delta : C \longrightarrow C \times C$ *is small.*
(2) Every regular monomorphism is small.
(3) if $g \circ f$ *is small, then so is* f*.*

PROOF. That 1 implies 2 follows from (S2), because every regular mono is a pullback of a diagonal.

To show 2 $\implies$ 3, suppose regular monos are small. Consider the following pullback diagram, with $g \circ f$ small:

$$\begin{tikzcd}
P \arrow[rr, "p_2"] \arrow[d, "p_1"'] & & B \arrow[d, "g"] \\
A \arrow[r, "f"'] & B \arrow[r, "g"'] & C
\end{tikzcd}$$

The arrow p_1 is a split epi, as can be seen by considering the pair $1 : A \longrightarrow A$ and $f : A \longrightarrow B$. Call the section $s : A \longrightarrow P$. This is a split mono, hence regular mono hence small. But p_2 is small by (S2). So $f = p_2 \circ s$ is small.

Finally, because identities are small, 3 implies that split monos are small. In particular, diagonals are small. $\square$

(S3) The equivalent conditions of Proposition 5.5 hold.

Note that a consequence of Proposition 5.5(3) is that if an object A is small then every morphism $f\colon A \longrightarrow B$ is a small map.

(S4) If $f \circ e$ is small and e is a regular epi, then f is small, as indicated in the diagram:

$$\begin{tikzcd}
A \arrow[r, two heads, "e"] \arrow[dr, "f \circ e"'] & B \arrow[d, "f"] \\
& C
\end{tikzcd}$$

(S5) Copairs of small maps are small. Thus if $f : A \longrightarrow C$ and $g : B \longrightarrow C$ are small, then so is $[f, g] : A + B \to C$.

Proposition 5.6 *Given (S1)–(S5), the following also hold:*

(1) The objects 0 *and* $1 + 1$ *are small.*
(2) If the maps $f : C \longrightarrow D$ *and* $f' : C' \longrightarrow D'$ *are small, then so is* $f + f' : C + C' \longrightarrow D + D'$.

PROOF. This follows easily from disjointness and stability of coproducts. $\square$

The final axiom of basic class structure requires every class to have a "powerclass" of all small subobjects (i.e., a *class* of all sub*sets*). Its formulation is similar to the defining property of powerobjects in toposes (Definition 3.1), only adjusted for small relations.

(P) Every object C has a *small powerobject*: an object $\mathcal{P}C$ with a small relation $\in_C \rightarrowtail C \times \mathcal{P}C$ (the *membership relation*) such that, for any object X and any small relation $R \rightarrowtail C \times X$, there is a unique arrow $\chi_R : X \longrightarrow \mathcal{P}C$ fitting into a pullback diagram of the form below.

$$\begin{tikzcd} R \arrow[r] \arrow[d, tail] & \in_C \arrow[d, tail] \\ C \times X \arrow[r, "1_C \times \chi_R"'] & C \times \mathcal{P}C \end{tikzcd}$$

Definition 5.7 (Basic class structure) A category with *basic class structure* is given by a positive Heyting category $\mathcal{C}$ together with a collection of small maps $\mathcal{S}$ satisfying axioms (S1)–(S5) and (P) above.

Definition 5.8 (Logical functor) A functor between categories with basic class structure is said to be *logical* if it preserves: positive Heyting category structure, small maps and membership relations.

As is standard, in this definition, we do not require the positive Heyting category structure and membership relations to be preserved "on the nose", but up to coherent isomorphism.

Later on we shall also need the natural category of sets associated with a category with basic class structure. We define this now.

Definition 5.9 (Category of sets) Given a category $\mathcal{C}$ with basic class structure $\mathcal{S}$, the associated *category of sets* $\mathcal{E}_{\mathcal{S}}(\mathcal{C})$ is the full subcategory of $\mathcal{C}$ on the small objects. Note that $\mathcal{E}_{\mathcal{S}}(\mathcal{C})$ is also a full subcategory of $\mathcal{S}$.

In the remainder of the section, we establish properties of categories with basic class structure. Assume that $\mathcal{C}$ is such a category with small maps $\mathcal{S}$.

For any small $A \xrightarrow{f} B$, the relation $\langle 1_A, f\rangle : A \longrightarrow A \times B$ is small, and we write $B \xrightarrow{f^{-1}} \mathcal{P}A$ for the unique morphism fitting into a pullback diagram

$$\begin{tikzcd} A \arrow[r] \arrow[d, "\langle 1_A, f\rangle"'] \arrow[dr, phantom, "\lrcorner", very near start] & \in_A \arrow[d, tail] \\ A \times B \arrow[r, "1 \times f^{-1}"] & A \times \mathcal{P}A \end{tikzcd} \tag{8}$$

Equivalently, f^{-1} is the unique morphism fitting into a pullback diagram

$$\begin{tikzcd} A \arrow[r] \arrow[d, "f"'] \arrow[dr, phantom, "\lrcorner", very near start] & \in_A \arrow[d, "\pi_A"] \\ B \arrow[r, "f^{-1}"] & \mathcal{P}A \end{tikzcd} \tag{9}$$

where π_A is the composite $\in_A \rightarrowtail A \times \mathcal{P}A \xrightarrow{\pi_2} \mathcal{P}A$, which is a small map. The lemma below will prove useful later.

Lemma 5.10 *If f is a small regular epi then f^{-1} is a small mono.*

PROOF. Suppose $A \xrightarrow{f} B$ is a small regular epi. Let $A \longrightarrow \in_A$ be as in diagram (8). By that diagram, the composite $A \longrightarrow \in_A \rightarrowtail A \times \mathcal{P}A \xrightarrow{\pi_1} A$ is the identity. Therefore $A \longrightarrow \in_A$ is a split mono, hence small by (S3). In the pullback diagram (9), the left edge is a regular epi and the top edge a mono. It is a property of regular categories that, in any such pullback square, the bottom edge is also a mono. Thus f^{-1} is a mono. It is small by (S4), because the top-right composite of (9) is small. □

By the existence of f^{-1} for small f, one sees that a a map $A \xrightarrow{f} B$ is small if and only if it can be obtained as a pullback of π_A. This fact allows the property of "smallness" to be expressed using the internal logic of $\mathcal{C}$ in the following sense, cf. [7, Proposition 1.6].

Proposition 5.11 *Every $f: A \longrightarrow B$ determines a subobject $m: B_f \rightarrowtail B$ satisfying: for any map $t: C \longrightarrow B$, it holds that the pullback $t^*(f)$ is small if and only if t factors through m.*

PROOF. Using the internal logic of $\mathcal{C}$, define B_f to be the subobject:

$$\{y: B \mid \exists Z: \mathcal{P}(A). \forall x: A. (x \in Z \leftrightarrow f(x) = y)\} .$$

That this has the right properties is easily verified using the Kripke-Joyal semantics of $\mathcal{C}$. □

Henceforth, we use the more suggestive $\{y: B \mid f^{-1}(y) \text{ is small}\}$ for the subobject B_f determined by the proposition.

A consequence of axioms (S1),(S2) and (S4) in combination is that small relations form a category under relational composition. Clearly the identity relation $\Delta : A \rightarrowtail A \times A$ is small. To see that relational composition preserves smallness, suppose $R \rightarrowtail A \times B$ and $R' \rightarrowtail B \times C$ are small relations. Recall that the composite relation $(R; R') \rightarrowtail A \times C$ is obtained by the factorization

$$R \times_B R' \twoheadrightarrow (R; R') \rightarrowtail A \times C$$

of the pair formed from the span below.

$$\begin{tikzcd}
R \times_B R' \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & R' \arrow[r] \arrow[d] & C \\
R \arrow[r] \arrow[d] & B & \\
A & &
\end{tikzcd} \tag{10}$$

We show that $R; R'$ is indeed a small relation. By assumption, we have that the morphisms $R \longrightarrow B$ and $R' \longrightarrow C$ in (10) are small. By (S2), the arrow $R \times_B R' \longrightarrow R'$ is also small. Thus, the composite

$$R \times_B R' \longrightarrow R' \longrightarrow C \quad = \quad R \times_B R' \longrightarrow A \times C \xrightarrow{\pi_2} C$$

is small. The required smallness of $(R' \circ R) \rightarrowtail A \times C \xrightarrow{\pi_2} C$ now follows from (S4).

We write $\mathcal{R}_{\mathcal{S}}(\mathcal{C})$ for the category with the same objects as $\mathcal{C}$ and with small relations $R \rightarrowtail A \times B$ as morphisms from A to B. There is an identity-on-objects functor $I\colon \mathcal{S} \to \mathcal{R}_{\mathcal{S}}(\mathcal{C})$ mapping any small $f\colon A \longrightarrow B$ to the small relation $\langle 1_A, f\rangle \rightarrowtail A \times B$. There is also an identity-on-objects functor $J\colon \mathcal{C}^{\mathrm{op}} \to \mathcal{R}_{\mathcal{S}}(\mathcal{C})$ mapping any $f\colon A \to B$ to $\langle f, 1_A\rangle \rightarrowtail B \times A$.

Axiom (P) is equivalent to asking for the functor

$$\mathcal{R}_{\mathcal{S}}(\mathcal{C})[A, J(-)]\colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$$

to be representable for every object A. That is, there is an isomorphism

$$\mathcal{R}_{\mathcal{S}}(\mathcal{C})[A, J(B)] \cong \mathcal{C}[B, \mathcal{P}A] \ ,$$

natural in B. Defining $\Omega = \mathcal{P}\mathbf{1}$, this specializes to

$$\mathcal{R}_{\mathcal{S}}(\mathcal{C})[\mathbf{1}, J(B)] \cong \mathcal{C}[B, \Omega] \ . \tag{11}$$

Easily, $\mathcal{R}_{\mathcal{S}}(\mathcal{C})[\mathbf{1}, J(B)]$ is isomorphic to the collection of small monomorphisms into B. Thus Ω classifies subobjects defined by small monomorphisms. By (S3), every regular mono is small. Conversely, every small mono $B' \rightarrowtail B$ is the equalizer of its classifier $B \longrightarrow \Omega$ with $\top \colon B \longrightarrow \Omega$, where $\top$ is classifies the identity 1_B. Thus a monomorphism is small if and only if it is regular. So Ω is a regular-subobject classifier. Note that every small subobject is represented by a small monomorphism, but small monomorphisms do not necessarily determine small subobjects. For example, 1_B is a small monomorphism for every B, but only determines a small subobject when B is small. However, in the case that B is a small object, a monomorphism $A \rightarrowtail B$ is small if and only if it presents A as a small subobject of B

Axiom (P) is also equivalent to asking for J to have a left adjoint. By composing the functor $J : \mathcal{C}^{\mathrm{op}} \to \mathcal{R}_{\mathcal{S}}(\mathcal{C})$ with its left adjoint, we obtain a comonad on $\mathcal{C}^{\mathrm{op}}$, hence monad on $\mathcal{C}$, whose underlying functor is a covariant small powerobject functor.

For future reference, we give explicit definitions of the covariant functor and the unit of the monad. The endofunctor maps an object A to $\mathcal{P}A$. Its action on morphisms maps $f : A \longrightarrow B$ to $f_! : \mathcal{P}A \longrightarrow \mathcal{P}B$ is defined as follows. Let $U \rightarrowtail B \times \mathcal{P}A$ be obtained as the mono part of the factorisation of:

$$\in_A \rightarrowtail A \times \mathcal{P}A \xrightarrow{f \times 1} B \times \mathcal{P}A$$

By (S4), $U \rightarrowtail B \times \mathcal{P}A$ represents a small relation. Accordingly, define $f_!$ to be the unique map fitting into the pullback below:

$$\begin{tikzcd} U \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \in_B \arrow[d, tail] \\ B \times \mathcal{P}A \arrow[r, "1 \times f_!"] & B \times \mathcal{P}B \end{tikzcd}$$

Proposition 5.12 *If $f \colon A \longrightarrow B$ is mono then so is $f_! \colon \mathcal{P}A \longrightarrow \mathcal{P}B$.*

PROOF. It is easily checked that $J \colon \mathcal{C}^{\mathrm{op}} \to \mathcal{R}_{\mathcal{S}}(\mathcal{C})$ preserves epis. Its left adjoint automatically preserves epis. Thus the composite endofunctor on $\mathcal{C}^{\mathrm{op}}$ preserves epis too, whence the corresponding endofunctor on $\mathcal{C}$ preserves monos. It is easily verified that $f \mapsto f_!$ is this endofunctor. $\square$

The unit of the monad is given by $\{\cdot\} : A \longrightarrow \mathcal{P}A$ defined by $\{\cdot\} = {1_A}^{-1}$. By Lemma 5.10, $\{\cdot\}$ is a small mono.

Next, following the argument outlined in [11], we establish a "descent" property, Proposition 5.14 below, which says that a map is small if it is small locally on a cover. This property was assumed as an axiom for small maps in [7]. First, we need a technical lemma, establishing an internal Beck-Chevalley property (cf. [9, p. 206]).

Lemma 5.13 *For any pullback diagram as on the left below with f small, the diagram on the right commutes.*

$$
\begin{tikzcd}
A' \arrow[r, "g'"] \arrow[d, "f'"'] \arrow[dr, phantom, very near start, "\lrcorner"] & A \arrow[d, "f"] \\
B' \arrow[r, "g"] & B
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
\mathcal{P}A' \arrow[r, "g'_!"] & \mathcal{P}A \\
B' \arrow[u, "f'^{-1}"] \arrow[r, "g"] & B \arrow[u, "f^{-1}"']
\end{tikzcd}
$$

PROOF. We prove that both sides of the right-hand square represent the small relation $\langle g', f'\rangle \colon A' \rightarrowtail A \times B'$. For $f^{-1} \circ g$ this is by a simple composition of pullbacks:

$$
\begin{tikzcd}
A' \arrow[r, "g'"] \arrow[d, tail, "{\langle g', f'\rangle}"'] \arrow[dr, phantom, very near start, "\lrcorner"] & A \arrow[r] \arrow[d, tail, "{\langle 1, f\rangle}"'] \arrow[dr, phantom, very near start, "\lrcorner"] & \in_A \arrow[d, tail] \\
A \times B' \arrow[r, "1 \times g"] & A \times B \arrow[r, "1 \times f^{-1}"] & B \times \mathcal{P}B
\end{tikzcd}
$$

For $g'_! \circ f'^{-1}$, we have the pullbacks below.

$$
\begin{tikzcd}
A' \arrow[r] \arrow[d, tail, "{\langle 1, f'\rangle}"'] \arrow[dr, phantom, very near start, "\lrcorner"] & \in_{A'} \arrow[d, tail] \\
A' \times B' \arrow[r, "1 \times f'^{-1}"] \arrow[d, "g' \times 1"'] \arrow[dr, phantom, very near start, "\lrcorner"] & A' \times \mathcal{P}A' \arrow[d, "g' \times 1"] \\
A \times B' \arrow[r, "1 \times f'^{-1}"] & A \times \mathcal{P}A'
\end{tikzcd}
$$

Let $U \rightarrowtail A \times \mathcal{P}A' \times A$ be the image of $\in_{A'} \rightarrowtail A' \times \mathcal{P}A' \xrightarrow{g' \times 1} A \times \mathcal{P}A'$. Then, by the stability of images and because $A' \xrightarrow{\langle g', f'\rangle} A \times B'$ is mono, the outer pullback square above implies that the left-hand square below is a pullback.

$$\begin{tikzcd}
A' \arrow[rr] \arrow[d, tail, "{\langle g', f' \rangle}"'] \arrow[dr, phantom, "\lrcorner", very near start] & & U \arrow[rr] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & & \in_A \arrow[d, tail] \\
A \times B' \arrow[rr, "1 \times f'^{-1}"] & & A \times \mathcal{P}A' \arrow[rr, "1 \times g'_!"] & & A \times \mathcal{P}A
\end{tikzcd}$$

The right-hand square is also a pullback, by the definition of $g'_!$. So $g'_! \circ f'^{-1}$ does indeed represent $\langle g', f' \rangle$. $\square$

Proposition 5.14 (Descent) *If g appears in a pullback diagram*

$$\begin{tikzcd}
A \arrow[r, two heads, "e'"] \arrow[d, "f"'] \arrow[dr, phantom, "\lrcorner", very near start] & C \arrow[d, "g"] \\
B \arrow[r, two heads, "e"] & D
\end{tikzcd}$$

where e is regular epi and f is small, then it follows that g is small.

PROOF. By defining $B' \overset{r_1}{\underset{r_2}{\rightrightarrows}} B$ to be the kernel pair of $B \overset{e}{\twoheadrightarrow} D$ and pulling back, we obtain:

$$\begin{tikzcd}
A' \arrow[r, shift left, "r'_1"] \arrow[r, shift right, "r'_2"'] \arrow[d, "f'"'] \arrow[dr, phantom, "\lrcorner", very near start] & A \arrow[r, two heads, "e'"] \arrow[d, "f"'] \arrow[dr, phantom, "\lrcorner", very near start] & C \arrow[d, "g"] \\
B' \arrow[r, shift left, "r_1"] \arrow[r, shift right, "r_2"'] & B \arrow[r, two heads, "e"] & D
\end{tikzcd}$$

where both rows are exact diagrams,[19] and each of the two left-hand squares is a pullback. By applying Lemma 5.13 to the two left-hand pullbacks we obtain: $e'_! \circ f^{-1} \circ r_1 = e'_! \circ (r'_1)_! \circ f'^{-1} = e'_! \circ (r'_2)_! \circ f'^{-1} = e'_! \circ f^{-1} \circ r_2$. So, by the coequaliser property of e, there exists $D \xrightarrow{h} \mathcal{P}C$ such that $h \circ e = e'_! \circ f^{-1}$.

[19] An *exact diagram* is a diagram $A \overset{r_1}{\underset{r_2}{\rightrightarrows}} B \overset{e}{\twoheadrightarrow} C$ where r_1, r_2 is the kernel pair of e and e is the coequalizer of r_1, r_2.

As in the proof of Lemma 5.13, we have pullbacks:

$$\begin{tikzcd}
A \arrow[r] \arrow[d, tail, "{\langle e', f \rangle}"'] \arrow[dr, phantom, "\lrcorner", very near start] & U \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \in_C \arrow[d, tail] \\
C \times B \arrow[r, "1 \times f^{-1}"] & C \times \mathcal{P}A \arrow[r, "1 \times e'_!"] & C \times \mathcal{P}C
\end{tikzcd}$$

showing that $e'_! \circ f^{-1}$ represents $\langle e', f \rangle \rightarrowtail C \times B$. Using the equality $h \circ e = e'_! \circ f^{-1}$, we reconstruct the outer rectangle above by pulling back in stages:

$$\begin{tikzcd}
A \arrow[r, two heads, "e''"] \arrow[d, tail, "{\langle e', f \rangle}"'] \arrow[dr, phantom, "\lrcorner", very near start] & X \arrow[r] \arrow[d, tail, "m"'] \arrow[dr, phantom, "\lrcorner", very near start] & \in_C \arrow[d, tail] \\
C \times B \arrow[r, two heads, "1 \times e"] & C \times D \arrow[r, "1 \times h"] & C \times \mathcal{P}C
\end{tikzcd} \tag{12}$$

But $C \times B' \overset{1 \times r_1}{\underset{1 \times r_2}{\rightrightarrows}} C \times B \xrightarrow{1 \times e} C \times D$ is an exact diagram. So, by pulling it back along m we obtain:

$$\begin{tikzcd}
A' \arrow[r, shift left, "r'_1"] \arrow[r, shift right, "r'_2"'] \arrow[d, "{\langle d, f' \rangle}"'] \arrow[dr, phantom, "\lrcorner", very near start] & A \arrow[r, two heads, "e''"] \arrow[d, tail, "{\langle e', f \rangle}"'] \arrow[dr, phantom, "\lrcorner", very near start] & X \arrow[d, "m"] \\
B' \times C \arrow[r, shift left, "1 \times r_1"] \arrow[r, shift right, "1 \times r_2"'] & C \times B \arrow[r, two heads, "1 \times e"] & C \times D
\end{tikzcd}$$

where $d = e' \circ r'_1 = e' \circ r'_2$. But then the top row is exact, hence e'' coequalizes r'_1, r'_2. Since e' also coequalizes r'_1, r'_2, the left-hand pullback square of (12) can be taken to be:

$$\begin{tikzcd}
A \arrow[r, two heads, "e'"] \arrow[d, "{\langle e', f \rangle}"'] \arrow[dr, phantom, "\lrcorner", very near start] & C \arrow[d, "{\langle 1, g \rangle}"] \\
C \times B \arrow[r, two heads, "1 \times e"] & C \times D
\end{tikzcd}$$

Then, by the right-hand pullback of (12), $m = \langle 1, g \rangle$ is a small relation. Hence g is indeed small (and also $h = g^{-1}$). $\square$

Finally in this section, we verify that basic class structure is preserved by taking slice categories. To establish this, we need an internal subset relation.

Definition 5.15 (Subset relation) For any object B, a *subset relation* is a relation $\subseteq_B : \mathcal{P}B \relbar\joinrel\mapstochar\joinrel\rightarrow \mathcal{P}B$ such that any morphism $\langle f, g\rangle : A \longrightarrow \mathcal{P}B \times \mathcal{P}B$ factors through $\subseteq_B \rightarrowtail \mathcal{P}B \times \mathcal{P}B$ if and only if, in the diagram below, $P \rightarrowtail B \times A$ factors through $Q \rightarrowtail B \times A$.

$$\begin{tikzcd}
P \arrow[r] \arrow[d, tail] \arrow[dr, phantom, very near start, "\lrcorner"] & \in_B \arrow[d, tail] & Q \arrow[l] \arrow[d, tail] \arrow[dl, phantom, very near start, "\llcorner"] \\
B \times A \arrow[r, "1 \times f"] & B \times \mathcal{P}B & B \times A \arrow[l, "1 \times g"']
\end{tikzcd}$$

The definition uniquely characterizes the subset relation, which indeed exists by the proposition below.

Proposition 5.16 *For every B, the subset relation $\subseteq_B : \mathcal{P}B \relbar\joinrel\mapstochar\joinrel\rightarrow \mathcal{P}B$ exists.*

PROOF. The relation $\subseteq_B \rightarrowtail \mathcal{P}B \times \mathcal{P}B$, can be constructed logically as:

$$\subseteq_B \;=\; [\![(y, z) : \mathcal{P}B \times \mathcal{P}B |\ \forall x : B.\ x \in y \Rightarrow x \in z]\!]$$

Here we use the canonical interpretation $[\![\ - \ ]\!]$ of first-order logic in the internal logic of $\mathcal{C}$, interpreting the atomic formula $x \in y$ using the membership relation $\in_B \rightarrowtail B \times \mathcal{P}B$. It is routine to verify that this has the required properties. $\Box$

For any slice category $\mathcal{C}/I$, define $\mathcal{S}_I$ to be the subcategory of those maps whose image under the forgetful to $\mathcal{C}$ is small. Our next result is the analogue for basic class structure of the "fundamental theorem" of topos theory, see e.g. [[REFS]].

Proposition 5.17

(1) $\mathcal{S}_I$ gives basic class structure on $\mathcal{C}/I$.
(2) For any $h : I \longrightarrow J$, the reindexing functor $h^* : \mathcal{C}/J \to \mathcal{C}/I$ is logical.
(3) For any small $h : I \longrightarrow J$, the reindexing functor $h^* : \mathcal{C}/J \to \mathcal{C}/I$ has a right adjoint $\Pi_h : \mathcal{C}/J \to \mathcal{C}/I$.

The proof is presented as a series of lemmas.

Lemma 5.18 *$\mathcal{S}_I$ gives basic class structure on $\mathcal{C}/I$.*

PROOF. It is standard that $\mathcal{C}/I$ satisfies (C1)–(C3), and easily checked that (S1)–(S5) hold for $\mathcal{S}_I$. We verify that (P) holds.

For any object $B \xrightarrow{g} I$ of $\mathcal{C}/I$, define $\mathcal{P}g$ to be the left edge of the diagram below

$$\begin{tikzcd}
V \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \subseteq_I \arrow[d, tail] \\
\mathcal{P}B \times I \arrow[r, "g_! \times \{\cdot\}"] \arrow[d, "\pi_2"'] & \mathcal{P}I \times \mathcal{P}I \\
I &
\end{tikzcd}$$

We must show that $\mathcal{C}/I[-, \mathcal{P}g] \cong \mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I)[g, J_I(-)]$ (where here we write J_I for the inclusion functor from $(\mathcal{C}/I)^{\mathrm{op}}$ to $\mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I)$). Consider any object $A \xrightarrow{f} I$ in $\mathcal{C}/I$. By the pullback defining $\mathcal{P}g$, we have that

$$\mathcal{C}/I[f, \mathcal{P}g] \cong \{A \xrightarrow{y} \mathcal{P}B \mid \langle g_! \circ y,\ \{\cdot\} \circ f\rangle \text{ factors through } \subseteq_I\ \}$$

But any $A \xrightarrow{y} \mathcal{P}B$ is contained in the above set if and only if the left-hand edge below factors through the right-hand edge.

$$\begin{tikzcd}
P \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & U \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \in_I \arrow[d, tail] & I \arrow[l] \arrow[d, tail, "\Delta"] \arrow[dl, phantom, "\llcorner", very near start] & A \arrow[l, "f"'] \arrow[d, tail, "{\langle f, 1\rangle}"] \arrow[dl, phantom, "\llcorner", very near start] \\
I \times A \arrow[r, "1 \times y"] & I \times \mathcal{P}B \arrow[r, "1 \times g_!"] & I \times \mathcal{P}I & I \times I \arrow[l, "1 \times \{\cdot\}"'] & I \times A \arrow[l, "1 \times f"']
\end{tikzcd}$$

By the definition of $g_!$, the mono $U \rightarrowtail I \times \mathcal{P}B$ is obtained as the image factorization of:

$$\in_B \rightarrowtail B \times \mathcal{P}B \xrightarrow{g \times 1} I \times \mathcal{P}B$$

So, by the stability of images under pullback, $P \rightarrowtail I \times A$ is the mono part of the factorization of:

$$h \;=\; R \rightarrowtail B \times A \xrightarrow{g \times 1} I \times A$$

where $R \rightarrowtail B \times A$ represents the small relation for which $\chi_R = y$. Then $P \rightarrowtail I \times A$ factors through $\langle f, 1\rangle$ if and only if h does. So we have shown that:

$$\mathcal{C}/I[f, \mathcal{P}g] \cong \{R \in \mathcal{R}_{\mathcal{S}}(\mathcal{C})[B, A] \mid f \circ r_2 = g \circ r_1\}$$

where $\langle r_1, r_2\rangle$ are the components of $R \rightarrowtail B \times A$. But the right-hand set above is $\mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I)[g, f]$ as required. The naturality of the isomorphism is routine to check, and is inherited from the naturality of the map from any $R \in \mathcal{R}_{\mathcal{S}}(\mathcal{C})[B, A]$ to $\chi_R : A \longrightarrow \mathcal{P}B$. $\square$

Lemma 5.19 *For any $h : I \longrightarrow J$, the reindexing functor $h^* : \mathcal{C}/J \to \mathcal{C}/I$ is logical.*

PROOF. Given $h : I \longrightarrow J$, the reindexing functor $h^* : \mathcal{C}/J \to \mathcal{C}/I$ has the usual left adjoint Σ_h given by postcomposition with h. We shall define functors $\mathcal{R}h^* : \mathcal{R}_{\mathcal{S}_J}(\mathcal{C}/J) \longrightarrow \mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I)$ and $\mathcal{R}\Sigma_h : \mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I) \longrightarrow \mathcal{R}_{\mathcal{S}_J}(\mathcal{C}/J)$ with $\mathcal{R}\Sigma_h$ right (sic) adjoint to $\mathcal{R}h^*$. The action of $\mathcal{R}h^*$ on objects is as for h^*. Given objects $A \xrightarrow{f} J$ and $B \xrightarrow{g} J$ in the slice $\mathcal{C}/J$, a small relation $R \rightarrowtail f \times g$ in $\mathcal{R}_{\mathcal{S}_J}(\mathcal{C}/J)$ represented by $R \rightarrowtail A \times_J B$ is mapped by $\mathcal{R}h^*$ to the relation $R' \rightarrowtail h^*(A) \times_I h^*(B)$ given by the pullback below.

$$\begin{tikzcd}
R' \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & R \arrow[d, tail] \\
h^*(A) \times_I h^*(B) \arrow[r] & A \times_J B
\end{tikzcd}$$

This is indeed a small relation by (S2).

The functor $\mathcal{R}\Sigma_h$ behaves like Σ_h on objects. Given $A \xrightarrow{f} I$ and $B \xrightarrow{g} I$ in $\mathcal{C}/I$, a morphism $R \rightarrowtail f \times g$ in $\mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I)$ represented by $R \rightarrowtail A \times_I B$ is mapped by $\mathcal{R}\Sigma_h$ to the relation from $\Sigma_h(f)$ to $\Sigma_h(g)$ in $\mathcal{R}_{\mathcal{S}_J}(\mathcal{C}/J)$ represented by the evident composite:

$$R \rightarrowtail A \times_I B \rightarrowtail A \times_J B$$

This is a small relation by (S3), because the composite

$$R \rightarrowtail A \times_I B \rightarrowtail A \times_J B \longrightarrow B$$

is small since $R \rightarrowtail f \times g$ was assumed a small relation in $\mathcal{C}/I$.

It is easily checked that the above definitions are good definitions of functors. To verify the adjunction, noting that the pullback square

$$\begin{tikzcd}
h^*A \arrow[r, "f^*h"] \arrow[d, "h^*f"'] \arrow[dr, phantom, "\lrcorner", very near start] & A \arrow[d, "f"] \\
I \arrow[r, "h"] & J
\end{tikzcd}$$

defines $(\mathcal{R}\Sigma_h)(\mathcal{R}h^*)(f) = h \circ h^* f$, the components of the unit of the adjunction are the maps $J_I(f^*h)$, where J_I is the functor $(\mathcal{C}/I)^{\mathrm{op}} \to \mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I)$. It is the contravariance of this functor that is reponsible for $\mathcal{R}\Sigma_h$ being a right adjoint.

It is straightforward from the definitions of $\mathcal{R}h^*$ and $\mathcal{R}\Sigma_h$ that both squares of functors below commute (up to equality).

$$\begin{tikzcd}
(\mathcal{C}/I)^{\mathrm{op}} \arrow[r, "J_I"] \arrow[d, shift left=2, "\Sigma_h"] \arrow[d, phantom, "\dashv"] & \mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I) \arrow[d, shift left=2, "\mathcal{R}\Sigma_h"] \arrow[d, phantom, "\dashv"] \\
(\mathcal{C}/J)^{\mathrm{op}} \arrow[r, "J_J"] \arrow[u, shift left=2, "h^*"] & \mathcal{R}_{\mathcal{S}_J}(\mathcal{C}/J) \arrow[u, shift left=2, "\mathcal{R}h^*"]
\end{tikzcd}$$

By Lemma 5.18, axiom (P) holds in $\mathcal{C}/I$ and $\mathcal{C}/J$, so J_I and J_J have left adjoints P_I and P_J respectively. Since the square of right adjoints above commutes, so does the square of associated left adjoints (up to natural isomorphism), i.e. $h^*P_J \cong P_I\mathcal{R}h^*$. But then we have $h^*P_J J_J \cong P_I \mathcal{R}h^* J_J = P_I J_I h^*$. So h^* maps the small powerclass $P_J J_J$ in $\mathcal{C}/J$ to the small powerclass $P_I J_I$ in $\mathcal{C}/I$. A similar argument shows that h^* preserves membership relations, as these are the components of the units of the $P \dashv J$ adjunctions. $\square$

Lemma 5.20 *If A is a small object in $\mathcal{C}$ then it is exponentiable, i.e. every exponential B^A exists.*

PROOF. Write $\langle\langle a,b\rangle, z\rangle$ for the components of the membership relation $\in_{A\times B} \rightarrowtail (A\times B)\times \mathcal{P}(A\times B)$. Since this is a small relation, z is a small map. But $z = \pi_2 \circ \langle a, z\rangle$ so, by (S3), $\langle a,z\rangle\colon\ \in_{A\times B} \longrightarrow A\times\mathcal{P}(A\times B)$ is also a small map. Therefore the mono $\langle b, \langle a,z\rangle\rangle$ determines a small relation $\in_{A\times B} \rightarrowtail B\times(A\times\mathcal{P}(A\times B))$. We write $r : A\times\mathcal{P}(A\times B) \longrightarrow \mathcal{P}B$ for its characteristic map. Now define $U \rightarrowtail A\times\mathcal{P}(A\times B)$ by pullback:

$$\begin{tikzcd}
U \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & B \arrow[d, tail, "\{\cdot\}"] \\
A\times\mathcal{P}(A\times B) \arrow[r, "r"] & \mathcal{P}B
\end{tikzcd} \tag{13}$$

By Lemma 5.10, $\{\cdot\}$ is a small mono, hence so is $U \rightarrowtail A\times\mathcal{P}(A\times B)$. The projection $\pi_2 : A\times\mathcal{P}(A\times B) \longrightarrow \mathcal{P}(A\times B)$ is small because A is a small object, so the composite map $U \longrightarrow \mathcal{P}(A\times B)$ is small, i.e. $U \rightarrowtail A\times\mathcal{P}(A\times B)$ is a small relation. We write $\chi_U : \mathcal{P}(A\times B) \longrightarrow \mathcal{P}A$ for the characteristic map of this small relation. Let $\ulcorner A\urcorner : \mathbf{1} \longrightarrow \mathcal{P}A$ be the inverse image $(!_A)^{-1}$

of $!_A \colon A \longrightarrow \mathbf{1}$, which exists because A is a small object. Define B^A as the pullback:

$$\begin{tikzcd}
B^A \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & \mathbf{1} \arrow[d, "\ulcorner A \urcorner"] \\
\mathcal{P}(A \times B) \arrow[r, "\chi_U"] & \mathcal{P}A
\end{tikzcd} \tag{14}$$

The above construction is similar to the construction of exponentials from powerobjects in a topos, see e.g. [[REFS]], only taking account of smallness. The verification that the construction indeed gives an exponential is also similar, and thus omitted. □

To prove Statement 3 of Proposition 5.17, suppose $h \colon I \longrightarrow J$ is small. The required Π_h maps any object $f \colon A \longrightarrow I$ in the slice $\mathcal{C}/I$ to the exponential $(h \circ f)^h$ calculated in $\mathcal{C}/J$. This exponential exists by relativizing Lemma 5.20 to the slice category $\mathcal{C}/J$, using that h is a small object in this category. A standard argument shows that Π_h indeed defines a right adjoint to h^*, cf. [[REFS]]. This completes the proof of Proposition 5.17.

6 Additional axioms

In this section, we consider several independent ways of adding additional properties and structure to categories with basic class structure. Throughout, let $\mathcal{C}$ be a category with collection of small maps $\mathcal{S}$ giving basic class structure.

6.1 Powerset

The notion of basic class structure provides a basis for considering category-theoretic models for a range of constructive set theories, including predicative set theories, see [[REFS]]. In the present paper, we are interested in models of the (impredicative) set theories associated with elementary toposes. This requires a further axiom on top of basic class structure: the Powerset axiom.

Proposition 6.1 *The following are equivalent.*

(1) $\subseteq_A \rightarrowtail \mathcal{P}A \times \mathcal{P}A$ *is a small relation.*
(2) In any slice $\mathcal{C}/I$*, the operation* $\mathcal{P}(-)$ *preserves small objects.*

PROOF. Assume 1. To show 2 we provide an alternative construction of small powerobjects in slice categories, available for small maps only. Given a small $g\colon B \longrightarrow I$, we claim that $\mathcal{P}g$ in $\mathcal{C}/I$ can be defined as the left edge of the diagram below.

$$\begin{tikzcd}
V \arrow[r] \arrow[d, tail] \arrow[dr, phantom, very near start, "\lrcorner"] & \subseteq_B \arrow[d, tail] \\
\mathcal{P}B \times I \arrow[r, "1 \times g^{-1}"] \arrow[d, "\pi_2"'] & \mathcal{P}B \times \mathcal{P}B \\
I &
\end{tikzcd} \tag{15}$$

This left edge is small by the assumption and (S2).

For the claim, we show that $\mathcal{C}/I[-,\mathcal{P}g] \cong \mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I)[g, J_I(-)]$. We have that

$$\mathcal{C}/I[f,\mathcal{P}g] \cong \{A \xrightarrow{y} \mathcal{P}B \mid \langle y,\, g^{-1} \circ f\rangle \text{ factors through } \subseteq_B\}.$$

But any $A \xrightarrow{y} \mathcal{P}B$ is contained in the above set if and only if the left-hand edge below factors through the right-hand edge.

$$\begin{tikzcd}
R \arrow[r] \arrow[d, tail] \arrow[dr, phantom, very near start, "\lrcorner"] & \in_B \arrow[d, tail] & B \arrow[l] \arrow[d, tail, "{\langle 1, g\rangle}"] \arrow[dl, phantom, very near start, "\llcorner"] & P \arrow[l] \arrow[d, tail] \arrow[dl, phantom, very near start, "\llcorner"] \\
B \times A \arrow[r, "1 \times y"] & B \times \mathcal{P}B & B \times I \arrow[l, "1 \times g^{-1}"'] & B \times A \arrow[l, "1 \times f"']
\end{tikzcd}$$

As maps $y : A \longrightarrow \mathcal{P}B$ are in one-to-one correspondence with small relations, it is immediate that

$$\mathcal{C}/I[f,\mathcal{P}g] \cong \{R \rightarrowtail B \times A \mid R \text{ a small relation and } f \circ r_2 = g \circ r_1\} ,$$

where $\langle r_1, r_2\rangle$ are the components of $R \rightarrowtail B \times A$. As required, the right-hand set above is $\mathcal{R}_{\mathcal{S}_I}(\mathcal{C}/I)[g, J_I(f)]$. The naturality of the established isomorphism is routine.

For the converse, we show below that

$$\subseteq_B \rightarrowtail \mathcal{P}B \times \mathcal{P}B \xrightarrow{\pi_2} \mathcal{P}B \tag{16}$$

is isomorphic to $\mathcal{P}(\pi)$ in $\mathcal{C}/(\mathcal{P}B)$, where π is

$$\in_B \rightarrowtail B \times \mathcal{P}B \xrightarrow{\pi_2} \mathcal{P}B\ . \tag{17}$$

As π is small, the smallness of the relation $\subseteq_B \rightarrowtail \mathcal{P}B \times \mathcal{P}B$ then follows from 1.

It remains to show that $\mathcal{P}(\pi)$ is isomorphic to (16). Consider π and $\pi' = \pi_2\colon B \times \mathcal{P}B \longrightarrow \mathcal{P}B$ as objects of $\mathcal{C}/(\mathcal{P}B)$, together with the associated mono $\pi \overset{\in_B}{\rightarrowtail} \pi'$. Applying the covariant small powerobject functor on $\mathcal{C}/(\mathcal{P}B)$, we obtain

$$\mathcal{P}(\pi) \overset{(\in_B)_!}{\rightarrowtail} \mathcal{P}(\pi') \ ,$$

which is a mono by Proposition 5.12. We shall prove that the subset relation is given by the image of this mono under the forgetful from $\mathcal{C}/(\mathcal{P}B)$ to $\mathcal{C}$.

By Lemma 5.19, the object $\mathcal{P}\pi'$ is (isomorphic to) $\mathcal{P}B \times \mathcal{P}B \xrightarrow{\pi_2} \mathcal{P}B$ with the membership relation:

$$\in_B \times \mathcal{P}B \overset{\in_B \times 1}{\rightarrowtail} (B \times \mathcal{P}B) \times \mathcal{P}B \ \cong \ (B \times \mathcal{P}B) \times_{\mathcal{P}B} (\mathcal{P}B \times \mathcal{P}B)$$

Thus $(\in_B)_!$ is indeed a binary relation on $\mathcal{P}B$.

Writing $\subseteq'_B \xrightarrow{\mathcal{P}(\pi)} \mathcal{P}B$ for the object $\mathcal{P}(\pi)$ of $\mathcal{C}/(\mathcal{P}B)$, we must show that the mono $\subseteq'_B \overset{(\in_B)_!}{\rightarrowtail} \mathcal{P}B \times \mathcal{P}B$ represents the subset relation. Accordingly, consider any $A \xrightarrow{\langle f,g\rangle} \mathcal{P}B \times \mathcal{P}B$. Define $P \overset{p}{\rightarrowtail} B \times A$ and $Q \overset{q}{\rightarrowtail} B \times A$ by pullback as in Definition 5.15. We must show that p factors through q if and only if $\langle f, g\rangle$ factors through $\subseteq'_B \overset{(\in_B)_!}{\rightarrowtail} \mathcal{P}B \times \mathcal{P}B$.

First, $p\colon P \rightarrowtail B \times A$ defines a mono $P \rightarrowtail \pi' \times g$ in $\mathcal{C}/(\mathcal{P}B)$, since $B \times A \cong \pi' \times_{\mathcal{P}B} g$. Using the explicit description of $\mathcal{P}(\pi')$ above, the pullback diagram below shows that $P \rightarrowtail \pi' \times g$ is a small relation with characteristic map $\langle f, g\rangle : g \longrightarrow \mathcal{P}(\pi')$ in $\mathcal{C}/(\mathcal{P}B)$.

$$\begin{tikzcd}
P \arrow[r, "{\langle 1,\, g \circ \pi_2 \circ p\rangle}"] \arrow[d, tail, "p"'] \arrow[dr, phantom, "\lrcorner", very near start] & P \times \mathcal{P}B \arrow[r] \arrow[d, tail, "p \times 1"'] \arrow[dr, phantom, "\lrcorner", very near start] & \in_B \times \mathcal{P}B \arrow[d, tail, "\in_B \times 1"] \\
B \times A \arrow[r, "{\langle 1, g \circ \pi_2\rangle}"] & (A \times B) \times \mathcal{P}B \arrow[r, "(1 \times f) \times 1"] & (\mathcal{P}B \times B) \times \mathcal{P}B
\end{tikzcd} \tag{18}$$

Next, consider the membership relation: $\in_\pi \ \rightarrowtail \pi \times_{\mathcal{P}B} \mathcal{P}(\pi)$ in $\mathcal{C}/(\mathcal{P}B)$. This exists in $\mathcal{C}$ as a mono $\in_\pi \ \rightarrowtail \ \in_B \times_{\mathcal{P}B} \subseteq'_B$. By the definition of the small powerobject functor, $\subseteq'_B \overset{(\in_B)_!}{\rightarrowtail} \mathcal{P}B \times \mathcal{P}B$ fits into the right-hand pullback below.

$$
\begin{tikzcd}
P \arrow[rr] \arrow[d, tail, "m"'] \arrow[dr, phantom, "\lrcorner", very near start] & & \in_\pi \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \in_B \times \mathcal{P}B \arrow[ddd, tail, "\in_B \times 1"] \\
Q \arrow[r, "\cong"] \arrow[dd, tail, "q"'] \arrow[dr, phantom, "\lrcorner", very near start] & \in_B \times_{\mathcal{P}B} A \arrow[r, "1 \times_{\mathcal{P}B} z"] & \in_B \times_{\mathcal{P}B} \subseteq'_B \arrow[d, "\in_B \times_{\mathcal{P}B} 1"] & \\
 & & (\mathcal{P}B \times B) \times_{\mathcal{P}B} \subseteq'_B \arrow[d, "\cong"] & \\
B \times A \arrow[rr, "1 \times z"] & & B \times \subseteq'_B \arrow[r, "1 \times (\in_B)_!"] & B \times \mathcal{P}B \times \mathcal{P}B
\end{tikzcd}
$$

Suppose now that p factors through q by $P \overset{m}{\rightarrowtail} Q$. By the pullback definition of q, we have $Q \cong \pi \times_{\mathcal{P}B} g$. The projection of $m\colon P \rightarrowtail \pi \times g$ to g is equal to the projection of the small relation $P \rightarrowtail \pi' \times g$ to g. Hence the relation $m\colon P \rightarrowtail \pi \times g$ in $\mathcal{C}/(\mathcal{P}B)$ is small, and its characteristic map $z\colon A \longrightarrow \subseteq'$ completes the two left-hand pullbacks in $\mathcal{C}$ above, where we write $\in_B \times_{\mathcal{P}B} A$ for the domain of $\pi \times g$ in $\mathcal{C}/(\mathcal{P}B)$. The outer pullback above shows that $(\in_B)_! \circ z$ is the characteristic map of $P \rightarrowtail \pi' \times g$. Hence $(\in_B)_! \circ z = \langle f, g \rangle$. So $\langle f, g \rangle$ does indeed factor through $(\in_B)_!$.

Conversely, suppose $\langle f, g \rangle$ factors through $(\in_B)_!$ by some $z\colon A \longrightarrow \subseteq'$. Then $(\in_B)_! \circ z$ is the characteristic map of $P \rightarrowtail \pi' \times g$. Thus the outer pullback of diagram (18) can be reconstructed in stages as in the diagram above to provide the required $m\colon P \rightarrowtail Q$ showing that p factors through q. $\square$

(Powerset) The equivalent conditions of Proposition 6.1 hold.

Since every slice category of a slice category $\mathcal{C}/I$ is itself a slice category of $\mathcal{C}$, it follows from statement 2 of Proposition 6.1 that if the Powerset axiom holds in $\mathcal{C}$ then it holds in every slice.

Proposition 6.2 *If the Powerset axiom holds then*

(1) If A, B are small objects then so is the exponential B^A.
(2) If $h\colon I \longrightarrow J$ is small, then $\Pi_h\colon \mathcal{C}/I \to \mathcal{C}/J$ preserves small objects.

PROOF. For 1, by the Powerset axiom, the object $\mathcal{P}(A \times B)$ is small. Hence any map out of $\mathcal{P}(A \times B)$ is small, including $\chi_U \colon \mathcal{P}(A \times B) \longrightarrow \mathcal{P}A$ from the proof of Lemma 5.20. That B^A is small now follows from (S2) using the pullback in diagram (14).

Statement 2 follows from statement 1, because $\Pi_h(f)$ is given by the exponential $(h \circ f)^h$ in $\mathcal{C}/J$. $\square$

Proposition 6.3 *If the Powerset axiom holds then the category of sets $\mathcal{E}_\mathcal{S}(\mathcal{C})$ is an elementary topos.*

PROOF. The category of sets has finite limits inherited from $\mathcal{C}$. It is cartesian closed by Proposition 6.2. As in the discussion below (11), the object $\Omega = \mathcal{P}\mathbf{1}$ classifies subobjects defined by small monomorphisms in $\mathcal{C}$. Since the inclusion $\mathcal{E}_\mathcal{S}(\mathcal{C}) \hookrightarrow \mathcal{C}$ preserves limits, it preserves monos. Thus, for a small object B, a mono $A \rightarrowtail B$ in $\mathcal{C}$ is small if and only if it is a mono in $\mathcal{E}_\mathcal{S}(\mathcal{C})$. By the Powerset axiom, Ω is itself a small object. Therefore it is a subobject classifier in $\mathcal{E}_\mathcal{S}(\mathcal{C})$. $\square$

A related consequence of the Powerset axiom is that the category $\mathcal{E}_\mathcal{S}(\mathcal{C})$ has basic class structure itself, with all maps small. Further, the inclusion $\mathcal{E}_\mathcal{S}(\mathcal{C}) \hookrightarrow \mathcal{C}$ is logical.

6.2 Separation

As observed in the discussion below equation (11), in a category with basic class structure, a monomorphism is small if and only if it is regular. This gives a restricted separation principle of the following form: if B is a small object and $A \rightarrowtail B$ is a regular subobject, then A is a small object. In other words, certain (i.e. regular) subclasses of sets are sets. The Separation axiom drops the restriction to regular subobjects, and asserts (in every slice) that arbitrary subobjects of small objects are small.

Proposition 6.4 *The following are equivalent.*

(1) Every monomorphism in $\mathcal{C}$ is small.
(2) Every monomorphism in $\mathcal{C}$ is regular.
(3) In every slice $\mathcal{C}/I$, subobjects of small objects are small.
(4) $\mathcal{C}$ has a subobject classifier.

PROOF. As monomorphisms are small if and only if regular, the equivalence of 1 and 2 is immediate. For 1 $\implies$ 3, if $A \rightarrowtail B$ is a subobject of a small object $B \longrightarrow I$ in $\mathcal{C}/I$, then $A \rightarrowtail I$ is indeed small as a composition of small maps. Conversely, every monomorphism $A \rightarrowtail B$ is a subobject of the small object $1_B \colon B \longrightarrow B$ of $\mathcal{C}/B$. That 1 $\implies$ 4, is immediate from the fact that $\Omega = \mathcal{P}\mathbf{1}$ classifies subobjects defined by small monos, see the discussion below equation (11). Finally, when $\mathcal{C}$ has a subobject classifier, it follows directly that every monomorphism is regular, thus 4 $\implies$ 2. □

(Separation) The equivalent conditions of Proposition 6.4 hold.

Using item 1 of Proposition 6.4, it is immediate that the Separation axiom is preserved by slicing.

In [11,**?**], categories with basic class structure satisfying both Powerset and Separation were considered under the description *categories with class(ic) structure*. As shown in [11], it is possible to give an economical axiomatization of such categories, using just axioms (C1),(S1), (S2) and (P) together with the Powerset and Separation axioms (the latter in the guise that all monomorphisms are small). Axioms (C2), (C3), (S3), (S4) and (S5) are then all derivable.

6.3 Infinity

Proposition 6.5 *If the Powerset axiom holds then the following are equivalent:*

(1) There is a small object I with a monomorphism $\mathbf{1} + I \rightarrowtail I$.
(2) The category $\mathcal{E}_{\mathcal{S}}(\mathcal{C})$ of small objects has a natural numbers object.

PROOF. $\mathcal{E}_{\mathcal{S}}(\mathcal{C})$ is an elementary topos. □

(Infinity) The equivalent conditions of Proposition 6.5 hold.

Using item 1 of Proposition 6.5, it is clear that the axiom of infinity is preserved by taking slice categories.

The axiom of infinity ensures that the category of sets has an nno. It does not follow that this is an nno in the category of classes, $\mathcal{C}$, which need not even possess an nno. This situation is analogous to the presence of restricted induction but not full induction in BIST, see Section 2. For BIST, the addition of the axiom of Separation is sufficient to ensure the derivability of full induction (Corollary 2.11). Similarly, as outlined in [11], when $\mathcal{C}$ satisfies Separation, it does hold that an nno in the category of sets is automatically an nno in the category of classes.

6.4 Collection

In set theory, see Section 2, the axiom of Collection asserts that, for every total relation R between a set A and a class Y (i.e. a relation satisfying $\forall x \in A. \exists y \in Y. R(x, y)$), there exists a subset B of Y such that R restricts to a total relation between A and B. Without loss of generality, in place of total relations, one can consider surjective class functions from a class X onto a set A. Trivially, any such class function is a total relation between A and X. Conversely, given a total relation between A and Y, one can use the class $X = \{(x, y) \mid x \in A, y \in Y, R(x, y)\}$, and consider its projection onto A. Collection, can now be rephrased as, for any surjective class function from a class X onto a set A, there exists a set B and a function $B \to X$ such that the composite $B \to X \to A$ is still surjective. (In the presence of Replacement, it is unnecessary to demand that B is a subset of X, since such a subset can be found by taking the image of $B \to X$.)

The above discussion, suggests formulating a collection property in $\mathcal{C}$ as follows. For any regular epi $X \twoheadrightarrow A$, where A is a small object, there exists a map $B \longrightarrow X$, where B is small, such that the composite $B \longrightarrow X \twoheadrightarrow A$ is a regular epi. However, this is not quite right. Logically, it should be sufficient for the existence of B and $B \longrightarrow X$ to hold in the internal logic of $\mathcal{C}$, and this does not require the real-world existence of corresponding external object and map. Second, as with any axiom, it is necessary to ensure that the property it asserts is preserved by slicing. Both modifications are taken account of simultaneously in property 1 of the proposition below, which is due to Joyal and Moerdijk [7].

Proposition 6.6 *The following are equivalent.*

(1) For every small $A \longrightarrow I$ and regular epi $X \twoheadrightarrow A$, there exists a quasi-pullback diagram[20]

$$\begin{tikzcd} B \arrow[r] \arrow[d] & X \arrow[r, two heads] & A \arrow[d] \\ J \arrow[rr, two heads] & & I \end{tikzcd} \tag{19}$$

[20] Diagram (19) is a *quasi-pullback* if it commutes and the canonical map $B \longrightarrow J \times_I A$ to the actual pullack is a regular epi.

with $J \twoheadrightarrow I$ regular epi and $B \longrightarrow J$ small.
(2) If $e\colon X \twoheadrightarrow Y$ is a regular epi then so is $e_!\colon \mathcal{P}X \twoheadrightarrow \mathcal{P}Y$.

This is Proposition 3.7 of [7], the proof of which goes through in our setting.

(Collection) The equivalent conditions of Proposition 6.6 hold.

It is a straightforward consequence of property 1 of Proposition 6.6 that the Collection axiom is preserved by taking slice categories.

6.5 Universes and universal objects

All the axioms we have considered up to now are compatible with the assumption that all objects of $\mathcal{C}$ (and hence all maps) are small, in which case $\mathcal{C}$ is an elementary topos and $\mathcal{P}X$ is the powerobject of X.

For elementary toposes, a version of Cantor's diagonalization argument shows that it is inconsistent to have an object X with a mono $\mathcal{P}X \rightarrowtail X$. Thus the following notions, introduced in [11], ensure that $\mathcal{C}$ (if consistent) has a non-small object U.

Definition 6.7 (Universe) A *universe* is an object U together with a mono $i\colon \mathcal{P}U \rightarrowtail U$.

Definition 6.8 (Universal object) A *universal object* is an object U such that, for every object X, there exists a mono $X \rightarrowtail U$.

The notion of universe captures the idea that $\mathcal{C}$, which may be seen as a "typed" world of classes, contains an object U, which may be considered as an "untyped" universe of "sets" and "non-sets" with the mono $\mathcal{P}U \rightarrowtail U$ singling out the subcollection of sets in U. We remark that this method of obtaining an untyped set-theoretic universe within a typed world of classes may be seen as analogous to Dana Scott's identification of models of the untyped λ-calculus as reflexive objects in cartesian closed categories [**?**].

The stronger notion of universal object enforces that every class can be seen as a subclass of the untyped universe U. This situation occurs naturally in first-order set theory, where classes are defined as subcollections of an assumed universe.

As observed in [11], any universe U acts as a universal object in a derived category with basic class structure. Indeed, defining $\mathcal{C}_{\leq U}$ and $\mathcal{S}_{\leq U}$ to be the full subcategories of $\mathcal{C}$ and $\mathcal{S}$ on subobjects (in $\mathcal{C}$) of U, we have:

Proposition 6.9 *If U is a universe then $\mathcal{C}_{\leq U}$ with $\mathcal{S}_{\leq U}$ has basic class struc-*

ture with universal object U.

PROOF (outline). The main points are to observe that $\mathcal{C}_{\leq U}$ is closed under finite product and $\mathcal{P}(-)$ in $\mathcal{C}$. The latter is a consequence of Proposition 5.12. For the former, Kuratowski pairing (cf. the proof of Lemma 3.9), defines a mono $U \times U \rightarrowtail \mathcal{P}\mathcal{P}U$, and we have just seen that $\mathcal{P}\mathcal{P}U \rightarrowtail U$. Thus we obtain a composite mono $U \times U \rightarrowtail U$, from which the closure of $\mathcal{C}_{\leq U}$ under finite products follows easily. □

By this result, universal objects are essentially just as general as universes, and so it is no real restriction to consider the former in preference to the latter.

In fact universal objects enjoy useful properties that do not hold of arbitrary universes. One such property, again taken from [11], is that the map $\pi_U \colon \in_U \longrightarrow \mathcal{P}U$, which one may think of as giving the $\mathcal{P}U$-indexed family of all sets, is a generic small map in the following sense.

Proposition 6.10 *If U is a universal object, then a map $f \colon X \longrightarrow Y$ is small if and only there exists $g \colon Y \longrightarrow \mathcal{P}U$ fitting into a pullback:*

$$\begin{tikzcd}
X \arrow[r] \arrow[d, "f"'] & \in_U \arrow[d, "\pi_U"] \\
Y \arrow[r, "g"] & \mathcal{P}U
\end{tikzcd}$$

PROOF. For the interesting direction, given f, take a mono $m \colon X \rightarrowtail U$ and define $g = m_! \circ f^{-1}$. It is easily checked that this determines a pullback as above. □

Obviously, every logical functor $F \colon \mathcal{C} \to \mathcal{C}'$ between categories with basic class structure preserves universes. It need not, howerver, preserve universal objects. We say that a functor F is *cofinal* if for every object Y of $\mathcal{C}'$ there exists $X \in \mathcal{C}$ with $Y \rightarrowtail FX$. Easily, a logical functor preserves universal objects if and only if it is cofinal.

It is readily checked that when $\mathcal{C}$ has a universal object then so does every slice category $\mathcal{C}/I$, and for every $f \colon I \longrightarrow J$ the reindexing functor $f^* \colon \mathcal{C}/J \to \mathcal{C}/I$ is cofinal.

Having now considered several additional properties one may require on top of basic class structure, we enshrine in a definition the main properties that will henceforth be relevant to our study of category-theoretic models of the set theory BIST^-.

Definition 6.11 (Category of classes) A *category of classes* is a positive Heyting category $\mathcal{C}$ together with basic class structure $\mathcal{S}$ satisfying the Powerset axiom and a universal object U.

This definition should be considered local to the present paper. In other situations, different combinations of basic properties might well be equally deserving of the appellation "category of classes".

7 Interpreting set theory in a category of classes

In this section, we interpret the first-order language of Section 2 in a category of classes. We use the universal object U as an untyped universe of sets (and non-sets), and interpret the logic using the internal logic of $\mathcal{C}$. We shall see that the set theory validated in this way is exactly BIST^-. Moreover, the additional axioms of Section 6 are related to their set-theoretic analogues from Section 2.

Given a category $\mathcal{C}$ with basic class structure and universe $i\colon \mathcal{P}U \longrightarrow U$, we interpret the first-order language of Section 2 in the internal logic of $\mathcal{C}$ by interpreting first-order variables and quantifiers as ranging over U. Thus a formula $\phi(x_1, \dots, x_k)$ is given an interpretation

$$[\![x_1, \dots, x_k \mid \phi]\!] \rightarrowtail \overbrace{U \times \cdots \times U}^{k} \ ,$$

which is determined by the interpretations of the basic relations $x \in y$ and $\mathsf{S}(x)$, defined as follows.

$$\begin{aligned} [\![x \mid \mathsf{S}(x)]\!] &= \mathcal{P}U \overset{i}{\rightarrowtail} U \\ [\![x, y \mid x \in y]\!] &= \in_U \rightarrowtail U \times \mathcal{P}U \xrightarrow{1_U \times i} U \times U \ . \end{aligned}$$

We write $\mathcal{C} \models \phi$ to mean that a sentence ϕ is validated in this interpretation, i.e. that $[\![\phi]\!] \cong \mathbf{1}$.[21]

[21] Strictly speaking, $\mathcal{C} \models \phi$ is an abuse of notation, since the interpretation is determined by all of $\mathcal{C}, \mathcal{S}, \mathcal{P}(-), U$ and i.

Theorem 7.1 (Soundness and completeness for class-category semantics) *For any theory $\mathcal{T}$ and sentence ϕ, the following are equivalent.*

(1) $\mathrm{BIST}^{-} + \mathcal{T} \vdash \phi$.
(2) $\mathcal{C} \models \phi$*, for all categories of classes* $\mathcal{C}$ *satisfying* $\mathcal{C} \models \mathcal{T}$.

7.1 Soundness of class-category semantics

To prove the soundness direction of Theorem 7.1, it is enough to verify that the axioms of BIST^{-} (Fig. 1) are validated by any category of classes, since the soundness of intuitionistic logic is a standard consequence of Proposition 5.4. We present a few illustrative cases.

Extensionality: $\mathsf{S}(x) \wedge \mathsf{S}(y) \wedge (\forall z.\, z \in x \leftrightarrow z \in y) \rightarrow x = y$

Suppose given arbitrary $\langle a, b\rangle : Z \to U \times U$ factoring through the subobject

$$[\![x, y \mid \mathsf{S}(x) \wedge \mathsf{S}(y) \wedge \forall z.(z \in x \leftrightarrow z \in y)]\!] \rightarrowtail U \times U$$

then by the first two conjuncts there are small relations $[\![z, x \mid z \in a(x)]\!]$ and $[\![z, y \mid z \in b(y)]\!]$ on $U \times Z$, and by the third these satisfy

$$[\![z, x \mid x \in a(z)]\!] \;=\; [\![z, y \mid y \in b(z)]\!]$$

as subobjects of $U \times U$. Whence $a = b$ by the uniqueness of characteristic maps in axiom (P) on basic class structure.

To verify the axioms involving the "set-many" quantifier, we make use of the following lemma.

Lemma 7.2 *For any formula $\varphi(x_1, \ldots, x_k, y)$, the subobject $[\![\vec{x} \mid \text{Ƨ}y.\, \phi]\!] \rightarrowtail U^k$ is given by*

$$\{z \colon U^k \mid p^{-1}(z) \text{ is small}\} \;,$$

using the notation introduced below Proposition 5.11, where p is the composite

$$[\![y, \vec{x} \mid \phi]\!] \rightarrowtail U \times U^k \xrightarrow{\pi_2} U^k \;.$$

PROOF. Routine verification using the definition of the Ƨ quantifier, and the Kripke-Joyal semantics of $\mathcal{C}$. $\square$

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

We must show that

$$[\![x, \vec{w} \mid \mathsf{S}(x) \wedge (\forall y \in x.\, \reflectbox{\textsf{S}} z.\, \phi)]\!] \leq [\![x, \vec{w} \mid \reflectbox{\textsf{S}} z.\, \exists y \in x.\, \phi]\!] \text{ in } \mathrm{Sub}(U \times U^k)$$

for formulas $\phi(z, y, x, \vec{w})$, where $\vec{w}$ abbreviates a vector of k variables. For notational convenience, we give the proof for empty $\vec{w}$. The same argument works in the general case.

Consider the projection maps in the diagram below.

$$\begin{tikzcd}
{[\![z, y, x \mid \mathsf{S}(x), (\forall y \in x.\, \reflectbox{\textsf{S}} z.\, \phi), y \in x, \phi]\!]} \arrow[r, "p_{y,x}"] \arrow[d, "q_{z,x}"'] & {[\![y, x \mid \mathsf{S}(x), (\forall y \in x.\, \reflectbox{\textsf{S}} z.\, \phi), y \in x]\!]} \arrow[d, "p_x"] \\
{[\![z, x \mid \mathsf{S}(x), (\forall y \in x.\, \reflectbox{\textsf{S}} z.\, \phi), \exists y \in x.\, \phi]\!]} \arrow[r, "q_x"] & {[\![x \mid \mathsf{S}(x), (\forall y \in x.\, \reflectbox{\textsf{S}} z.\, \phi)]\!]}
\end{tikzcd}$$

The map p_x is small, because $\mathsf{S}(x)$ holds. By Lemma 7.2 and Proposition 5.11, $p_{y,x}$ is small, because for (y, x) in the codomain it holds that $\reflectbox{\textsf{S}} z.\, \phi$. Thus the composite, $p_x \circ p_{y,x}$ is small. By (S4), q_x is small. The required inclusion of subobjects now follows from Lemma 7.2 and Proposition 5.11.

The other axioms involving the "set-many" quantifier are similarly reduced to Lemma 7.2 and Proposition 5.11. Indeed, Emptyset and Pairing hold by Proposition 5.6(1), the latter also requiring (S4). The Equality axiom follows from (S3). Finally, the Powerset axiom of BIST$^-$ is a consequence of its namesake for small maps.

7.2 Completeness of class-category semantics

In fact, we shall prove the stronger statement that there exists a single category of classes $\mathcal{C}_\mathcal{T}$ such that, for any formula φ:

$$\mathcal{C}_\mathcal{T} \models \phi \quad \text{implies} \quad \mathrm{BIST}^- + \mathcal{T} \vdash \phi \ .$$

The category $\mathcal{C}_\mathcal{T}$ is constructed similarly to the *syntactic category* of the first-order theory BIST$^-$ + $\mathcal{T}$, cf. [6], D1.4. In our setting, due to the first-order definability of finite products of classes (cf. Section 2), it suffices to build the category out of formulas with at most one free variable.

Definition 7.3 The category $\mathcal{C}_\mathcal{T}$ is defined as follows.

objects $\{x|\phi\}$, where ϕ is a formula with at most x free, identified up to α-equivalence (i.e. $\{x|\phi\}$ and $\{y|\phi[y/x]\}$ are identified).

arrows $[\theta] : \{x|\phi\} \longrightarrow \{y|\psi\}$ are equivalence classes of formulas $\theta(x, y)$ that are "provably functional relations", i.e. the following hold in BIST$^-$+ $\mathcal{T}$:

$$\begin{aligned} \theta(x, y) &\to \phi(x) \wedge \psi(y) \\ \phi(x) &\to \exists y.\, \theta(x, y) \\ \theta(x, y) \wedge \theta(x, y') &\to y = y' \end{aligned}$$

with two such θ and θ' identified if $\theta \leftrightarrow \theta'$ holds in BIST$^-$+ $\mathcal{T}$.

identity $1_{\{x|\phi\}} = [x = y \wedge \phi] : \{x|\phi\} \longrightarrow \{y|\phi[y/x]\}$

composition $[\theta'(y, z)] \circ [\theta(x, y)] = [\exists y.\, \theta(x, y) \wedge \theta'(y, z)]$.

Lemma 7.4 *The syntactic category $\mathcal{C}_\mathcal{T}$ is a positive Heyting category.*

PROOF. Finite products and coproducts are given by (co)product classes as defined in Section 2. For equalizers and the regular and Heyting structures, standard arguments from categorical logic apply, cf. [6] D1.4.10. $\square$

For later use, we remark that in the proof of the lemma, one characterizes a map $[\theta(x, y)]\colon \{x|\phi\} \longrightarrow \{y|\psi\}$ in $\mathcal{C}_\mathcal{T}$ as being a regular epi if and only if it holds in BIST$^-$+ $\mathcal{T}$ that:

$$\psi(y) \to \exists x.\, \theta(x, y) \ . \tag{20}$$

Similarly, $[\theta(x, y)]$ is a mono if and only if:

$$\theta(x, y) \wedge \theta(x', y) \to x = x' \ . \tag{21}$$

Now define a map $[\theta] : \{x|\phi\} \longrightarrow \{y|\psi\}$ in $\mathcal{C}_0$ to be *small* if in BIST$^-$+ $\mathcal{T}$:

$$\psi(y) \to \text{Ƨ}x.\, \theta(x, y) \ .$$

Note that this definition is indeed independent of the choice of representative formula θ. We write $\mathcal{S}_\mathcal{T}$ for the collection of small maps.

Lemma 7.5 *The small maps $\mathcal{S}_\mathcal{T}$ in $\mathcal{C}_\mathcal{T}$ satisfy axioms (S1)–(S5).*

PROOF. For (S1) we need to show that the small maps form a subcategory. An identity map

$$[x = x' \wedge \phi(x)] : \{x|\phi(x)\} \longrightarrow \{x'|\phi(x')\}$$

is small because in bIST:

$$\phi(x) \to \text{Ƨ}x'.\, (x = x' \wedge \phi(x)) \ .$$

For composition, suppose we have the arrows:

$$[\theta(x,y)] : \{x|\phi_1\} \longrightarrow \{y|\phi_2\} \qquad\qquad [\theta'(y,z)] : \{y|\phi_2\} \longrightarrow \{z|\phi_3\}$$

and we know that:

$$\phi_2(y) \;\rightarrow\; \text{Ƨ}x.\, \theta(x,y) \qquad\qquad\qquad \phi_3(z) \;\rightarrow\; \text{Ƨ}y.\, \theta'(y,z) \ \ .$$

Then, by Indexed-Union, one has:

$$\phi_3(z) \;\rightarrow\; \text{Ƨ}x.\, \exists y.\, \theta(x,y) \wedge \theta'(y,z) \ \ ,$$

as required.

Axiom (S2) concerns pullbacks, which in $\mathcal{C}_{\mathcal{T}}$ are constructed as follows. Given $[\theta_1(x,z)]\colon \{x|\phi_1\} \longrightarrow \{z|\psi\}$ and $[\theta_2(y,z)]\colon \{y|\phi_2\} \longrightarrow \{z|\psi\}$, the pullback has vertex

$$\{(x,y) \mid \exists z.\, \theta_1(x,z) \,\wedge\, \theta_2(y,z)\} \ \ ,$$

using Kuratowski pairing, as in the definition of product classes. The pullback cone maps are the projections. Now suppose that $[\theta_2(y,z)]$ is small, i.e.

$$\psi(z) \;\rightarrow\; \text{Ƨ}y.\, \theta_2(y,z) \ \ . \tag{22}$$

We must show that the pullback along $[\theta_1(x,y)]$ is small, i.e. that

$$\phi_1(x) \;\rightarrow\; \text{Ƨ}y.\, \exists z.\, \theta_1(x,z) \,\wedge\, \theta_2(y,z) \ \ ,$$

but this follows directly from (22), because $\phi_1(x)$ implies there exists a unique z such that $\theta_1(x,z)$.

Axiom (S3) requires the diagonal $\Delta_\phi : \{x|\phi\} \longrightarrow \{x|\phi\} \times \{x|\phi\}$ to be small. But Δ_ϕ is represented by the formula $\theta(x,p)$:

$$\phi(x) \,\wedge\, p = (x,x) \ \ .$$

For this to be small, we require:

$$(\exists y,z.\, \phi(y) \wedge \phi(z) \wedge p = (y,z)) \;\rightarrow\; \text{Ƨ}x.\, p = (x,x) \ \ ,$$

equivalently:

$$\phi(y) \wedge \phi(z) \;\rightarrow\; \text{Ƨ}x.\, x = y \wedge x = z \ \ ,$$

and this follows from the Equality axiom of BIST^-.

For axiom (S4), suppose we have

$$[\theta(x,y)] : \{x|\phi_1\} \longrightarrow \{y|\phi_2\} \qquad\qquad [\theta'(y,z)] : \{y|\phi_2\} \longrightarrow \{z|\phi_3\} \ \ ,$$

with $[\theta'(y,z)] \circ [\theta(x,y)]$ small and $[\theta(x,y)]$ regular epi. By (20), the latter condition amounts to

$$\phi_2(y) \;\rightarrow\; \exists x.\, \theta(x,y) \;\; . \tag{23}$$

The former condition gives

$$\phi_3(z) \;\rightarrow\; \text{Ƨ}x.\, \exists y.\, \theta(x,y) \wedge \theta'(y,z) \;\; .$$

Thus, for any z such that $\phi_3(z)$, there is a set $\{x \mid \exists y.\, \theta(x,y) \wedge \theta'(y,z)\}$. Moreover, it holds that

$$\phi_3(z) \;\rightarrow\; \forall x \in \{x \mid \exists y.\, \theta(x,y) \wedge \theta'(y,z)\}.\, \text{Ƨ}y.\, \theta(x,y) \wedge \theta'(y,z) \;\; , \tag{24}$$

because there is in fact a unique such y. We must show that

$$\phi_3(z) \;\rightarrow\; \text{Ƨ}y.\, \theta'(y,z) \;\; .$$

By (23), the above property is equivalent to

$$\phi_3(z) \;\rightarrow\; \text{Ƨ}y.\, \exists x.\, \theta(x,y) \wedge \theta'(y,z) \;\; ,$$

which indeed follows from (24), by the Indexed-Union axiom of BIST^-.

The remaining case (S5) is left to the reader. $\square$

Using the characterization of monomorphisms (21), one easily shows that, up to isomorphism, every binary relation $R \rightarrowtail \{x|\phi\} \times \{y|\psi\}$ in $\mathcal{C}_\mathcal{T}$ is of the form $R = \{(x,y)|\rho(x,y)\}$, where $\rho(x,y)$ is a formula satisfying

$$\rho(x,y) \;\rightarrow\; \phi(x) \wedge \psi(y) \;\; ,$$

with the evident inclusion map for the morphism part. Further, the relation is small if and only if:

$$\psi(y) \;\rightarrow\; \text{Ƨ}x.\, \rho(x,y) \;\; . \tag{25}$$

Small powerobjects in $\mathcal{C}_\mathcal{T}$ are defined in the expected way by,

$$\mathcal{P}\{x|\phi\} \;=\; \{y|\mathsf{S}(y) \wedge \forall x \in y.\, \phi\} \;\; ,$$

with the membership relation given, as above, by the formula:

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

The smallness of the membership relation follows easily from (25).

Lemma 7.6 *$\mathcal{C}_\mathcal{T}$ satisfies axiom (P).*

PROOF. Suppose $R \rightarrowtail \{x|\phi\} \times \{y|\psi\}$ is a small relation, defined by $\rho(x, y)$ as above. The required map $\chi_R \colon \{y|\psi\} \longrightarrow \mathcal{P}\{x|\phi\}$ is given by $[\theta(y, z)]$ where θ is the formula:

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

The routine verification that this has the required property is left to the reader. □

Lemma 7.7 *$\mathcal{C}_{\mathcal{T}}$ satisfies the Powerset axiom.*

PROOF. The subset relation $\subseteq \ \rightarrowtail \mathcal{P}\{x|\phi\} \times \mathcal{P}\{x|\phi\}$, is given by the formula $\rho(y, z)$:

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

The smallness of this relation follows from (25) using the Powerset axiom of BIST$^-$. □

Lemma 7.8 *$\mathcal{C}_{\mathcal{T}}$ has universal object $U = \{u|u = u\}$.*

PROOF. For any object $\{x|\phi\}$, there is a canonical morphism

$$i_\phi = [\phi(x) \wedge x = u] : \{x|\phi\} \longrightarrow U \ ,$$

which is a mono by (21). □

In combination, Lemmas 7.4–7.8 show that $\mathcal{C}_{\mathcal{T}}$ is a category of classes in the sense of Definition 6.11.

To prove completeness, it is necessary to analyse the validity of first-order formulas in $\mathcal{C}_{\mathcal{T}}$. The interpretation of the first-order language in $\mathcal{C}_{\mathcal{T}}$, with respect to the canonical mono $\mathcal{P}U \rightarrowtail U$ yields, for each formula $\phi(x_1, \dots, x_n)$ a subobject

$$[\![x_1, \dots, x_n \mid \phi]\!] \rightarrowtail U^n \ ,$$

as in Section 7.1. On the other hand, ϕ also determines an object of $\mathcal{C}_{\mathcal{T}}$:

$$\{p \mid \exists x_1, \dots, x_n.\, p = (x_1, \dots, x_n) \ \wedge \ \phi\},$$

using a suitable n-ary tupling. Henceforth, we write $\{x_1, \dots, x_n \mid \phi\}$ for the above object. There is an evident mono

$$i_\varphi : \{x_1, \dots, x_n \mid \phi\} \rightarrowtail U^n \ ,$$

given by inclusion.

Lemma 7.9 *For any formula $\phi(x_1, \ldots, x_n)$,*

$$[\![x_1, \ldots, x_n \mid \phi]\!] = \{x_1, \ldots, x_n \mid \phi\}$$

as subobjects of U^n. This subobject is isomorphic to U^n if and only if

$$\mathrm{BIST}^- + \mathcal{T} \vdash \forall x_1, \ldots, x_n. \phi \ .$$

PROOF. The equality of subobjects is proved by a straightforward but tedious induction on the structure of ϕ. For the second part, it follows easily from the definition of equality between morphisms in $\mathcal{C}_\mathcal{T}$ that $\{x_1, \ldots, x_n \mid \phi\} \cong U^n$ if and only if $\mathrm{BIST}^- + \mathcal{T} \vdash \forall x_1, \ldots, x_n. \phi$. □

The completeness direction of Theorem 7.1 now follows. By Lemma 7.9, we have that $\mathcal{C}_\mathcal{T} \models \phi$ if and only if $\mathrm{BIST}^- + \mathcal{T} \vdash \phi$, for sentences ϕ. By the right-to-left implication, $\mathcal{C}_\mathcal{T}$ does indeed satisfy $\mathcal{C}_\mathcal{T} \models \mathcal{T}$. Completeness then follows from the left-to-right implication.

7.3 Additional axioms

In this section, we extend the soundness and completeness of Theorem 7.1 to relate the additional axioms on categories of classes introduced in Section 6 to the corresponding axioms extending BIST^- from Section 2.

Proposition 7.10 *For any theory $\mathcal{T}$ and sentence ϕ, the following are equivalent.*

(1) $\mathrm{BIST}^- + \mathrm{Sep} + \mathcal{T} \vdash \phi$ *(i.e.* $\mathrm{IST}^- + \mathcal{T} \vdash \phi$*).*
(2) For all categories of classes $\mathcal{C}$ satisfying Separation, $\mathcal{C} \models \mathcal{T}$ implies $\mathcal{C} \models \phi$.

PROOF. For the soundness direction, suppose $\mathcal{C}$ is a category of classes. Using Lemma 7.2, one shows that $[\![\vec{x} \mid !\phi]\!] \cong U^k$ if and only if the monomorphism $[\![\vec{x} \mid \phi]\!] \rightarrowtail U^k$ is small. If $\mathcal{C}$ satisfies the Separation axiom then all monos are small, hence indeed $\mathcal{C} \models !\phi$ for all ϕ, i.e. $\mathcal{C} \models \mathrm{Sep}$.

Conversely, for completeness, one verifies straightforwardly that if $\mathcal{T}$ contains all instances of Separation then the syntactic category $\mathcal{C}_\mathcal{T}$, defined in Section 7.2, satisfies Separation. □

One might hope for a stronger completeness theorem of the form if $\mathcal{C} \models \mathrm{Sep}$ then the category of classes $\mathcal{C}$ satisfies Separation. However, this does not hold.

The reason is that the validity of the Separation axiom of set theory only requires first-order definable monomorphisms in $\mathcal{C}$ to be small, from which it need not follow that all monos are small. We give a concrete example after Theorem 9.3 below.

Proposition 7.11 *For any theory $\mathcal{T}$ and sentence ϕ, the following are equivalent.*

(1) BIST+ $\mathcal{T} \vdash \phi$.
(2) For all categories of classes $\mathcal{C}$ satisfying infinity, $\mathcal{C} \models \mathcal{T}$ implies $\mathcal{C} \models \phi$.

PROOF. If $\mathcal{C}$ has a small natural numbers object N, then the mono $N \rightarrowtail U$ generates points

$$\begin{aligned} I &= \mathbf{1} \xrightarrow{N} \mathcal{P}U \rightarrowtail U \\ 0 &= \mathbf{1} \xrightarrow{0} N \rightarrowtail U \\ s &= \mathbf{1} \xrightarrow{s} N^N \rightarrowtail \mathcal{P}U \rightarrowtail U\,. \end{aligned}$$

With these, it is easily verified that $\mathcal{C} \models \mathrm{Inf}$.

Conversely, for completeness, suppose that $\mathcal{T}$ contains the Infinity axiom. Consider the syntactic category $\mathcal{C}_{\mathcal{T}}$. This need not satisfy infinity. However, consider the object

$$\begin{aligned} X \;=\; \{I, 0, s \mid 0 \in I \;\wedge\; s \in I^I \;\wedge\; (\forall x \in I.\, s(x) \neq 0) \\ \wedge\; (\forall x, y \in I.\, s(x) = s(y) \;\rightarrow\; x = y)\}\ , \end{aligned}$$

using the notation established above Lemma 7.9. Then it is easily seen that the slice category $\mathcal{C}_{\mathcal{T}}/X$ does satisfy condition 1 of Proposition 6.5, and hence the infinity axiom. Let ϕ be a sentence in the language of set theory. Then, writing $\mathit{Inf}(I, 0, s)$ for the formula used to define X,

$$\begin{aligned} \mathcal{C}_{\mathcal{T}}/X \models \phi \quad &\text{iff} \quad \mathrm{BIST}^{-}\!+ \mathcal{T} + \mathit{Inf}(I, 0, s) \vdash \phi \\ &\text{iff} \quad \mathrm{BIST}+ \mathcal{T} \vdash \phi\ . \end{aligned}$$

Here, the first equivalence follows from Lemma 7.9, and the second holds because $\mathit{Inf}(I, 0, s)$ is the only formula containing $I, 0, s$ as free variables. Thus the category $\mathcal{C}_{\mathcal{T}}/X$ demonstrates the required completeness property. □

We remark that, in fact, for a category of classes $\mathcal{C}$, it holds that $\mathcal{C} \models \mathrm{Inf}$ if and only if there exists an object X with global support [22] such that the slice category $\mathcal{C}/X$ satisfies the infinity axiom of Section 6.3. Thus an alternative

[22] An object X has *global support* if the unique map $X \longrightarrow \mathbf{1}$ is a regular epi.

approach to modelling the Infinity axiom of set theory would be to weaken the infinity axiom on categories of classes to merely require a small infinite object in some globally supported slice. This has the disadvantage of being less natural, and we shall not consider it further.

It is worth commenting that the completeness theorem for the theory IST of [11, Theorem 11] follows immediately from the combination of Propositions 7.10 and 7.11 above.

Proposition 7.12 *For any theory $\mathcal{T}$ and sentence ϕ, the following are equivalent.*

(1) $\mathrm{BIST} + \mathrm{Coll} + \mathcal{T} \vdash \phi$.
(2) For all categories of classes $\mathcal{C}$ satisfying Collection, $\mathcal{C} \models \mathcal{T}$ implies $\mathcal{C} \models \phi$.

PROOF. The proof of soundness is a simple verification that when $\mathcal{C}$ satisfies Collection, it holds that $\mathcal{C} \models \mathrm{Coll}$. The argument is essentially given by Joyal and Moerdijk [7, Proposition 5.1].

For the converse, suppose $\mathcal{T}$ contains all instances of Collection. One verifies easily that the covariant small powerobject functor in the syntactic category $\mathcal{C}_{\mathcal{T}}$ preserves regular epis. Thus $\mathcal{C}_{\mathcal{T}}$ satisfies Collection axiom. Hence completeness holds. $\square$

8 Categories of ideals

Proposition 6.3 showed that, in any category of classes, the full subcategory of small objects is a topos. In this section we prove that conversely every topos occurs as the category of small objects in a category of classes, in fact in a category of classes satisfying Collection. By Theorem 3.10, we can, without loss of generality, work with toposes endowed with a directed structural system of inclusions, i.e. a dssi as defined in Section 3. Given such a topos, we build a category of classes whose objects are ideals of objects in the topos under the inclusion order. The small objects turn out to be exactly the principal ideals, and thus essentially the same as the objects of original topos. Moreover, the resulting category of ideals automatically satisfies the Collection axiom.

We also give a variation on the ideal construction in the case of a topos endowed with a superdirected structural system of inclusions (i.e. an sdssi), using which we embed the topos in a category of classes satisfying both Collection and Separation axioms

Throughout this section, let $\mathcal{E}$ be a fixed topos with dssi $\mathcal{I}$. For convenience, we assume that $\mathcal{I}$ partially orders $\mathcal{E}$, see Proposition 3.3 and ensuing discussion. By an *ideal* in $\mathcal{E}$ we then mean an order ideal with respect to the inclusion ordering, i.e. a non-empty collection $\mathbf{C}$ of objects of $\mathcal{E}$, such that $A, B \in \mathbf{C}$ and $A' \hookrightarrow A$ implies $A \cup B \in \mathbf{C}$ and $A' \in \mathbf{C}$. A *morphism of ideals* consists of an order-preserving function,

$$\mathbf{f} : \mathbf{C} \to \mathbf{D}$$

together with a family of epimorphisms in $\mathcal{E}$,

$$\mathbf{f}_C : C \twoheadrightarrow \mathbf{f}(C) \qquad \text{for all } C \in \mathbf{C}$$

satisfying the naturality condition that, whenever $C' \hookrightarrow C$ in $\mathbf{C}$, the following diagram commutes in $\mathcal{E}$.

$$\begin{tikzcd}
C \arrow[r, two heads, "\mathbf{f}_C"] & \mathbf{f}(C) \\
C' \arrow[u, hook] \arrow[r, two heads, "\mathbf{f}_{C'}"'] & \mathbf{f}(C') \arrow[u, hook]
\end{tikzcd}$$

With the obvious identities and composition, these morphisms form the *category of ideals* in the topos $\mathcal{E}$with dssi $\mathcal{I}$, denoted $\mathbf{Idl}_{\mathcal{I}}(\mathcal{E})$. Usually, we omit explicit mention of $\mathcal{I}$, and we simply write $\mathbf{f}$ for the morphism $(\mathbf{f}, (\mathbf{f}_C)_{C \in \mathbf{C}})$.

Because epi-inclusion factorizations in $\mathcal{E}$ are unique, the values $\mathbf{f}(C)$ and $\mathbf{f}_C$ determine the values $\mathbf{f}(C')$ and $\mathbf{f}_{C'}$ for all $C' \hookrightarrow C$. Indeed, locally (i.e. on the segment below any fixed $C \in \mathbf{C}$) the mapping $\mathbf{f}$ is essentially the same as the direct image functor

$$(\mathbf{f}_C)_! : \mathrm{Sub}(C) \to \mathrm{Sub}(\mathbf{f}(C))$$

This implies the following.

Lemma 8.1 *Every morphism of ideals* $\mathbf{f} : \mathbf{C} \to \mathbf{D}$ *preserves unions,*

$$\mathbf{f}(A \cup B) = \mathbf{f}(A) \cup \mathbf{f}(B)$$

for all $A, B \in \mathbf{C}$*. Moreover,* $\mathbf{f}$ *is "locally surjective" in the sense that for every* $C \in \mathbf{C}$ *and* $D \hookrightarrow \mathbf{f}(C)$*, there is some* $C' \hookrightarrow C$ *with* $\mathbf{f}(C') = D$*.*

Next, observe that taking principal ideals determines a functor,

$$\downarrow : \mathcal{E} \to \mathbf{Idl}(\mathcal{E})$$

as follows: for any $f : A \to B$ in $\mathcal{E}$, we define:

$$\downarrow(f)(A' \hookrightarrow A) \;=\; f_!(A') \hookrightarrow B$$

where $f_!(A')$ is the image of A' under f, given by the unique epi-inclusion factorization, as indicated in:

$$\begin{tikzcd}
A \arrow[r, "f"] & B \\
A' \arrow[u, hook] \arrow[r, two heads, "f'"'] & f_!(A') \arrow[u, hook]
\end{tikzcd}$$

Moreover, we can then let $\downarrow(f)_{A'} = f'$, where f' is the indicated epi part of the factorization.

Proposition 8.2 *The principal ideal functor is full and faithful.*

PROOF. Given any morphism of ideals $\mathbf{f} : \downarrow(A) \longrightarrow \downarrow(B)$, consider the composite map:

$$T(\mathbf{f}) = i \circ \mathbf{f}_A : A \twoheadrightarrow \mathbf{f}(A) \hookrightarrow B$$

where $i : \mathbf{f}(A) \hookrightarrow B$ is the canonical inclusion. Then by naturality, the value of $\mathbf{f}$ on every $A' \hookrightarrow A$ is just $T(\mathbf{f})_!(A')$, and $\mathbf{f}_{A'} = \downarrow(T(\mathbf{f}))_{A'} : A' \longrightarrow T(\mathbf{f})_!(A')$. Thus $\mathbf{f} = \downarrow(T(\mathbf{f}))$. Since clearly $T(\downarrow(f)) = f$ for any morphism $f : A \longrightarrow B$, this proves the proposition. $\square$

Our main objective in this section is to prove that the category of ideals is a category of classes. The intuition is that each ideal represents a class in terms of its approximating subsets, and that class functions are similarly represented by their approximating functions on subsets. Accordingly, it is natural to define a map in $\mathbf{Idl}(\mathcal{E})$ to be small if it has an inverse image that maps approximating subsets of the codomain to approximating subsets of the domain.

Formally, we define a morphism of ideals $\mathbf{f} : \mathbf{A} \longrightarrow \mathbf{B}$ to be *small* if, for every $B \in \mathbf{B}$, the collection

$$\{A \in \mathbf{A} \mid f(A) \hookrightarrow B\}$$

has a greatest element under inclusion, and we write $\mathbf{f}^{-1}(B)$ for the largest A such that $f(A) \hookrightarrow B$. Equivalently, $\mathbf{f}$ is small if and only the function from $\mathbf{A}$ to $\mathbf{B}$ given by its mapping part has a right adjoint $\mathbf{f}^{-1}$ as a function between partial orders.

Theorem 8.3 *The category* $\mathbf{Idl}(\mathcal{E})$ *is a category of classes satisfying the Collection axiom. Moreover, the small objects in* $\mathbf{Idl}(\mathcal{E})$ *are exactly the principal ideals, and so the principal ideal embedding* $\downarrow\colon \mathcal{E} \hookrightarrow \mathbf{Idl}(\mathcal{E})$ *exibits* $\mathcal{E}$ *as the full subcategory of sets in* $\mathbf{Idl}(\mathcal{E})$.

The proof requires a lengthy verification of the axioms for class structure, which we present as a series of lemmas.

Lemma 8.4 *The category* $\mathbf{Idl}(\mathcal{E})$ *of ideals is a positive Heyting category.*

PROOF. The terminal ideal is $\downarrow(\mathbf{1})$, as is easily verified. The product of two ideals $\mathbf{A}$ and $\mathbf{B}$ is the collection:[23]

$$\mathbf{A} \times \mathbf{B} = \{C \hookrightarrow A \times B \mid A \in \mathbf{A}, B \in \mathbf{B}\}$$

which is an ideal because, if $C \hookrightarrow A \times B$ and $C' \hookrightarrow A' \times B'$, then we have:

$$C \cup C' \hookrightarrow (A \times B) \cup (A' \times B') \hookrightarrow (A \cup A') \times (B \cup B')$$

since products preserve inclusions. The projection $\pi_{\mathbf{1}} : \mathbf{A} \times \mathbf{B} \longrightarrow \mathbf{A}$ is defined by factoring as indicated in the following diagram:

$$\begin{tikzcd}
C \arrow[r, hook] \arrow[d, "\pi_{\mathbf{1}C}"'] & A \times B \arrow[d, "\pi_1"] \\
\pi_{\mathbf{1}}(C) \arrow[r, hook] & A
\end{tikzcd}$$

To see that this is well-defined, suppose also $C \hookrightarrow A' \times B'$ and consider $(A \cup A') \times (B \cup B')$. Then since products preserve inclusions, the image $\pi_{\mathbf{1}}(C)$ can equally well be computed with respect to $(A \cup A') \times (B \cup B')$, as indicated in the following:

$$\begin{tikzcd}
C \arrow[r, hook] \arrow[d, "\pi_{\mathbf{1}C}"'] & A \times B \arrow[r, hook] \arrow[d, "\pi_1"] & (A \cup A') \times (B \cup B') \arrow[d, "\pi_1"] \\
\pi_{\mathbf{1}}(C) \arrow[r, hook] & A \arrow[r, hook] & A \cup A'
\end{tikzcd}$$

Since the same is true for $\pi_{\mathbf{1}}(C)$ computed according to $C \hookrightarrow A' \times B'$, the two must agree. The second projection $\pi_{\mathbf{2}}$ is defined analogously. To see that this specification is indeed the product in $\mathbf{Idl}(\mathcal{E})$, given any ideal $\mathbf{C}$ and maps

[23] By this notation we of course mean the collection of objects C included in $A \times B$, not the collection of inclusion maps.

$\mathbf{f} : \mathbf{C} \longrightarrow \mathbf{A}$ and $\mathbf{g} : \mathbf{C} \longrightarrow \mathbf{B}$, let $\langle \mathbf{f}, \mathbf{g} \rangle : \mathbf{C} \longrightarrow \mathbf{A} \times \mathbf{B}$ take $C \in \mathbf{C}$ to the image in the diagram below.

$$\begin{tikzcd}
C \arrow[d, "\langle \mathbf{f}, \mathbf{g} \rangle_C"'] \arrow[dr, "\langle \mathbf{f}_C, \mathbf{g}_C \rangle"] & \\
\langle \mathbf{f}, \mathbf{g} \rangle(C) \arrow[r, hook] & f(C) \times g(C)
\end{tikzcd}$$

Then $\pi_{\mathbf{1}}(\langle \mathbf{f}, \mathbf{g} \rangle(C)) = \mathbf{f}(C)$ since $\mathbf{f}_C$ is an epi. We omit the verification of uniqueness.

For equalizers, given $\mathbf{f}, \mathbf{g} : \mathbf{A} \rightrightarrows \mathbf{B}$, their equalizer is the evident inclusion into $\mathbf{A}$ of:

$$\{A \in \mathbf{A} \mid \mathbf{f}(A) = \mathbf{g}(A), \mathbf{f}_A = \mathbf{g}_A\} \ .$$

This is clearly down-closed, and if $A, A' \in \mathbf{A}$ are both in it, then so is $A \cup A'$ since $\mathbf{f}$ and $\mathbf{g}$ both preserve unions.

Combining the foregoing two cases, we obtain the following description of pullbacks. Given $\mathbf{f} : \mathbf{A} \longrightarrow \mathbf{C}$ and $\mathbf{g} : \mathbf{B} \longrightarrow \mathbf{C}$, the pullback consists of (the evident projection morphisms on) the object:

$$\mathbf{A} \times_{\mathbf{C}} \mathbf{B} = \{D \hookrightarrow A \times B \mid A \in \mathbf{A}, B \in \mathbf{B}, \mathbf{f}(A) = \mathbf{g}(B), \mathbf{f}_A \circ d_1 = \mathbf{g}_B \circ d_2\}$$

where $d_1 : D \longrightarrow A$ and $d_2 : D \longrightarrow B$ are the two components of $D \hookrightarrow A \times B$.

In order to show that $\mathbf{Idl}(\mathcal{E})$ is a regular category, we characterize the regular epis as those morphisms $\mathbf{e} : \mathbf{A} \longrightarrow \mathbf{B}$ for which the mapping part $A \mapsto \mathbf{e}(A)$ is a surjective function from $\mathbf{A}$ to $\mathbf{B}$. Clearly such maps are indeed epis. Below, we show that, for any morphism $\mathbf{f} : \mathbf{A} \longrightarrow \mathbf{B}$, the subcollection,

$$\mathbf{f}(\mathbf{A}) = \{\mathbf{f}(A) \mid A \in \mathbf{A}\} \subseteq \mathbf{B} \tag{26}$$

is the coequalizer of the kernel pair of $\mathbf{f}$. It follows that the surjective mappings are indeed regular epis. Conversely, if $\mathbf{f}$ is a regular epi then it also coequalizes its own kernel pair, so we have $\mathbf{B} \cong \{\mathbf{f}(A) \mid A \in \mathbf{A}\} \subseteq \mathbf{B}$, whence $\mathbf{f}$ is surjective. It remains to prove that (26) indeed coequalizes the kernel pair of $\mathbf{f}$. According to the description of pullbacks above, the kernel-pair of $\mathbf{f}$ is:

$$\mathbf{K} = \{D \hookrightarrow A \times A' \mid A, A' \in \mathbf{A}, \mathbf{f}(A) = \mathbf{f}(A'), \mathbf{f}_A \circ d_1 = \mathbf{f}_{A'} \circ d_2\}$$

with the two evident projections $\pi_{\mathbf{1}}, \pi_{\mathbf{2}} : \mathbf{K} \longrightarrow \mathbf{A}$. But this ideal agrees with the following one:

$$\mathbf{K} = \{D \hookrightarrow A \times_{f(A)} A' \mid A, A' \in \mathbf{A}, \mathbf{f}(A) = \mathbf{f}(A')\}$$

where the indicated pullbacks are taken using the maps $f_A : A \twoheadrightarrow f(A)$ and $f_{A'} : A' \twoheadrightarrow f(A') = f(A)$. Given any morphism $\mathbf{g} : \mathbf{A} \longrightarrow \mathbf{C}$ with $\mathbf{g} \circ \pi_\mathbf{1} = \mathbf{g} \circ \pi_\mathbf{2}$, one can then define the required extension $\mathbf{g}' : \mathbf{f}(\mathbf{A}) \longrightarrow \mathbf{C}$ simply by setting:

$$\begin{aligned} \mathbf{g}'(f(A)) &= \mathbf{g}(A), \\ {\mathbf{g}'}_{f(A)} &= \mathbf{g}_A : A \twoheadrightarrow \mathbf{g}(A) \ . \end{aligned}$$

Having now characterized the regular epis in $\mathbf{Idl}(\mathcal{E})$ as the morphisms whose mappings are surjective, it is straightforward to verify that regular epis are stable under pullback. Thus $\mathbf{Idl}(\mathcal{E})$ is indeed a regular category.

Using the coproduct in $\mathcal{E}$ defined above Proposition 3.14, the coproduct of ideals $\mathbf{A}$ and $\mathbf{B}$ is defined by:

$$\mathbf{A} + \mathbf{B} \ = \ \{A + B \mid A \in \mathbf{A}, B \in \mathbf{B}\}$$

with the injection morphisms $A \mapsto A + \emptyset$ and $B \mapsto \emptyset + B$. It follows easily from Proposition 3.14 that this is indeed an ideal. The coproduct property of $\mathbf{A} + \mathbf{B}$ in $\mathbf{Idl}(\mathcal{E})$ is straightforward to verify.

Finally, the dual image along $\mathbf{f} : \mathbf{C} \longrightarrow \mathbf{D}$ of a subideal $\mathbf{A} \rightarrowtail \mathbf{C}$ is calculated as follows. Without loss of generality, we can assume that $\mathbf{A} \subseteq \mathbf{C}$. Then let:

$$\forall_\mathbf{f}(\mathbf{A}) = \{D \in \mathbf{D} \mid \text{for all } C \in \mathbf{C},\ \mathbf{f}(C) \hookrightarrow D \text{ implies } C \in \mathbf{A}\} \ .$$

To see that this works, note that the condition determining the elements D in $\forall_\mathbf{f}(\mathbf{A})$ is equivalent to $\forall_\mathbf{f}(\downarrow(D)) \subseteq \mathbf{A}$. □

Lemma 8.5 *The following characterizations hold in the category* $\mathbf{Idl}(\mathcal{E})$*:*

(1) The small objects *are exactly the principal ideals* $\downarrow E$ *for* $E \in \mathcal{E}$*.*
(2) Every morphism $\mathbf{f} : \downarrow E \longrightarrow \downarrow F$ *between small objects is of the form:* $\mathbf{f} = \downarrow f$ *for a unique* $f : E \longrightarrow F$ *in* $\mathcal{E}$*, and is therefore small.*
(3) The small subobjects $\mathbf{C}' \rightarrowtail \mathbf{C}$ *are exactly those isomorphic to subobjects of the form* $\downarrow C \subseteq \mathbf{C}$ *for some* $C \in \mathbf{C}$*.*
(4) A morphism $\mathbf{f} : \mathbf{A} \to \mathbf{B}$ *is small if, whenever* $S \rightarrowtail \mathbf{B}$ *is a small subobject, then* $\mathbf{f}^{-1}(S) \rightarrowtail \mathbf{A}$ *is also small.*

PROOF. Straighforward. □

Lemma 8.6 *The small maps so defined satisfy axioms (S1)–(S5).*

PROOF.

(S1) Small maps form a subcategory, since adjoints compose.

(S2) Suppose we have a pullback

$$\begin{tikzcd}
\mathbf{A} \times_{\mathbf{C}} \mathbf{B} \arrow[r, "\mathbf{q}"] \arrow[d, "\mathbf{p}"'] & \mathbf{B} \arrow[d, "\mathbf{g}"] \\
\mathbf{A} \arrow[r, "\mathbf{f}"'] & \mathbf{C}
\end{tikzcd}$$

with $\mathbf{g}$ small. To show $\mathbf{p}$ small, we need to find $\mathbf{p}^{-1}(A) \in \mathbf{A} \times_{\mathbf{C}} \mathbf{B}$ for each $A \in \mathbf{A}$. Consider the pullback diagram:

$$\begin{tikzcd}
T'' \arrow[r, two heads] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & T' \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & A \arrow[d, "\mathbf{f}(A)"] \\
T \arrow[r, two heads, "\mathbf{g}_T"'] & \mathbf{g}(T) \arrow[r, hook] & \mathbf{f}(A)
\end{tikzcd}$$

in which $T = \mathbf{g}^{-1}(\mathbf{f}(A))$. It follows that the subobject $T'' \hookrightarrow T' \times T$ is in the pullback $\mathbf{A} \times_{\mathbf{C}} \mathbf{B}$. Define: $\mathbf{p}^{-1}(A) = T''$. We omit the easy verification that this has the right properties.

(S3) Given $\Delta : \mathbf{C} \longrightarrow \mathbf{C} \times \mathbf{C}$ and $T \hookrightarrow A \times B$ in $\mathbf{C} \times \mathbf{C}$, we take the pullback:

$$\begin{tikzcd}
T' \arrow[rr] \arrow[d, hook] & & T \arrow[d, hook] \\
A \cap B \arrow[r, tail, "\Delta_{A \cap B}"'] & (A \cap B) \times (A \cap B) \arrow[r, hook] & A \times B
\end{tikzcd}$$

Define $\Delta^{-1}(T) = T'$. Again, we omit the straightforward verification.

(S4) Suppose the diagram below commutes

$$\begin{tikzcd}
\mathbf{A} \arrow[r, two heads, "\mathbf{e}"] \arrow[dr, "\mathbf{g}"'] & \mathbf{B} \arrow[d, "\mathbf{f}"] \\
& \mathbf{C}
\end{tikzcd}$$

where $\mathbf{g}$ is small and $\mathbf{e}$ is a regular epi. As in the proof of Lemma 8.4, the mapping part of $\mathbf{e}$ is a surjective function. To show that $\mathbf{f}$ is small, for

$C \in \mathbf{C}$, define $\mathbf{f}^{-1}(C) = \mathbf{e}(\mathbf{g}^{-1}(C)$. That this has the required properties follows from the smallness of $\mathbf{g}$ and the surjectivity property of $\mathbf{e}$.

(S5) Given small maps $\mathbf{f}$ and $\mathbf{g}$ as below, we must show that $[\mathbf{f}, \mathbf{g}]$ is small.

$$\begin{tikzcd} \mathbf{A} \arrow[r] \arrow[dr, "\mathbf{f}"'] & \mathbf{A} + \mathbf{B} \arrow[d, "{[\mathbf{f},\mathbf{g}]}"] & \mathbf{B} \arrow[l] \arrow[dl, "\mathbf{g}"] \\ & \mathbf{C} & \end{tikzcd}$$

For $C \in \mathbf{C}$, define: $[\mathbf{f}, \mathbf{g}]^{-1}(C) = \mathbf{f}^{-1}(C) + \mathbf{g}^{-1}(C)$. We omit the straightforward verification that this has the required properties.

□

Next we define small powerobjects in $\mathbf{Idl}(\mathcal{E})$. Given any ideal $\mathbf{C}$, define:

$$\mathcal{P}\mathbf{C} = \{S \hookrightarrow \mathcal{P}C \mid C \in \mathbf{C}\} \ .$$

This is indeed an ideal because, given $S \hookrightarrow \mathcal{P}C$, and $S' \hookrightarrow \mathcal{P}C'$, it holds that $S \cup S' \hookrightarrow \mathcal{P}C \cup \mathcal{P}C' \hookrightarrow \mathcal{P}(C \cup C')$.

Because $\mathcal{I}$ is a dssi on $\mathcal{E}$, for each object C of $\mathcal{E}$, the membership relation is given by an incluion $\in_C \hookrightarrow C \times \mathcal{P}C$. For an ideal $\mathbf{C}$, the membership relation on small objects is defined by the inclusion of

$$\in_{\mathbf{C}} = \{S \hookrightarrow \ \in_C \mid C \in \mathbf{C}\}$$

in the ideal $\mathbf{C} \times \mathcal{P}\mathbf{C}$. It is easily verified that $\in_{\mathbf{C}}$ is indeed an ideal.

Lemma 8.7 *The category* $\mathbf{Idl}(\mathcal{E})$ *satisfies axiom (P).*

PROOF. Since this is the trickiest case in the verification that $\mathbf{Idl}(\mathcal{E})$ is a category of classes, we give the proof in detail. Suppose we have a small relation $\mathbf{R} \subseteq \mathbf{C} \times \mathbf{A}$, with components $\mathbf{r_1} \colon \mathbf{R} \longrightarrow \mathbf{C}$ and $\mathbf{r_2} \colon \mathbf{R} \longrightarrow \mathbf{A}$. Because $\mathbf{R}$ is a small relation, $\mathbf{r_2}$ is a small map. We must show that there is a unique map $\chi_{\mathbf{R}} \colon \mathbf{A} \longrightarrow \mathcal{P}\mathbf{C}$ fitting into a pullback diagram:

$$\begin{tikzcd} \mathbf{R} \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & \in_{\mathbf{C}} \arrow[d] \\ \mathbf{C} \times \mathbf{A} \arrow[r, "\mathbf{1} \times \chi_{\mathbf{R}}"] & \mathbf{C} \times \mathcal{P}\mathbf{C} \end{tikzcd} \tag{27}$$

First we define $\chi_{\mathbf{R}} : \mathbf{A} \longrightarrow \mathcal{P}\mathbf{C}$. For $A \in \mathbf{A}$, take $\mathbf{r_2}^{-1}(A) \in \mathbf{R}$, which has the form $\mathbf{r_2}^{-1}(A) \rightarrowtail C \times A$ for some $C \in \mathbf{C}$. Using the characteristic property of powerobjects in $\mathcal{E}$ together with an image factorization, define $\chi_{\mathbf{R}}(A)$ and $(\chi_{\mathbf{R}})_A \colon A \twoheadrightarrow \chi_{\mathbf{R}}(A)$ to be the unique object and epimorphism fitting into a pullback diagram:

$$
\begin{tikzcd}
\mathbf{r_2}^{-1}(A) \arrow[rrr] \arrow[d, hook] \arrow[dr, phantom, "\lrcorner", very near start] & & & \in_C \arrow[d, hook] \\
C \times A \arrow[rr, two heads, "1_C \times (\chi_{\mathbf{R}})_A"] & & C \times \chi_{\mathbf{R}}(A) \arrow[r, hook] & C \times \mathcal{P}C
\end{tikzcd}
\tag{28}
$$

We show that this is independent of C. Accordingly, take another $C' \in \mathbf{C}$ such that $\mathbf{r_2}^{-1}(A) \hookrightarrow C' \times A$. Without loss of generality, we can assume $C \hookrightarrow C'$ (otherwise apply the following argument twice to show that the objects C, $C \cup C'$ and C' all determine the same $\chi_{\mathbf{R}}$). Then, composing pullback squares, we obtain:

$$
\begin{tikzcd}
\mathbf{r_2}^{-1}(A) \arrow[rrr] \arrow[d, hook] \arrow[dr, phantom, "\lrcorner", very near start] & & & \in_C \arrow[r, hook] \arrow[d, hook] \arrow[dr, phantom, "\lrcorner", very near start] & \in_{C'} \arrow[dd, hook] \\
C \times A \arrow[rr, two heads, "1_C \times (\chi_{\mathbf{R}})_A"] \arrow[d, hook] \arrow[dr, phantom, "\lrcorner", very near start] & & C \times \chi_{\mathbf{R}}(A) \arrow[r, hook] & C \times \mathcal{P}C \arrow[d, hook] & \\
C' \times A \arrow[rr, two heads, "1_{C'} \times (\chi_{\mathbf{R}})_A"] & & C' \times \chi_{\mathbf{R}}(A) \arrow[r, hook] & C' \times \mathcal{P}C \arrow[r, hook] & C' \times \mathcal{P}C'
\end{tikzcd}
$$

where the outer pullback shows that the same action of $\chi_{\mathbf{R}}$ on A is determined by the inclusion $\mathbf{r_2}^{-1}(A) \hookrightarrow C' \times A$.

We must verify that $\chi_{\mathbf{R}}$ indeed makes diagram (27) into a pullback and that it is the unique map doing so. This requires an analysis of the pullback property itself. For any map $\mathbf{g} \colon \mathbf{A} \longrightarrow \mathcal{P}\mathbf{C}$, the pullback

$$
\begin{tikzcd}
\mathbf{P_g} \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \in_{\mathbf{C}} \arrow[d, tail] \\
\mathbf{C} \times \mathbf{A} \arrow[r, "\mathbf{1} \times \mathbf{g}"] & \mathbf{C} \times \mathcal{P}\mathbf{C}
\end{tikzcd}
$$

can be defined by

$$\mathbf{P_g} \;=\; \{S \hookrightarrow C \times A \mid C \in \mathbf{C},\ A \in \mathbf{A},\ (\mathbf{1} \times \mathbf{g})(S) \;\in\; \in_{\mathbf{C}}\} \;,$$

with the map from $\mathbf{P_g}$ to $\mathbf{C} \times \mathbf{A}$ given by the evident inclusion. As in the proof of Lemma 8.4, the object $(\mathbf{1} \times \mathbf{g})(S)$ of $\mathcal{E}$ is given by the factorization:

$$\begin{tikzcd} S \arrow[r, two heads] \arrow[d, hook] & (\mathbf{1} \times \mathbf{g})(S) \arrow[d, hook] \\ C \times A \arrow[r, two heads, "1 \times \mathbf{g}_A"] & C \times \mathbf{g}(A) \end{tikzcd} \tag{29}$$

which is independent of C and A. For $S \hookrightarrow C \times A$ if $S \in \mathbf{P_g}$ then we have $(\mathbf{1} \times \mathbf{g})(S) \hookrightarrow \in_{C'}$ for some $C' \in \mathbf{C}$. Also, $\mathbf{g}_A \hookrightarrow \mathcal{P}C''$ for some $C'' \in \mathbf{C}$. By redefining C to be $C \cup C' \cup C''$, and applying the remarks before Proposition 3.12, we have that if $S \in \mathbf{P_g}$ then there exists $C \in \mathbf{C}$ such that the bottom composite below factors through the right-hand edge:

$$\begin{tikzcd} S \arrow[rr, dotted, "f"] \arrow[d, hook] & & \in_C \arrow[d, hook] \\ C \times A \arrow[r, two heads, "1 \times \mathbf{g}_A"] & C \times \mathbf{g}(A) \arrow[r, hook] & C \times \mathcal{P}C \end{tikzcd} \tag{30}$$

Conversely, by the uniqueness of the factorization (29) defining $(\mathbf{1} \times \mathbf{g})(S)$, any $S \hookrightarrow C \times A$, for which there exists an f making the diagram above commute, is contained in $\mathbf{P_g}$. Thus we have:

$$\mathbf{P_g} \;=\; \{S \hookrightarrow C \times A \mid \mathbf{g}(A) \hookrightarrow \mathcal{P}(C),\ \exists f.\ (30)\text{ commutes}\}\ . \tag{31}$$

At last, we show that diagram (27) is indeed a pullback with the defined $\chi_{\mathbf{R}}$. We must show that $\mathbf{R} = \mathbf{P}_{\chi_\mathbf{R}}$. Suppose that $S \hookrightarrow C \times A$ is in $\mathbf{R}$. Then $S \hookrightarrow \mathbf{r_2}^{-1}(A)$. Interpreting the pullback of (28) as an instance of diagram (30), we see that $\mathbf{r_2}^{-1}(A) \in \mathbf{P}_{\chi_\mathbf{R}}$, by (31). So $S \in \mathbf{P}_{\chi_\mathbf{R}}$, because $\mathbf{P}_{\chi_\mathbf{R}}$ is down-closed. Conversely, suppose that $S \hookrightarrow C \times A$ is in $\mathbf{P}_{\chi_\mathbf{R}}$. By (31), we have a commuting diagram (30) with $\mathbf{g} = \chi_\mathbf{R}$, whose span part is thus a cone for the pullback of (28). By the pullback property, $S \hookrightarrow \mathbf{r_2}^{-1}(A)$. Thus indeed $S \in \mathbf{R}$.

Finally, suppose $\mathbf{g}\colon \mathbf{A} \longrightarrow \mathcal{P}\mathbf{C}$ is such that $\mathbf{R} = \mathbf{P_g}$. We must show that $\mathbf{g} = \chi_\mathbf{R}$. For any $A \in \mathbf{A}$, we have $\mathbf{r_2}^{-1}(A) \in \mathbf{R} = \mathbf{P_g}$. Hence, by (31), there is a commuting diagram of the form (30), with $S = \mathbf{r_2}^{-1}(A)$. Let $T \hookrightarrow C \times A$ be the pullback of the right edge along the bottom. Again by (31), $T \in \mathbf{P_g} = \mathbf{R}$. By the pullback property of T, we have $\mathbf{r_2}^{-1}(A) \hookrightarrow T$. Conversely, $T \hookrightarrow \mathbf{r_2}^{-1}(A)$ follows from the defining property of $\mathbf{r_2}^{-1}(A)$, because $T \in \mathbf{R}$. Thus diagram (30) with $S = \mathbf{r_2}^{-1}(A)$ is itself a pullback, and hence identical to diagram (28). So indeed $\mathbf{g}(A) = \chi_\mathbf{R}(A)$ and $\mathbf{g}_A = (\chi_\mathbf{R})_A$. □

Lemma 8.8 *The category* $\mathbf{Idl}(\mathcal{E})$ *satisfies the Powerset axiom.*

PROOF. The subset relation $\subseteq_{\mathbf{C}} \rightarrowtail \mathcal{P}\mathbf{C} \times \mathcal{P}\mathbf{C}$ is given by the subideal:

$$\subseteq_{\mathbf{C}} = \{S \mid S \hookrightarrow \subseteq_C \hookrightarrow \mathcal{P}C \times \mathcal{P}C,\ C \in \mathbf{C}\}$$

with the evident inclusion. To see that the second projection

$$\mathbf{q} : \subseteq_{\mathbf{C}} \rightarrowtail \mathcal{P}\mathbf{C} \times \mathcal{P}\mathbf{C} \xrightarrow{\pi_2} \mathcal{P}\mathbf{C}$$

is small, take any $S \hookrightarrow \mathcal{P}C$ in $\mathcal{P}\mathbf{C}$, and form the pullback:

$$\begin{tikzcd}
S' \arrow[rr] \arrow[d, hook] & & S \arrow[d, hook] \\
\subseteq_C \arrow[r, hook] & \mathcal{P}C \times \mathcal{P}C \arrow[r, "\pi_2"] & \mathcal{P}C
\end{tikzcd}$$

Define $\mathbf{q}^{-1}(S) = S'$. We omit the verification that this has the required properties. $\square$

Lemma 8.9 *The category* $\mathbf{Idl}(\mathcal{E})$ *satisfies the Collection axiom.*

PROOF. We verify that the covariant small powerobject functor preserves regular epis (property 2 of Proposition 6.6). Accordingly, suppose $\mathbf{e} \colon \mathbf{A} \longrightarrow \mathbf{B}$ is regular epi. As in the proof of Lemma 8.4, this means that the mapping part of $\mathbf{e}$ is a surjective function. We must show that the same property holds of $\mathbf{e}_! \colon \mathcal{P}\mathbf{A} \longrightarrow \mathcal{P}\mathbf{B}$. The map $\mathbf{e}_!$ has the following explicit description. For $A \in \mathbf{A}$ and $S \hookrightarrow \mathcal{P}A$, the object $(\mathbf{e}_!)(S)$ and map $(\mathbf{e}_!)_S \colon S \twoheadrightarrow (\mathbf{e}_!)(S)$ are given by the image factorization:

$$\begin{tikzcd}
S \arrow[r, two heads, "(\mathbf{e}_!)_S"] \arrow[d, hook] & (\mathbf{e}_!)(S) \arrow[d, hook] \\
\mathcal{P}A \arrow[r, two heads, "(\mathbf{e}_A)_!"] & \mathcal{P}(\mathbf{e}(A))
\end{tikzcd}$$

To show surjectivity of the mapping part, suppose $T \hookrightarrow \mathcal{P}B$ for some $B \in \mathbf{B}$. We must show that there exists $S \in \mathcal{P}\mathbf{A}$ with $(\mathbf{e}_!)(S) = T$.

By the surjectivity of $\mathbf{e}$, there exists $A \in \mathbf{A}$ with $\mathbf{e}(A) = B$. But then $(\mathbf{e}_A)_! \colon \mathcal{P}A \longrightarrow \mathcal{P}B$ is an epi, since the covariant powerobject functor in a topos preserves epis. Defining S by pullback:

$$\begin{tikzcd} S \arrow[r, two heads] \arrow[d, hook] \arrow[dr, phantom, "\lrcorner", very near start] & T \arrow[d, hook] \\ \mathcal{P}A \arrow[r, two heads, "(\mathbf{e}_A)_!"] & \mathcal{P}B \end{tikzcd}$$

we see that the factorization defining $(\mathbf{e}_!)(S)$ yields $(\mathbf{e}_!)(S) = T$, as required. □

Lemma 8.10 *The category* $\mathbf{Idl}(\mathcal{E})$ *has a universal object*

PROOF. The total ideal $\mathbf{U} = \{E \mid E \in \mathcal{E}\}$ is a universal object in $\mathbf{Idl}(\mathcal{E})$ because $\mathbf{C} \subseteq \mathbf{U}$ for every ideal $\mathbf{C}$. □

In combination, Lemma 8.4–8.10 prove Theorem 8.3.

Corollary 8.11 *The category* $\mathbf{Idl}(\mathcal{E})$ *satisfies the infinity axiom if and only if* $\mathcal{E}$ *has a natural numbers object.*

PROOF. Immediate from Theorem 8.3 and Proposition 6.5. □

Thus far, in Part III, we have avoided discussing meta-theoretic issues altogether. This was justifiable in Sections 5–7, where the development was entirely elementary and easily formalizable in any reasonable meta-theory, including BIST. In this section, the construction of categories of ideals is less elementary, and meta-theoretic issues do arise exactly parallel to those discussed in Section 4. Again, we take BIST itself as our primary meta theory. In the case that $\mathcal{E}$ is a small topos, there is no problem in doing so, as, by the Powerset axiom, the category of ideals is again small (taking ideals to be sets of objects). In the case that $\mathcal{E}$ is only locally small, an ideal has to be taken to be a subclass of the class of objects. In this case, the category $\mathbf{Idl}(\mathcal{E})$ is not itself a locally small category. The collection of morphisms between two objects $\mathbf{A}$ and $\mathbf{B}$ may form a class, and the collection of all objects need not even form a class (just as there is no class of all classes). In this case, it is best to look at $\mathbf{Idl}(\mathcal{E})$ as a "meta-category" in the following sense: its objects and hom-classes are individually definable as classes, but we never need to gather them together in a single collection. Instead, the results above should be understood schematically as applying to the relevant objects on an individual basis.

We end this section with a variation on the construction of $\mathbf{Idl}(\mathcal{E})$, which requires the collection $\mathcal{I}$ of inclusions on $\mathcal{E}$ to be a *superdirected* structural system of inclusions (sdssi). Under these circumstances, a *superideal* is a (necessarily nonempty) down-closed collection $\mathbf{A}$ of objects in $\mathcal{I}$ such that every subset of $\mathbf{A}$ has an upper bound in $\mathbf{A}$. We write $\mathbf{sIdl}(\mathcal{E})$ for the full subcategory of $\mathbf{Idl}(\mathcal{E})$ consisting of superideals. And we define a map in $\mathbf{sIdl}(\mathcal{E})$ to be *small* if it is small in $\mathbf{Idl}(\mathcal{E})$.

Theorem 8.12 (IST+ Coll) *If $\mathcal{I}$ is an sdssi on a locally small category $\mathcal{E}$, then the category $\mathbf{sIdl}(\mathcal{E})$ of superideals is a category of classes satisfying both Collection and Separation axioms. Once again, the small objects in $\mathbf{sIdl}(\mathcal{E})$ are exactly the principal ideals, and so the principal ideal embedding $\downarrow\colon \mathcal{E} \hookrightarrow \mathbf{sIdl}(\mathcal{E})$ exibits $\mathcal{E}$ as the full subcategory of sets in $\mathbf{sIdl}(\mathcal{E})$. Moreover, the inclusion functor $\mathbf{sIdl}(\mathcal{E}) \hookrightarrow \mathbf{Idl}(\mathcal{E})$ is logical.*

(Because toposes with sdssi's are not small, in the proof below, the meta-theory IST+ Coll is being used in the schematic sense discussed above.)

PROOF. Suppose that $\mathcal{I}$ is an sdssi on $\mathcal{E}$. To show that the category of superideals is a category of classes satisfying Separation, we use the economical axiomatization of such categories from [11]. For this, it suffices to verify axioms (C1),(S1), (S2) and (P) together with the Powerset and Separation axioms.

For all but the Separation axiom, we verify that the structure already defined on the category of ideals $\mathbf{Idl}(\mathcal{E})$, preserves the property of being a superideal. The most interesting case is to show that $\mathbf{sIdl}(\mathcal{E})$ is a regular category, for which we establish that superideals are closed under images in the category of ideals. Accordingly, suppose that $\mathbf{A}$ is a superideal and $\mathbf{e}\colon \mathbf{A} \twoheadrightarrow \mathbf{B}$ is a regular epi in the category of ideals. We show that the ideal $\mathbf{B}$ is a superideal. Suppose then that $\mathcal{B}$ is a subset of $\mathbf{B}$. As $\mathbf{e}$ is a regular epi, for each $B \in \mathcal{B}$, there exists $A \in \mathbf{A}$ with $\mathbf{e}(A) = B$. By Collection in the meta-theory, there exists a set $\mathcal{A} \subseteq \mathbf{A}$ such that, for all $B \in \mathcal{B}$, there exists $A \in \mathcal{A}$ with $\mathbf{e}(A) = B$. As $\mathbf{A}$ is a superideal, there exists an upper bound $U \in \mathbf{A}$ for $\mathcal{A}$. Then $\mathbf{e}(U)$ is the required upper bound for $\mathcal{B}$ in $\mathbf{B}$.

To show the Separation axiom, suppose that $\mathbf{m}\colon \mathbf{A} \rightarrowtail \mathbf{B}$ is a mono in $\mathbf{sIdl}(\mathcal{E})$. Without loss of generality $\mathbf{A} \subseteq \mathbf{B}$. To show that the mono is small, take any $B \in \mathbf{B}$. Consider the collection $\mathcal{A} = \{A \in \mathbf{A} \mid A \hookrightarrow B\}$. Because $\mathcal{E}$ is locally small, the collection $\{A \in \mathbf{B} \mid A \hookrightarrow B\}$ is a set, so, by full Separation in the meta-theory, $\mathcal{A}$ is a set. As $\mathbf{A}$ is a superideal, $\mathcal{A}$ has an upper bound $U \in \mathbf{A}$. But then $U \cap B \in \mathbf{A}$ is the required object $\mathbf{m}^{-1}(B)$ showing that $\mathbf{m}$ is small.

Finally, the inclusion functor is logical because the structure on $\mathbf{sIdl}(\mathcal{E})$ is all inherited directly from $\mathbf{Idl}(\mathcal{E})$. $\square$

The reader may have noticed that the above proof shares similarities with the proof of Proposition 4.9. In common with that proof, we mention that it is not straightforward to verify directly that superideals are closed under dual images in $\mathbf{Idl}(\mathcal{E})$. Thus the economical axiomatization of [11] is helpful in enabling the simple proof above.

9 Ideal models of set theory

The ideal construction of the previous section shows that every topos with dssi embeds in a category of classes satisfying the Collection axiom. Using the interpretation of set theory in a category of classes from Section 7, one thereby obtains a model of the set theory BIST$^-$+ Coll. On the other hand, in Section 3, we gave a direct interpretation of the language of set theory in a topos with dssi, using the forcing interpretation defined over the inclusions, which again modelled BIST$^-$+ Coll. In this short section, we show that these two interpretations of set theory coincide.

Theorem 9.1 *If $\mathcal{E}$ is an elementary topos with dssi $\mathcal{I}$ then the following are equivalent for a sentence ϕ in the first-order language of Section 2.*

(1) $\mathbf{Idl}(\mathcal{E}) \models \phi$*, using the class category interpretation of Section 7.*
(2) $(\mathcal{E}, \mathcal{I}) \models \phi$*, using the forcing semantics of Section 3.*

The theorem is proved by induction on the structure of ϕ, and hence we need to establish a generalized equivalence for formulas with free variables.

Suppose we have such an open formula $\phi(x_1, \ldots, x_n)$. Then the interpretation from Section 7 of ϕ in $\mathbf{Idl}(\mathcal{E})$ defines:

$$[\![x_1, \ldots, x_k \mid \phi]\!] \rightarrowtail \mathbf{U}^k \ ,$$

where $\mathbf{U}$ is the universal ideal of Lemma 8.10. However, the object $\mathbf{U}^k$ in $\mathbf{Idl}(\mathcal{E})$ is given by the ideal

$$\mathbf{U}^k \ = \ \{S \hookrightarrow A_1 \times \cdots \times A_k \mid A_1, \ldots, A_k \text{ objects of } \mathcal{E}\} \ ,$$

and subobjects of $\mathbf{U}^k$ are simply subideals of this (i.e. down-closed subcollections closed under binary union). Henceforth in this section, we write $[\![x_1, \ldots, x_k \mid \phi]\!]$ to mean such a subideal. In the case that ϕ is a sentence, then $[\![\phi]\!]$ is a subideal of $\downarrow\mathbf{1}$. By definition, $[\![\phi]\!] = \downarrow\mathbf{1}$ if and only if $\mathbf{Idl}(\mathcal{E}) \models \phi$.

We next observe that the forcing semantics of Section 3 also associates a subideal of $\mathbf{U}^k$ to $\phi(x_1, \ldots, x_n)$, namely:

$$[\![x_1, \ldots, x_k \mid \phi]\!]' \ = \ \{S \hookrightarrow A_1 \times \cdots \times A_k \mid S \Vdash_\rho \phi\} \ ,$$

where ρ_{x_i} is the projection $S \hookrightarrow A_1 \times \cdots \times A_k \longrightarrow A_i$. As above, when ϕ is a sentence, $[\![\phi]\!]'$ is a subideal of $\downarrow\mathbf{1}$. By the remarks above Theorem 4.6, it holds that $[\![\phi]\!]' = \downarrow\mathbf{1}$ if and only if $(\mathcal{E}, \mathcal{I}) \models \phi$.

By the discussion above on the two interpretations $[\![\phi]\!]$ and $[\![\phi]\!]'$ of a sentence ϕ, Theorem 9.1 is an immediate consequence of the lemma below.

Lemma 9.2 *If $\mathcal{E}$ is an elementary topos with dssi $\mathcal{I}$ then, for any formula $\phi(x_1, \ldots, x_k)$, it holds that $[\![x_1, \ldots, x_k \mid \phi]\!] = [\![x_1, \ldots, x_k \mid \phi]\!]'$.*

PROOF. The proof is a straightforward induction on the structure of ϕ. We present one illustrative case.

Assuming $[\![\vec{x}, y \mid \phi]\!] = [\![\vec{x}, y \mid \phi]\!]'$, we show that:

$$[\![\vec{x} \mid \exists y.\, \phi]\!] = [\![\vec{x} \mid \exists y.\, \phi]\!]' \; . \tag{32}$$

By the semantics of existential quantification in the internal logic of $\mathbf{Idl}(\mathcal{E})$, the ideal $[\![\vec{x} \mid \exists y.\, \phi]\!]$ is given by the following image factorization in $\mathbf{Idl}(\mathcal{E})$.

$$\begin{tikzcd}
[\![\vec{x}, y \mid \phi]\!] \arrow[r, two heads, "\mathbf{e}"] \arrow[d, tail] & {[\![\vec{x} \mid \exists y.\, \phi]\!]} \arrow[d, tail] \\
\mathbf{U}^k \times \mathbf{U} \arrow[r, "\pi_1"] & \mathbf{U}^k
\end{tikzcd} \tag{33}$$

To show the $\subseteq$ inclusion of (32), suppose that $T \hookrightarrow X_1 \times \cdots \times X_k$ is contained in $[\![\vec{x} \mid \exists y.\, \phi]\!]$. Applying the characterization of regular epis in $\mathbf{Idl}(\mathcal{E})$, as morphisms whose mapping part is surjective, to the above factorization, there exist Y and $S \hookrightarrow X_1 \times \cdots \times X_k \times Y$ such that $S \in [\![\vec{x}, y \mid \phi]\!] = [\![\vec{x}, y \mid \phi]\!]'$ and $\mathbf{e}(S) = T$. Hence the epi $\mathbf{e}_S \colon S \twoheadrightarrow T$ together with the projection $S \hookrightarrow X_1 \times \cdots \times X_k \times Y \longrightarrow Y$ are the data required by the forcing semantics for showing that $T \in [\![\vec{x} \mid \exists y.\, \phi]\!]'$.

For the converse inclusion, suppose that $T \hookrightarrow X_1 \times \cdots \times X_k$ is contained in $[\![\vec{x} \mid \exists y.\, \phi]\!]'$. Then, by the forcing semantics, there exists maps $U \overset{t}{\twoheadrightarrow} T$ and $U \xrightarrow{a} Y$ in $\mathcal{E}$ such that $U \Vdash_{\rho \circ t[a/y]} \phi$ (where ρ is built from the evident projections). Define S by taking the image factorization of the unique map $U \longrightarrow X_1 \times \cdots \times X_k \times Y$ making the solid arrows in the diagram below commute.

$$
\begin{tikzcd}
U \arrow[rr, "a"] \arrow[d, two heads, "t"'] \arrow[dr, two heads] & & Y \\
T \arrow[d, hook] & S \arrow[l, dotted, two heads] \arrow[dr, hook] & \\
X_1 \times \cdots \times X_k & & X_1 \times \cdots \times X_k \times Y \arrow[ll, "\pi_{X_1 \times \cdots \times X_k}"'] \arrow[uu, "\pi_Y"']
\end{tikzcd}
\tag{34}
$$

Because the left edge is the image factorization of the bottom-left triangle, there exists an epi $S \twoheadrightarrow T$ as indicated. Since $U \Vdash_{\rho \circ t[a/y]} \phi$, it follows from Lemma 4.2 that $S \Vdash_{\rho'} \phi$, where ρ' again consists of the evident projections away from $S \hookrightarrow X_1 \times \cdots \times X_k \times Y$. Thus $S \in [\![\vec{x}, y \mid \phi]\!]' = [\![\vec{x}, y \mid \phi]\!]$. However, it follows from the bottom quadrilateral of (34) that the epi $S \twoheadrightarrow T$ is a component of the bottom-left composite of (33). So, by the definition of $[\![\vec{x} \mid \exists y.\, \phi]\!]$ as the factorization of this composite, indeed $T \in [\![\vec{x} \mid \exists y.\, \phi]\!]$. $\square$

We now have the promised second proof of the soundness direction of Theorem 4.6. Indeed the result is a consequence of Theorem 9.1 together with the soundness direction of Theorem 7.1. Thus the direct proof of the soundness of the forcing semantics in Section 4 has been rendered redundant.

At this point, we return to the issue of the conservativity of the forcing semantics over the internal logic of $\mathcal{E}$, discussed around Proposition 4.10. Using the tools we have now established, there is a much neater formulation of this. Since $\mathcal{E}$ is a topos, it can itself be considered as a category with basic class structure, and, as already discussed at the end of Section 6.1, the embedding $\mathcal{E} \hookrightarrow \mathbf{Idl}(\mathcal{E})$ is logical and reflects isomorphisms. This expresses in an elegant way that the first-order logic of quantification over classes of the internal logic of $\mathbf{Idl}(\mathcal{E})$ is conservative over the internal logic of $\mathcal{E}$. By Theorem 9.1. the forcing semantics of BIST$^-$ is equivalent to the semantics determined by class quantification over the universal ideal in the internal logic of $\mathbf{Idl}(\mathcal{E})$. Hence the forcing semantics is in general conservative over the internal logic of $\mathcal{E}$. In particular, when $\mathcal{E}$ has a natural numbers object N, the properties of first-order arithmetic valid in $\mathcal{E}$ are the same as those valid for $\downarrow N$ in $\mathbf{Idl}(\mathcal{E})$ (it is irrelevant that $\downarrow N$ is not a natural numbers object in $\mathbf{Idl}(\mathcal{E})$). Proposition 4.10 follows.

We end the section by observing that, in the case of a topos with superdirected system of inclusions, the forcing semantics of set theory also coincides with the interpretation in the category of superideals.

Theorem 9.3 *If $\mathcal{E}$ is an elementary topos with sdssi $\mathcal{I}$ then the following are equivalent for a sentence ϕ in the first-order language of Section 2.*

(1) $\mathbf{sIdl}(\mathcal{E}) \models \phi$*, using the class category interpretation of Section 7.*
(2) $(\mathcal{E}, \mathcal{I}) \models \phi$*, using the forcing semantics of Section 3.*

PROOF. As the inclusion $\mathbf{sIdl}(\mathcal{E}) \hookrightarrow \mathbf{Idl}(\mathcal{E})$ is logical (Theorem 8.12), the interpretation of the language of set theory in $\mathbf{sIdl}(\mathcal{E})$ coincides with the interpretation in $\mathbf{Idl}(\mathcal{E})$. The forcing semantics is anyway unchanged for a superdirected system of inclusions. Thus the result is an immediate consequence of Theorem 9.1. $\square$

Finally, we remark that since the interpretation of set theory in $\mathbf{Idl}(\mathcal{E})$ and $\mathbf{sIdl}(\mathcal{E})$ coincide, when $\mathcal{E}$ carries an sdssi, it holds that $\mathbf{Idl}(\mathcal{E}) \models$ Sep even though the Separation axiom for categories of classes does not hold in $\mathbf{Idl}(\mathcal{E})$. This justifies a comment made after Proposition 7.10 above.

10 Ideal completeness

We have seen that every topos $\mathcal{E}$ with dssi gives rise to a category of ideals $\mathbf{Idl}(\mathcal{E})$ in which the universal object models BIST$^-$+ Coll. The aim of this section is to strengthen the completeness direction of Theorem 7.1 by showing that completeness still holds if the quantification over categories of classes is restricted to categories of ideals. In particular, BIST$^-$+ Coll is a complete axiomatization of the sentences valid in all categories of ideals.

Theorem 10.1 (Ideal completeness) *For any theory $\mathcal{T}$ and sentence ϕ, if:*

for every topos $\mathcal{E}$ with dssi, it holds that $\mathbf{Idl}(\mathcal{E}) \models \mathcal{T}$ *implies* $\mathbf{Idl}(\mathcal{E}) \models \phi$*,*

then BIST$^-$+ Coll+ $\mathcal{T} \vdash \phi$.

As an immediate consequence of the theorem, we finally obtain the missing implication of Theorem 4.6, the completeness of the forcing semantics.

Corollary 10.2 *The completeness implication of Theorem 4.6 holds.*

PROOF. Immediate from Theorems 9.1 and 10.1. $\square$

The rest of this section is devoted to the proof of Theorem 10.1. The strategy is to derive Theorem 10.1 from the completeness direction of Theorem 7.1, by showing that, for every category of classes $\mathcal{C}$ satisfying Collection, it is possible to "conservatively" embed $\mathcal{C}$ in a category of ideals. Here, the conservativity of the embedding means that the category of ideals does not validate any properties in the internal logic of $\mathcal{C}$ that are not already valid in $\mathcal{C}$. Clearly this is enough to obtain completeness.

In order to construct the embedding, we start with a small category of classes $\mathcal{C}$ satisfying Collection, and we work in ZFC as the meta-theory. The construction of the embedding of $\mathcal{C}$ into a category of ideals proceeds in two steps.

Step 1: Any small category of classes $\mathcal{C}$ satisfying the axiom of Collection has a conservative logical functor,

$$\mathcal{C} \to \mathcal{C}^*$$

into another one $\mathcal{C}^*$ that is "saturated" with small objects.

Step 2: The saturated class category $\mathcal{C}^*$ has a conservative logical functor,

$$\mathcal{C}^* \to \mathbf{Idl}(\mathcal{E})$$

into the category of ideals in a topos $\mathcal{E}$.

The topos $\mathcal{E}$ in step 2 is equivalent to the subcategory of small objects in $\mathcal{C}^*$. Step 1 is required to ensure there are enough such objects.

Before proceeding with the two steps, we prepare some necessary machinery from the general model theory of categories of classes. First we define the required notion of conservative functor. As is standard, we say that a subobject $X' \rightarrowtail X$ is *proper* if its representing mono is not an isomorphism.

Definition 10.3 (Conservative functor) A functor $F : \mathcal{C} \to \mathcal{D}$, between categories of classes, is called *conservative* if it is both logical and preserves proper subobjects.

Often we shall use the conservativity of a functor $F\colon \mathcal{C} \to \mathcal{C}'$ as follows. Given a property ϕ expressed in the internal logic of $\mathcal{C}$, one obtains a translation $F\phi$ in the internal logic of $\mathcal{C}'$. When F is logical, it holds that $C \models \phi$ implies $\mathcal{C}' \models F\phi$. When F is conservative, it also holds that $C \not\models \phi$ implies $\mathcal{C}' \not\models F\phi$.

By applying the above argument to suitable internal formulas, one sees that conservative functors are faithful and reflect monos, epis and isos. Clearly, a logical functor is conservative if and only if it reflects isos. We remark that we do not know if faithful logical functors between categories of classes are automatically conservative.

Recall (see footnote 22) that an object X is said to have *global support* if the unique map $X \longrightarrow \mathbf{1}$ is a regular epi.

Lemma 10.4 *If $\mathcal{C}$ is a category of classes and X has global support. then the reindexing functor $X^*\colon \mathcal{C} \to \mathcal{C}/X$ is conservative.*

PROOF. X^* is logical by Proposition 5.17. That it preserves proper subobjects is an easy consequence of X having global support. $\Box$

Because categories of classes have universal objects, we shall be interested in functors preserving these, i.e in *cofinal* functors in the sense of Section 6.5. As in the discussion there, all functors $X^*\colon \mathcal{C} \to \mathcal{C}/X$ are indeed cofinal. The next lemma allows us to build filtered colimits of cofinal conservative functors between class categories.

Lemma 10.5 (ZFC) *If $(\mathcal{C}_i)_{i\in I}$ is a filtered diagram of cofinal logical functors between small categories of classes, then the colimit category,*

$$\varinjlim_i \mathcal{C}_i$$

is also a category of classes, and is a colimit in the large category of categories of classes and (cofinal) logical functors. If each $\mathcal{C}_i$ has Collection, then so does $\varinjlim_i \mathcal{C}_i$. Moreover, if each functor $\mathcal{C}_i \to \mathcal{C}_j$ is conservative, then so is each canonical inclusion $\mathcal{C}_i \to \varinjlim_i \mathcal{C}_i$.

PROOF. A routine verification. Note that the axiom of choice is required to define the class category structure on $\varinjlim_i \mathcal{C}_i$. Also, the cofinality of the functors is required to define a universal object in $\varinjlim_i \mathcal{C}_i$. $\Box$

10.1 Saturating a category of classes

Definition 10.6 (Saturated category) A category of classes $\mathcal{C}$ is said to be *saturated* if it satisfies the following conditions:

> *Small covers:* given any regular epi $C \twoheadrightarrow A$ with A small, there is a small subobject $B \rightarrowtail C$ such that the restriction $B \rightarrowtail C \twoheadrightarrow A$ is still a regular epi.

$$\begin{tikzcd} B \arrow[r, tail, dotted] \arrow[dr, two heads] & C \arrow[d, two heads] \\ & A \end{tikzcd}$$

Small generators: given any subobject $B \rightarrowtail C$, if every small subobject $A \rightarrowtail C$ factors through B, then $B \cong C$.

Recall that an object X of a regular category is said to be *(regular) projective* if, for every regular epi $e\colon Y \twoheadrightarrow Z$ and map $z\colon X \longrightarrow Z$, there exists a map $y\colon X \longrightarrow Y$ such that $z = e \circ y$. A straightforward pullback argument shows that X is projective if and only if every regular epimorphism $e\colon Y \twoheadrightarrow X$ splits (i.e. there exists $s\colon X \longrightarrow Y$ with $e \circ s = 1_X$).

We require a strengthened notion of projectivity.

Definition 10.7 (Strong projectivity) An object X in a category of classes is said to be *strongly projective* if, for every regular epi $e\colon Y \twoheadrightarrow X$ and proper subobject $Y' \rightarrowtail Y$, there exists a splitting $X \longrightarrow Y$ of e that does not factor through $Y' \rightarrowtail Y$.

Classically, strong projectivity implies ordinary projectivity because, for any regular epi $e\colon Y \twoheadrightarrow X$, either $\mathbf{0} \rightarrowtail Y$ is a proper subobject or $\mathbf{0} \cong Y \cong X$. In the first case e splits by strong projectivity, in the second e is an iso.

Lemma 10.8 (ZFC) *Every small category of classes $\mathcal{C}$ has a cofinal conservative functor, $\mathcal{C} \to \mathcal{C}^*$ into a category of classes $\mathcal{C}^*$ in which the terminal object $\mathbf{1}$ is strongly projective. Also, if $\mathcal{C}$ satisfies Collection, then so does $\mathcal{C}^*$.*

PROOF. First we observe the following fact. If $m\colon C \rightarrowtail X$ is a proper subobject in a category of classes $\mathcal{C}$, then there exists a map $x\colon \mathbf{1} \longrightarrow X^*X$ in $\mathcal{C}/X$ (where X^*X is the reindexing of X along $X^*\colon \mathcal{C} \to \mathcal{C}/X$) such that the property $\exists c\colon X^*C.\, x = (X^*m)(c)$ does not hold in the internal logic of $\mathcal{C}/X$. To see this, define $x\colon \mathbf{1} \longrightarrow X^*X$ to be the "generic point" of X^*X in $\mathcal{C}/X$, given by diagonal $\Delta\colon 1_x \longrightarrow \pi_2$ (recall that $X^*X = \pi_2\colon X \times X \longrightarrow X$). If the above property were valid in the internal logic of $\mathcal{C}/X$ then, by the genericity of x, we would have $\mathcal{C} \models \forall x\colon X.\, \exists c\colon C.\, x = m(c)$, which contradicts that m is a proper subobject. So indeed it holds that $\mathcal{C}/X \not\models \exists c\colon X^*C.\, x = (X^*m)(c)$.

Now we turn to the construction of $\mathcal{C}^*$ required by the lemma. This is done in two stages.

First, using the axiom of choice, let $(X_\alpha)_{\alpha<\kappa}$ be a well-ordering of the objects of $\mathcal{C}$ that have global support, indexed by ordinals $< \kappa$. We construct a sequence $\{C_\alpha\}_{\alpha\leq\kappa}$ of categories of classes together with conservative functors

$\{J_{\alpha,\beta}\colon \mathcal{C}_\alpha \to \mathcal{C}_\beta\}_{\alpha\leq\beta<\kappa}$, forming a filtered diagram, as follows.

$$\begin{aligned}
\mathcal{C}_0 &= \mathcal{C} & J_{0,0} &= \mathrm{Id} & \\
\mathcal{C}_{\beta+1} &= \mathcal{C}_\beta/J_{0,\beta}X_\beta & J_{\alpha,\beta+1} &= (J_{0,\beta}X_\beta)^* \circ J_{\alpha\beta} & \\
\mathcal{C}_\lambda &= \varinjlim_{\alpha<\lambda} \mathcal{C}_\alpha & J_{\alpha,\lambda} &= \text{colimit injection} & (\lambda \text{ a limit ordinal})
\end{aligned}$$

Here, the functors $J_{\alpha,\beta+1}$ are conservative by Lemma 10.4. Similarly, the functors $J_{\alpha,\lambda}$ are conservative by Lemma 10.5. (The diagrams $\{C_\alpha\}_{\alpha<\lambda}$ are always filtered by construction.) Moreover, as the Collection axiom is preserved by slicing and by filtered colimits, if $\mathcal{C}$ satisfies Collection then so does every $\mathcal{C}_\alpha$.

Define $\mathcal{C}_* = \mathcal{C}_\kappa$ and $J = J_{0,\kappa}\colon \mathcal{C} \to \mathcal{C}_*$. We claim that, for any X_α with global support and proper subobject $m\colon C \rightarrowtail X_\alpha$ in $\mathcal{C}$, there is an arrow $d : \mathbf{1} \to JX_\alpha$ in $\mathcal{C}_*$ such that $\mathcal{C}_* \not\models \exists c\colon JC.\, d = (Jm)(c)$. Indeed, since $J_{0,\alpha}$ is conservative, $J_{0,\alpha}m$ is a proper subobject in $\mathcal{C}_\alpha$. Thus, by the observation at the start of the proof, there exists $x : \mathbf{1} \longrightarrow J_{0,\alpha+1}X_\alpha$ in $\mathcal{C}_{\alpha+1} = \mathcal{C}_\alpha/J_{0,\alpha}X_\alpha$ such that $\mathcal{C}_{\alpha+1} \not\models \exists c\colon J_{0,\alpha+1}C.\, x = (J_{0,\alpha+1}m)(c)$. Since $J_{\alpha+1,\kappa}$ is conservative and $J = J_{\alpha+1,\kappa} \circ J_{0,\alpha}X_\alpha$, defining $d = J_{\alpha+1,\kappa}x$, we obtain that indeed $\mathcal{C}_* \not\models \exists c\colon JC.\, d = (Jm)(c)$.

For the second stage, define categories

$$\begin{aligned}
\mathcal{C}^0 &= \mathcal{C} & \\
\mathcal{C}^{n+1} &= (\mathcal{C}^n)_* & \text{using the construction above} \\
\mathcal{C}^* &= \varinjlim_{i<\omega} \mathcal{C}^i &
\end{aligned}$$

Again by Lemma 10.5, these categories are all categories of classes with conservative functors between them, and they all satisfy Collection whenever $\mathcal{C}$ does.

We show that $\mathbf{1}$ is strongly projective in $\mathcal{C}^*$. Suppose X has global support and $m\colon C \rightarrowtail X$ is proper in $\mathcal{C}^*$. Then the same is already the case in some $\mathcal{C}^n$, whence by the argument above there is an arrow $d : \mathbf{1} \longrightarrow X$ in $\mathcal{C}^{n+1}$ such that $\mathcal{C}_{n+1} \not\models \exists c\colon C.\, d = m(c)$ (here omitting explicit mention of the mediating functors $C^n \to C^{n+1} \to \mathcal{C}^*$). Since the functor $C^{n+1} \to \mathcal{C}^*$ is conservative, also $\mathcal{C}^* \not\models \exists c\colon C.\, d = m(c)$. It follows immediately that d does not factor through m in $\mathcal{C}^*$. □

Lemma 10.9 *If* $\mathbf{1}$ *is strongly projective in a category of classes* $\mathcal{C}$*, then* $\mathcal{C}$ *has small generators.*

PROOF. Suppose we have any proper subobject $B \rightarrowtail C$ in $\mathcal{C}$, and consider its image $\mathcal{P}B \rightarrowtail \mathcal{P}C$ under the small powerobject functor. Since $B \rightarrowtail C$

is proper, so is $\mathcal{P}B \rightarrowtail \mathcal{P}C$. Since $\mathcal{P}C \twoheadrightarrow \mathbf{1}$ and $\mathbf{1}$ is strongly projective, there is a point $a : \mathbf{1} \longrightarrow \mathcal{P}C$ that does not factor through $\mathcal{P}B$. Then a classifies a small subobject $A \rightarrowtail C$ that does not factor through B. $\square$

Lemma 10.10 *If* $\mathbf{1}$ *is projective in a category of classes* $\mathcal{C}$ *with Collection, then* $\mathcal{C}$ *has small covers.*

PROOF. Suppose $e\colon C \twoheadrightarrow A$ is a regular epi with A small. By the Collection axiom $e_!\colon \mathcal{P}C \twoheadrightarrow \mathcal{P}A$ is also a regular epi. Since A is small, let $\lceil A \rceil\colon \mathbf{1} \longrightarrow \mathcal{P}A$ be the set $A \in \mathcal{P}A$ (as in the proof of Lemma 5.20). Because $\mathbf{1}$ is projective, there exists $b\colon \mathbf{1} \longrightarrow \mathcal{P}C$ such that $e \circ b = \lceil A \rceil$. Then letting $B \rightarrowtail C$ be the small subobject of C classified by b, we indeed have that the composite $B \rightarrowtail C \twoheadrightarrow A$ is a regular epi. $\square$

Combining Lemmas 10.8–10.10 yields the desired first step of the completeness proof.

Proposition 10.11 (ZFC) *Every small category of classes* $\mathcal{C}$ *satisfying Collection has a cofinal conservative functor* $\mathcal{C} \to \mathcal{C}^*$ *into a saturated class category* $\mathcal{C}^*$*.*

10.2 The derivative functor

Let $\mathcal{C}$ be a category of classes, with universal object U. We wish to map $\mathcal{C}$ to a category of ideals over the topos of small objects in $\mathcal{C}$. To do so, we require a system of inclusions on the topos. To obtain this, we define a category of classes $\mathcal{C}_{\hookrightarrow}$ equivalent to $\mathcal{C}$ that itself has a system of inclusions defined upon it.

The objects of $\mathcal{C}_{\hookrightarrow}$ are subobjects $A \rightarrowtail U$ in $\mathcal{C}$. Using choice, we assume that each object is represented by a chosen representative mono $A \rightarrowtail U$. Then the morphisms from $A \rightarrowtail U$ to $B \rightarrowtail U$ are just the morphisms from A to B in $\mathcal{C}$.

By the defining property of a universal object, it is clear that $\mathcal{C}_{\hookrightarrow}$ is a category equivalent to $\mathcal{C}$. In particular, using choice, there is an equivalence functor $\mathcal{C} \to \mathcal{C}_{\hookrightarrow}$.

We say that a map $A \longrightarrow B$ from $A \rightarrowtail U$ to $B \rightarrowtail U$ is an *inclusion* in

$\mathcal{C}_{\hookrightarrow}$ if the triangle below commutes

$$\begin{tikzcd}
A \arrow[rr] \arrow[dr, tail] & & B \arrow[dl, tail] \\
& U &
\end{tikzcd}$$

in which case $A \longrightarrow B$ is clearly a mono. It is easily seen that these inclusion maps have finite meets and joins, and thus give rise to associated finite intersection and union operations on objects of $\mathcal{C}_{\hookrightarrow}$.

Because $\mathcal{C}_{\hookrightarrow}$ is equivalent to $\mathcal{C}$, it too is a category of classes. We define the required structure in a way that is compatible with the inclusion maps. For this, we use a chosen subobject $\mathcal{P}U \rightarrowtail U$ to define natural products, coproducts, equalizers, and small powerobjects in $\mathcal{C}_{\hookrightarrow}$, using the same definitions used for the analogous constructions on classes in Section 2. As is easily verified, all these constructions preserve inclusion maps in a sense identical to the "structural" property of dssi's in Section 3. We refrain from going into the, at this stage tedious, details.

We write $\mathcal{E}_{\hookrightarrow}$ for the full subcategory of small objects in $\mathcal{C}_{\hookrightarrow}$. We define the topos structure on $\mathcal{E}_{\hookrightarrow}$ using the class structure described above. Then, as is easily verified, the inclusion maps in $\mathcal{E}_{\hookrightarrow}$, inherited from $\mathcal{C}_{\hookrightarrow}$, form a dssi on the topos $\mathcal{E}_{\hookrightarrow}$.

Definition 10.12 Let $\mathcal{C}$ be a category of classes. The *derivative* functor,

$$\mathbf{d} : \mathcal{C}_{\hookrightarrow} \to \mathbf{Idl}(\mathcal{E}_{\hookrightarrow})$$

is defined as follows.

> $\mathbf{d}C = \{A \hookrightarrow C \mid A \text{ small}\}$
> $\mathbf{d}f : \mathbf{d}C \longrightarrow \mathbf{d}D$, given $f : C \longrightarrow D$, is defined by factoring, as indicated in the following diagram:

$$\begin{tikzcd}
C \arrow[r, "f"] & D \\
A \arrow[u, hook] \arrow[r, dotted, "(\mathbf{d}f)_A"'] & \mathbf{d}f(A) \arrow[u, hook]
\end{tikzcd}$$

Lemma 10.13 *For any category of classes $\mathcal{C}$, the derivative functor,*

$$\mathbf{d} : \mathcal{C}_{\hookrightarrow} \to \mathbf{Idl}(\mathcal{E}_{\hookrightarrow})$$

preserves the following structure.

(1) finite limits and coproducts,
(2) small maps,
(3) powerobjects $\mathcal{P}C$,
(4) and the universal object U.

PROOF. Routine verification. Briefly:

(1) Given $C \times D$ in $\mathcal{C}_{\hookrightarrow}$,

$$\begin{aligned}\mathbf{d}(C \times D) &= \{S \hookrightarrow C \times D \mid S \in \mathcal{E}_{\hookrightarrow}\} \\ &= \{S \hookrightarrow C' \times D' \mid S, C', D' \in \mathcal{E}_{\hookrightarrow},\ C' \hookrightarrow C,\ D' \hookrightarrow D\} \\ &= \mathbf{d}(C) \times \mathbf{d}(D)\end{aligned}$$

by factoring any small subobject $S \rightarrowtail C \times D$ into $S \rightarrowtail C' \times D' \rightarrowtail C \times D$ with $C' \rightarrowtail C$ and $D \rightarrowtail D'$ small subobjects. The other cases are similar.

(2) Let $f : C \longrightarrow D$ be a small map. Since for any small subobject $B \rightarrowtail D$, the pullback $f^{-1}(B) \rightarrowtail C$ is also a small subobject, we can define an inverse image for $\mathbf{d}f : \mathbf{d}C \longrightarrow \mathbf{d}D$ by setting:

$$(\mathbf{d}f)^{-1}(B) = f^{-1}(B)$$

It is easily checked this satisfies the required property.

(3) For any $C \in \mathcal{C}_{\hookrightarrow}$ and small $A \hookrightarrow \mathcal{P}C$, the subobject $\bigcup A \hookrightarrow C$ is also small, and it satisfies that $A \hookrightarrow \mathcal{P}X$ iff $\bigcup A \hookrightarrow X$ for all $X \hookrightarrow C$, cf. Proposition 3.13. Thus any small subobject $A \hookrightarrow \mathcal{P}C$ can be factored as $A \hookrightarrow \mathcal{P}B \hookrightarrow \mathcal{P}C$ for some small $B \hookrightarrow C$, namely $B = \bigcup A$. We therefore have:

$$\begin{aligned}\mathbf{d}(\mathcal{P}C) &= \{A \hookrightarrow \mathcal{P}C \mid A \in \mathcal{E}_{\hookrightarrow}\} \\ &= \{A \hookrightarrow \mathcal{P}B \mid A, B \in \mathcal{E}_{\hookrightarrow},\ B \hookrightarrow C\} \\ &= \{A \hookrightarrow \mathcal{P}B \mid A \in \mathcal{E}_{\hookrightarrow},\ B \in \mathbf{d}C\} \\ &= \mathcal{P}(\mathbf{d}C)\end{aligned}$$

(4) For the universal object U, the ideal $\mathbf{d}\,U = \mathcal{E}_{\hookrightarrow}$ is a universal object in $\mathbf{Idl}(\mathcal{E}_{\hookrightarrow})$.

□

Lemma 10.14 *Let $\mathcal{C}$ be a category of classes and $\mathbf{d} : \mathcal{C}_{\hookrightarrow} \to \mathbf{Idl}(\mathcal{E}_{\hookrightarrow})$ the derivative functor.*

(1) If $\mathcal{C}$ has small covers, then $\mathbf{d}$ preserves regular epis.

(2) *If $\mathcal{C}$ has small generators, then* $\mathbf{d}$ *preserves dual images and proper subobjects.*

PROOF. To prove 1 , suppose $\mathcal{C}$ has small covers. As $\mathcal{C}_{\hookrightarrow}$ is equivalent to $\mathcal{C}$, it also has small covers. Take any regular epimorphism $e : C \twoheadrightarrow D$ in $\mathcal{C}_{\hookrightarrow}$. To show that the morphism $\mathbf{d}e \colon \mathbf{d}C \longrightarrow \mathbf{d}D$ is a regular epi in $\mathbf{Idl}(\mathcal{E}_{\hookrightarrow})$, we must show that the mapping part $A \mapsto (\mathbf{d}e)(A)$ is surjective. For this, take $B \in \mathbf{d}D$ and pull back the inclusion $B \hookrightarrow D$ along e as in the right-hand square below.

$$\begin{tikzcd}
A \arrow[r, hook] \arrow[dr, two heads] & B' \arrow[r, hook] \arrow[d, two heads] \arrow[dr, phantom, "\lrcorner", very near start] & C \arrow[d, two heads, "e"] \\
 & B \arrow[r, hook] & D
\end{tikzcd}$$

Since B is small, by small covers, there exists small subobject $A \hookrightarrow B'$ such that the composite $A \hookrightarrow B' \twoheadrightarrow B$ is a regular epi. Thus we have $A \in \mathbf{d}C$ with $(\mathbf{d}e)(A) = B$, as required.

For 2, suppose $\mathcal{C}$,and hence $\mathcal{C}_{\hookrightarrow}$, has small generators. Consider the following situation in $\mathcal{C}_{\hookrightarrow}$:

$$\begin{tikzcd}
S \arrow[d, hook] & \forall_f S \arrow[d, hook] \\
C \arrow[r, "f"] & D
\end{tikzcd}$$

We want to show:

$$\mathbf{d}(\forall_f S) = \forall_{\mathbf{d}f}\, \mathbf{d}S$$

While we know:

$$\begin{aligned}
\mathbf{d}(\forall_f S) &= \{B \hookrightarrow D \ \mid f^{-1}B \hookrightarrow S\} \ , \\
\forall_{\mathbf{d}f}\, \mathbf{d}S &= \{B \hookrightarrow D \ \mid \forall \text{ small } A \hookrightarrow f^{-1}B, \text{ it holds that } A \hookrightarrow S\} \ ,
\end{aligned}$$

the latter by the explicit description of dual images in $\mathbf{Idl}(\mathcal{E}_{\hookrightarrow})$ given in the proof of Lemma 8.4. The inclusion $\mathbf{d}(\forall_f S) \subseteq \forall_{\mathbf{d}f}\, \mathbf{d}S$ is easy. For the converse, suppose that $B \in \forall_{\mathbf{d}f}\, \mathbf{d}S$, i.e. every small subobject of $f^{-1}B$ is included in S. Using the small generator property, it follows that $(f^{-1}B) \cap S = S$, equivalently $f^{-1}B \hookrightarrow S$. Thus indeed $B \in \mathbf{d}(\forall_f S)$.

Finally, to show that $\mathbf{d}$ preserves proper subobjects, suppose that $i \colon C \hookrightarrow D$

in $C_{\hookrightarrow}$ is such that $\mathbf{d}i$ is an isomorphism. Then the inclusion

$$\{A \hookrightarrow C \mid A \text{ small}\} \subseteq \{A \hookrightarrow D \mid A \text{ small}\}$$

is an equality. So, by the small generator property, i is also an iso as required. $\square$

Combining the last two lemmas, we have that if $\mathcal{C}$ is saturated then the derivative functor $\mathbf{d}\colon C_{\hookrightarrow} \to \mathbf{Idl}(\mathcal{E}_{\hookrightarrow})$ is logical, cofinal and conservative. Thus, composing with the equivalence (obtained using choice) $\mathcal{C} \to \mathcal{C}_{\hookrightarrow}$, we have completed the desired step 2.

Proposition 10.15 (ZFC) *If $\mathcal{C}$ is saturated, then there is a cofinal conservative functor $\mathcal{C} \to \mathbf{Idl}(\mathcal{E}_{\hookrightarrow})$.*

10.3 The ideal embedding theorem

Putting together the results of Sections 10.1 and 10.2, we have proved the following embedding theorem for categories of classes with Collection.

Theorem 10.16 (ZFC) *For any small category of classes $\mathcal{C}$ satisfying Collection, there exists a small topos $\mathcal{E}$ and a cofinal conservative functor $\mathcal{C} \to \mathbf{Idl}(\mathcal{E})$.*

PROOF. Combine Propositions 10.11 and 10.15. $\square$

As a corollary, finally, we have the proof of Theorem 10.1.

PROOF of Theorem 10.1. Suppose that $\mathrm{BIST}^- + \mathrm{Coll} + \mathcal{T} \nvdash \phi$. By (the proof of) Theorem 7.1, there exists a (small) category of classes $\mathcal{C}$ for which $\mathcal{C} \models \mathcal{T}$ but $\mathcal{C} \not\models \phi$. By the foregoing embedding theorem, there exists a small topos $\mathcal{E}$ and a cofinal conservative functor $\mathcal{C} \to \mathbf{Idl}(\mathcal{E})$. Then, as required, $\mathbf{Idl}(\mathcal{E}) \models \mathcal{T}$ but $\mathbf{Idl}(\mathcal{E}) \not\models \phi$ $\square$

PART IV — CONSTRUCTING SYSTEMS OF INCLUSIONS

11 Elementary and cocomplete toposes

In this section we give the postponed proofs of Theorems 3.10 and 3.18.

For any locally small topos $\mathcal{E}$, we need to construct an equivalent topos carrying a dssi. In fact, many different such constructions are possible. (By equation (2), this corresponds to there being many different ways of modelling BIST$^-$ in $\mathcal{E}$.) We take a two stage approach. First, in Section 11.1, we construct an equivalent topos carrying a directed system of inclusions. Second, in Section 11.2, we use the new topos as a basis for the construction of another equivalent topos in which the system of inclusions is also structural. The two steps are combined in Section 11.3 to yield the proof of Theorem 3.10.

Theorem 3.18 is proved simultaneously. For each step, we explain the minor modifications needed to obtain a superdirected system of inclusions in the case that $\mathcal{E}$ is a cocomplete topos.

Actually, in Section 11.2, we give two different constructions of dssi's. Each validates, under the forcing interpretation, additional set-theoretic axioms not included in BIST. In Section 11.4 we present the two set theories BIZFA (Basic Intuitionistic Zermelo-Fraenkel with Atoms), and BINWFA (Basic Intuitionistic Non-Well-Founded set theory with Atoms) that are modelled by the different constructions.

11.1 Obtaining a (super)directed system of inclusions

The main goal of this section is to construct, for any locally small topos $\mathcal{E}$, an equivalent topos $\mathcal{E}'$ carrying a directed system of inclusions. We shall do this using BIST itself as the meta-theory, an enterprise which requires us to make one extra technical assumption on the topos $\mathcal{E}$, which would be vacuous were a classical meta-theory used instead.

Our construction makes essential use of finite sets. Because the meta-theory is intuitionistic, there are inequivalent notions of finiteness available. The one we adopt is the standard notion of "Kuratowski finite". Specifically, a set X is said to be *finite* if there exists $n \in N$ and a surjection $e \colon \{x \in N \mid x < n\} \twoheadrightarrow X$. Because equality on X need not be a decidable relation, it is not possible in general to assume that e is bijective. Also, because subsets $Y \subseteq X$ need not be decidable, it is not necessarily the case that a subset of a finite set is again finite. See e.g. [[REFS]] for further discussion of Kuratowski finiteness.

Let $\mathcal{E}$ be a locally small topos. We say that $\mathcal{E}$ *interprets equality of objects* if, for every $X, Y \in |\mathcal{E}|$, the category $\mathcal{E}$ possesses a chosen copower:

$$\coprod_{\{\emptyset | X=Y\}} \mathbf{1} \ ,$$

i.e. the copower of the terminal object $\mathbf{1}$ indexed by the subsingleton set $\{x \mid x = \emptyset \wedge X = Y\}$. Then, for any object Z of $\mathcal{E}$, the copower

$$\coprod_{\{\emptyset | X=Y\}} Z$$

exists, since it can be defined as $Z \times \coprod_{\{\emptyset | X=Y\}} \mathbf{1}$. We henceforth write $Z\lceil_{X=Y}$ for such a copower. By the universal property of this object, $X\lceil_{X=Y}$ is always isomorphic to $Y\lceil_{X=Y}$, and there are thus canonical maps $X\lceil_{X=Y} \longrightarrow X$ and $X\lceil_{X=Y} \longrightarrow Y$.

The condition that $\mathcal{E}$ interprets equality of objects may seem obscure. Fortunately, in many situations it is vacuous. For example, if REM is assumed in the meta-theory, then every topos interprets equality of objects, since a case analysis on $X = Y$ yields that the required copower is either an initial or terminal object. (Readers who are happy to use a classical meta-theory for the construction in this section are thus advised to simply ignore all further issues concerning interpreting equality of objects.) More generally, whenever the class $|\mathcal{E}|$ has decidable equality, the topos $\mathcal{E}$ interprets equality of objects. Since the free elementary topos with natural numbers object has decidable equality between objects (it can be constructed so that its set of objects is in bijection with the natural numbers), this gives an example topos that, provably in BIST, interprets equality of objects. Observe also that any cocomplete topos trivially interprets equality of objects.

We can now formulate the main result of this section, which we shall prove in BIST.

Proposition 11.1 *For any locally small topos $\mathcal{E}$ interpreting equality of objects, there exists an equivalent topos $\mathcal{E}'$ carrying a directed system of inclusions.*

To prove Proposition 11.1, we suppose, for the rest of Section 11.1, that $\mathcal{E}$ is a locally small topos interpreting equality of objects.

We first observe that, for any finite set $\mathcal{X}$ of objects of $\mathcal{E}$, there exists a coproduct diagram

$$\{A \xrightarrow{\mathrm{in}_A} C \mid A \in \mathcal{X}\}$$

in $\mathcal{E}$. Indeed, such a diagram can be constructed straightforwardly from any surjection $A_{(-)} \colon \{x \in N \mid x < n\} \twoheadrightarrow X$ witnessing the finiteness of $\mathcal{X}$. For this, first take the coproduct $\coprod_{0\leq i<n} A_i$, which is easily defined using empty and binary coproducts in $\mathcal{E}$. Then consider the two maps:

$$\coprod_{0\leq i,j<n} A_i\lceil_{A_i=A_j} \rightrightarrows \coprod_{0\leq i<n} A_i \ , \tag{35}$$

whose components at i, j are respectively

$$A_i\lceil_{A_i=A_j} \longrightarrow A_i \xrightarrow{\mathrm{in}_i} \coprod_{0\leq i<n} A_i$$
$$A_i\lceil_{A_i=A_j} \longrightarrow A_j \xrightarrow{\mathrm{in}_j} \coprod_{0\leq i<n} A_i \ ,$$

defined using the canonical maps mentioned above. The desired coproduct object C is given by the coequalizer of the parallel pair (35). Then for $A \in \mathcal{X}$, the injection $\mathrm{in}_A\colon A \longrightarrow C$ is the composite:

$$A_i \xrightarrow{\mathrm{in}_i} \coprod_{0\leq i<n} A_i \twoheadrightarrow C \ ,$$

where i is such that $A_i = A$. This injection is easily seen to be independent of the choice of i.

The above construction shows the existence of coproduct diagram for any finite set $\mathcal{X}$ of objects of $\mathcal{E}$. Note that the coproduct diagram obtained depends upon the presentation of the finite set $\mathcal{X}$ by a finite enumeration. In general, in BIST, there is no natural way to select a chosen coproduct for $\mathcal{X}$.

We now define the topos $\mathcal{E}'$. The objects are tuples $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$, where:

(1) $\mathcal{X}$ is a finite set of objects of $\mathcal{E}$,
(2) $p_A\colon A \longrightarrow \Omega$ (we write $P_A \rightarrowtail A$ for the associated subobject),
(3) C is an object of $\mathcal{E}$, and
(4) $\{\mathrm{in_A}\colon P_A \longrightarrow C\}_{A\in\mathcal{X}}$ is a coproduct diagram in $\mathcal{E}$.

The morphisms from $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$ to $(\mathcal{Y}, \{p_B\}_{B\in\mathcal{Y}}, D, \{\mathrm{in}_B\}_{B\in\mathcal{Y}})$ are simply the morphisms from C to D in $\mathcal{E}$.

Proposition 11.2 *$\mathcal{E}'$ is equivalent to $\mathcal{E}$ and hence an elementary topos.*

PROOF. The equivalence functor from $\mathcal{E}'$ to $\mathcal{E}$ maps $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$ to C. That in the opposite direction takes C to $(\{C\}, \{x \mapsto \top\}, C, \{1_C\})$. $\square$

It remains to define a directed system of inclusions on $\mathcal{E}'$. In order to do this, given an object $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$ of $\mathcal{E}'$, it is necessary to extend the family $\{p_A\}_{A\in\mathcal{X}}$ to a family $\{p_A\colon A \longrightarrow \Omega\}_{A\in|\mathcal{E}|}$ indexed by the class of all objects in $\mathcal{E}$. In doing so, we make essential use of the requirement that $\mathcal{E}$ interprets equality of objects. Given any object $B \in |\mathcal{E}|$ consider the composite

$$\coprod_{A\in\mathcal{X}} P_A\lceil_{A=B} \rightarrowtail \coprod_{A\in\mathcal{X}} A\lceil_{A=B} \rightarrowtail B \ , \tag{36}$$

where the first map is the sum of the monos $P_A\lceil_{A=B} \rightarrowtail A\lceil_{A=B}$, and the second has as components the canonical maps $A\lceil_{A=B} \longrightarrow B$. Clearly the first map is a mono. The second is also a mono because it factors as

$$\coprod_{A\in\mathcal{X}} A\lceil_{A=B} \rightarrowtail \coprod_{A\in\mathcal{X}\cup\{B\}} A\lceil_{A=B} \ \cong\ B \ ,$$

where the first component is the evident mono, and the isomorphism exists as a consequence of the universal properties of the coproducts involved. The subobject of B defined by (36) above is independent of the choice of coproduct in its definition, and hence determines $p_B\colon B \to \Omega$, whence a canonical $P_B \rightarrowtail B$. There is also a canonical $\mathrm{in}_B\colon P_B \rightarrowtail C$, given by:

$$P_B \cong \coprod_{A\in\mathcal{X}} P_A\lceil_{A=B} \rightarrowtail \coprod_{A\in\mathcal{X}} P_A \cong C \ .$$

One easily shows that, for $B \in \mathcal{X}$, the map p_B defined above is equal to that originally specified by the family $\{p_A\}_{A\in\mathcal{X}}$, hence the defined P_B also coincides with the corresponding P_A. Also, the canonical $\mathrm{in}_B\colon P_B \rightarrowtail C$ defined above is equal to the coproduct injection specified in the diagram $\{\mathrm{in}_A\}_{A\in\mathcal{X}}$. Thus there is no ambiguity in the notation. Furthermore, for any finite $\mathcal{Z} \supseteq \mathcal{X}$, it holds that $\{\mathrm{in}_A\}_{A\in\mathcal{Z}}$ is a coproduct diagram with vertex C. This is a straightforward consequence of the definition of P_A as a coproduct, as in (36) above.

A morphism f from $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$ to $(\mathcal{Y}, \{q_B\}_{B\in\mathcal{Y}}, D, \{\mathrm{in}'_B\}_{B\in\mathcal{Y}})$ is defined to be an *inclusion* if, for every $A \in \mathcal{X}$ there exists a (necessarily monomorphic) map $i_A\colon P_A \rightarrowtail Q_A$ fitting into the diagram below.

$$\begin{tikzcd}
& P_A \arrow[r, tail, "\mathrm{in}_A"] \arrow[dl, tail] \arrow[d, dotted, "i_A"'] & C \arrow[d, "f"] \\
A & Q_A \arrow[l, tail] \arrow[r, tail, "\mathrm{in}'_A"'] & D
\end{tikzcd} \tag{37}$$

where Q_A is obtained by extending the family $\{q_B\}_{B\in\mathcal{Y}}$ to $\{q_B\}_{B\in|\mathcal{E}|}$, as above. When f is an inclusion, then, for arbitrary $A \in |\mathcal{E}|$, there in fact exists $i_A\colon P_A \rightarrowtail Q_A$ fitting into the diagram above (where, now, P_A is also from the extended family). This follows easily from the definition of P_A as a coproduct in (36) above. Since $\{in_A\}_{A\in\mathcal{X}\cup\mathcal{Y}}$ and $\{in'_B\}_{B\in\mathcal{X}\cup\mathcal{Y}}$ are coproduct diagrams with vertices C and D respectively, it follows that $f = \coprod_{A\in\mathcal{X}\cup\mathcal{Y}} i_A$; hence, by stability of coproducts, the square in (37) above is always a pullback.

Proposition 11.3 *The inclusion maps defined above provide a directed system of inclusions on $\mathcal{E}'$.*

PROOF. For (si1), we have observed above that, any inclusion map f is a coproduct $\coprod_{A\in\mathcal{X}\cup\mathcal{Y}} i_A$ of monomorphisms, and hence itself a monomorphism.

For (si2), we must show that there is at most one inclusion between any two objects. But, given objects $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$ and $(\mathcal{Y}, \{q_B\}_{B\in\mathcal{Y}}, D, \{\mathrm{in}'_B\}_{B\in\mathcal{Y}})$, each monos i_A, for $A \in \mathcal{X}$, in diagram (37), is uniquely determined by the left-hand triangle. The inclusion f can thus exist only in the case that all the uniquely determined i_A maps exist, in which case f is itself uniquely determined by the C being the vertex of the coproduct diagram $\{\mathrm{in}_A\}_{A\in\mathcal{X}}$.

For (si3), consider any monomorphism m into $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$ given by a monomorphism $m\colon P \rightarrowtail C$ in $\mathcal{E}$. For each $A \in \mathcal{X}$ define $p'_A\colon A \to \Omega$ to be the characteristic map for the top edge of the diagram below, constructed by pullback.

$$\begin{tikzcd}
P'_A \arrow[r, tail] \arrow[d, tail, "\mathrm{in}'_A"'] \arrow[dr, phantom, "\lrcorner", very near start] & P_A \arrow[r, tail] \arrow[d, tail, "\mathrm{in}_A"] & A \\
P \arrow[r, tail, "m"'] & C &
\end{tikzcd}$$

Then, as is easily seen, the mono m is an inclusion map from the object $(\mathcal{X}, \{p'_A\}_{A\in\mathcal{X}}, P, \{\mathrm{in}'_A\}_{A\in\mathcal{X}})$ to $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$.

For (si4), suppose we have

$$\begin{aligned}
f\colon (\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}}) &\longrightarrow (\mathcal{Y}, \{q'_A\}_{A'\in\mathcal{Y}}, D, \{\mathrm{in}'_{A'}\}_{A'\in\mathcal{Y}}) \\
g\colon (\mathcal{Y}, \{q_{A'}\}_{A'\in\mathcal{Y}}, D, \{\mathrm{in}'_{A'}\}_{A'\in\mathcal{Y}}) &\longrightarrow (\mathcal{Z}, \{r_{A''}\}_{A''\in\mathcal{Z}}, E, \{\mathrm{in}''_{A''}\}_{A''\in\mathcal{Z}}) \ ,
\end{aligned}$$

with $g \circ f$ and g inclusions. Thus, for every $A \in |\mathcal{E}|$, there exist maps k_A and

j_A fitting into the diagrams below.

$$\begin{tikzcd}
 & P_A \arrow[r, tail, "\mathrm{in}_A"] \arrow[dl, tail] \arrow[d, "k_A"'] \arrow[dr, phantom, "\lrcorner", very near start] & C \arrow[d, "g \circ f"] \\
A & R_A \arrow[l, tail] \arrow[r, tail, "\mathrm{in}''_A"'] & E
\end{tikzcd}
\qquad
\begin{tikzcd}
 & Q_A \arrow[r, tail, "\mathrm{in}'_A"] \arrow[dl, tail] \arrow[d, "j_A"'] \arrow[dr, phantom, "\lrcorner", very near start] & D \arrow[d, "g"] \\
A & R_A \arrow[l, tail] \arrow[r, tail, "\mathrm{in}''_A"'] & E
\end{tikzcd}$$

By the pullback on the right, each k_A is of the form $j_A \circ i_A$ for a unique $i_A\colon P_A \rightarrowtail Q_A$. These i_A are easily seen to have the required properties to show that f is an inclusion.

Finally, we show directedness. Consider objects $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$ and $(\mathcal{Y}, \{q_B\}_{B\in\mathcal{Y}}, D, \{\mathrm{in}'_B\}_{B\in\mathcal{Y}})$. The required upper bound (in fact the union) is given simply by $(\mathcal{X}\cup\mathcal{Y}, \{r_A\}_{A\in\mathcal{X}\cup\mathcal{Y}}, E, \{\mathrm{in}''_A\}_{A\in\mathcal{X}\cup\mathcal{Y}})$, where: $r_A\colon A \to \Omega$ is the characteristic map for the subobject $P_A \cup Q_A \rightarrowtail A$ (where, of course, P_A and Q_A are taken from the extended families $\{P_A\}_{A\in|\mathcal{E}|}, \{Q_A\}_{A\in|\mathcal{E}|}$); and E and $\{\mathrm{in}''_A\}_{A\in\mathcal{X}\cup\mathcal{Y}}$ is a coproduct cocone for the family . $\{R_A\}_{A\in\mathcal{X}\cup\mathcal{Y}}$. It is easy to see that both $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$ and $(\mathcal{Y}, \{q_B\}_{B\in\mathcal{Y}}, D, \{\mathrm{in}'_B\}_{B\in\mathcal{Y}})$ are included in this object. There is one remaining niggle. We have been assuming throughout the paper that directedness should supply a specified upper bound for any two objects. Thus we have to specify a canonical coproduct $(E, \{\mathrm{in}''_A\}_{A\in\mathcal{X}\cup\mathcal{Y}})$. This is achieved as follows. Any coproduct E induces an obvious canonical epi $C + D \twoheadrightarrow E$; so simply take E to be the canonical quotient in $|\mathcal{E}|$ of the unique equivalence relation on $C + D$ induced by such quotients. $\square$

This completes the proof of Proposition 11.1.

To end this section, we turn to the special case of cocomplete toposes.

Proposition 11.4 *For any locally small cocomplete topos $\mathcal{E}$, there exists an equivalent topos $\mathcal{E}'$ carrying a superdirected system of inclusions.*

Perhaps surprisingly, only BIST itself is required as the meta-theory for this result.

PROOF. To avoid unnecessary repetition, we simply indicate the modifications required to the proof of Proposition above. The first main change is to the construction of $\mathcal{E}'$, where objects are now tuples $(\mathcal{X}, \{p_A\}_{A\in\mathcal{X}}, C, \{\mathrm{in}_A\}_{A\in\mathcal{X}})$, where $\mathcal{X}$ is an *arbitrary* set of objects of $\mathcal{E}$, and conditions (2)–(4) on objects remain as before. The definition of the inclusion maps also remains unaltered, as does the proof that inclusion maps form a system of inclusions. It remains to prove superdirectedness. Suppose then that $\{(\mathcal{X}_i, \{p^i_A\}_{A\in\mathcal{X}_i}, C_i, \{\mathrm{in^i}_A\}_{A\in\mathcal{X}_i})\}_{i\in I}$ is a family of objects of $\mathcal{E}'$. The required upper bound is then:

$$(\bigcup_{i\in I} \mathcal{X}_i,\ \{r_A\}_{A\in\bigcup_{i\in I}\mathcal{X}_i},\ E,\ \{\mathrm{in}_A\}_{A\in\bigcup_{i\in I}\mathcal{X}_i})\ ,$$

where $r_A\colon A \to \Omega$ is the characteristic map for the subobject $(\bigcup_{i\in I} P^i_A) \rightarrowtail A$, obtainned using the cocompleteness of $\mathcal{E}$, and where $E = \coprod_{A\in\bigcup_{i\in I}\mathcal{X}_i} R_A$, with $\{\mathrm{in}_A\}_{A\in\bigcup_{i\in I}\mathcal{X}_i}$ the corresponding coproduct diagram, again obtained using cocompleteness. Note that, this time, there is no difficulty in obtaining a canonical E since it can simply be taken to be the specified coproduct available via cocompleteness. □

11.2 Implementing the structural property

In this section, we prove the proposition below.

Proposition 11.5 *For any locally small topos $\mathcal{E}$ with directed system of inclusions $\mathcal{I}$, there exists an equivalent topos $\mathcal{E}_{nwf}$ carrying a directed structural system of inclusions $\mathcal{I}_{nwf}$. Moreover, if $\mathcal{I}$ is superdirected and $\mathcal{E}$ is cocomplete then $\mathcal{I}_{nwf}$ is also superdirected.*

For the remainder of Section 11.2, let $\mathcal{E}$ be an elementary topos with directed system of inclusions $\mathcal{I}$.

A *membership graph* is a structure $G = (|G|, A_G, r_G)$ where $|G|$ and A_G are objects of $\mathcal{E}$ and $r_G\colon |G| \longrightarrow A_G + \mathcal{P}|G|$ is a morphism in $\mathcal{E}$. One thinks of $|G|$ as a set of *vertices* with each vertex $x \in |G|$ being either, in the case that $r_G(x) = \mathrm{inl}(a)$, an *atom* $a\colon A_G$, or, in the case that $r_G(x) = \mathrm{inr}(d)$, a *branching vertex* with adjacency set $d \subseteq |G|$.

The relation of *bisimilarity* between two membership graphs G, H is defined, internally in $\mathcal{E}$, as the greatest element $\sim_{G,H}\colon \mathcal{P}(|G| \times |H|)$ satisfying:

$$\begin{aligned} x \sim_{G,H} y \text{ iff } & (\exists\, a\colon A_G,\, b\colon A_H.\ r_G(x) = \mathrm{inl}(a) \ \wedge\ r_H(y) = \mathrm{inl}(b) \ \wedge\ a =_{A_G,A_H} b) \\ & \vee\ (\exists\, d\colon \mathcal{P}|G|,\, e\colon \mathcal{P}|H|.\ r_G(x) = \mathrm{inr}(d) \ \wedge\ r_H(y) = \mathrm{inr}(e) \ \wedge \\ & \qquad (\forall x' \in d.\, \exists y' \in e.\ x' \sim_{G,H} y') \ \wedge \\ & \qquad (\forall y' \in e.\, \exists x' \in d.\ x' \sim_{G,H} y')\ )\ , \end{aligned}$$

making use of the heterogeneous equality on $\mathcal{E}$, supplied by its directed system of inclusions, as defined above Lemma 3.7.

As is standard, $\sim_{G,H}$ is in fact the largest relation satisfying just the left-to-right implication of the above equivalence. Using this fact, and Lemma 3.7, one easily proves the lemma below.

Lemma 11.6 *For membership graphs G, H, I, the following hold internally in $\mathcal{E}$.*

(1) $x \sim_{G,G} x$.
(2) $x \sim_{G,H} y$ *implies* $y \sim_{H,G} x$.
(3) if $x \sim_{G,H} y$ *and* $y \sim_{H,I} z$ *then* $x \sim_{G,I} z$.

One might say that the family of $\sim$ relations is a heterogeneous equivalence relation over membership graphs. In particular, each $\sim_{G,G}$ is an equivalence relation, which we henceforth write more simply as $\sim_G$.

We now define the topos $\mathcal{E}_{\mathrm{nwf}}$ over which we shall construct a directed structural system of inclusions. An object of $\mathcal{E}_{\mathrm{nwf}}$ is a triple (D, m, G) where G is a membership graph, D is an object of $\mathcal{E}$, and $m \colon D \rightarrowtail |G|$ is a monomorphism in $\mathcal{E}$. Since D is a subobject of $|G|$, the equivalence relation $\sim_G$ restricts to an equivalence relation on D, and we write simply $d \sim_G d'$ rather than $m(d) \sim_G m(d')$. Similarly, for notational simplicity, we usually write (D, G) for an object, leaving the monomorphism m implicit. Of course, rather than resorting to such notational devices, an alternative would be to simply require $D \hookrightarrow |G|$, so that D determines m. However, since this does not lead to any real simplification, we find it clearer to restrict usage of the inclusion maps in $\mathcal{E}$ to the one place where they are really essential: the definition of bisimilarity above.

A morphism from (D, G) to (E, H) is given by a binary relation F between D and E satisfying, internally in $\mathcal{E}$:

(F1) if $d \sim_G d'$, $e \sim_H e'$ and $F(d, e)$ then $F(d', e')$; and
(F2) for all $d \colon D$ there exists $e \colon E$ such that $F(d, e)$;
(F3) if $F(d, e)$ and $F(d, e')$ then $e \sim_H e'$.

The first condition says that F is saturated under the equivalence relation; the second that F is a total relation; and the third that F is single-valued up to equivalence. Such relations are easily seen to be closed under relational composition, and this defines composition in the category $\mathcal{E}_{\mathrm{nwf}}$.

Proposition 11.7 *$\mathcal{E}_{nwf}$ is equivalent to $\mathcal{E}$ and hence an elementary topos.*

PROOF. The equivalence functor from $\mathcal{E}_{\mathrm{nwf}}$ to $\mathcal{E}$ maps (D, m, G) to the quotient $D/{\sim_G}$. The functor in the opposite direction maps an object A of $\mathcal{E}$ to $(A, 1_A, \Delta_A)$ where Δ_A is the "discrete" membership graph (A, A, inl). $\Box$

Next we define the inclusion maps $\mathcal{I}_{\mathrm{nwf}}$ in $\mathcal{E}_{\mathrm{nwf}}$. A morphism F from (D, G) to (E, H) is defined to be an *inclusion* if, internally in $\mathcal{E}$:

(I) $F(d, e)$ implies $d \sim_{G,H} e$.

In fact, if F is an inclusion then also $d \sim_{G,H} e$ implies $F(d, e)$, by a straightforward argument combining (F1), (F2), (I) and Lemma 11.6.

Proposition 11.8 *$\mathcal{I}_{nwf}$ is a system of inclusions on $\mathcal{E}_{nwf}$.*

PROOF. For (si1), one notes that, in general, a morphism F from (D, G) to (E, H) is a mono if and only if $F(d, e)$, $F(d', e')$ and $e \sim_H e'$ together imply that $d \sim_G d'$. This holds for inclusions by Lemma 11.6.

For (si2), for any inclusion F, we have $F(d, e)$ if and only if $d \sim_{G,H} e$. Obviously this determines F uniquely.

For (si3), suppose that F from (D, G) to (E, H) is a monomorphism. Let E' be the subobject of E' defined internally as $\{e \colon E \mid \exists d \colon D.\ F(d, e)\}$. Easily, the identity relation on E' is an inclusion from (E', H) to (E, H), and F factors through this inclusion. The inverse map from (E', H) to (D, G) is given by the relation $\{(e, d) \colon E' \times D \mid F(d, e)\}$, which has the required properties by the characterization of monomorphisms in the proof of (si1) above.

For (si4), suppose we have $F \colon (D, G) \longrightarrow (D', G')$ and $I \colon (D', G') \longrightarrow (D'', G'')$ with I and $I \circ F$ inclusions. To show that F is an inclusion, suppose that $F(d, d')$. By (F2), there exists d'' with $I(d', d'')$, so also $(I \circ F)(d, d'')$. Then $d' \sim_{G',G''} d''$ and $d \sim_{G,G''} d''$ because I and $I \circ F$ are inclusions. Thus indeed $d \sim_{G,G'} d'$, by Lemma 11.6. $\square$

Proposition 11.9 *The system of inclusions $\mathcal{I}_{nwf}$ is directed. Moreover, if $\mathcal{I}$ is superdirected and $\mathcal{E}$ is cocomplete then it is superdirected.*

PROOF. Consider any two objects (D, G) and (D', G'). Take B to be the specified upper bound for A_G and $A_{G'}$, using the directedness of $\mathcal{I}$; sp we have inclusions $i \colon A_G \hookrightarrow B$ and $i' \colon A_{G'} \hookrightarrow B$. Define a membership graph H as follows.

$$\begin{aligned} |H| &= |G| + |G'| \\ A_H &= B \\ r_H &= [(i + \mathcal{P}(\mathrm{inl})) \circ r_G,\ (i' + \mathcal{P}(\mathrm{inr})) \circ r_{G'}] \colon |G| + |G'| \longrightarrow B + \mathcal{P}(|G| + |G'|) \end{aligned}$$

Then $(D + D', H)$ is easily seen to be an upper bound for (D, G) and (D', G').

The construction in the case that $\mathcal{I}$ is superdirected and $\mathcal{E}$ is cocomplete is similar. Specifically, superdirectedness is required to find an object of atoms containing all objects of atoms of the component graphs, and cocompleteness is used to construct the required membership graph as a coproduct of the original graphs. $\square$

It remains to implement the structural property of inclusions. For this, we define an appropriate powerobject functor on $\mathcal{E}_{\mathrm{nwf}}$ such that property (ssi4) of Definition 3.8 is satisfied. Given a membership graph, G, we define its *powergraph* $\mathcal{P}_{\mathrm{gr}}\, G$ by:

$$\begin{aligned} |\mathcal{P}_{\mathrm{gr}}\, G| &= A_G + \mathcal{P}|G| \\ A_{\mathcal{P}_{\mathrm{gr}}\, G} &= A_G \\ r_{\mathcal{P}_{\mathrm{gr}}\, G} &= 1_{A_G} + \mathcal{P}(r_G)\colon A_G + \mathcal{P}|G| \longrightarrow A_G + \mathcal{P}(A_G + \mathcal{P}|G|) \ . \end{aligned}$$

Thus:

$$\begin{aligned} r_{\mathcal{P}_{\mathrm{gr}}\, G}(\mathrm{inl}(a)) &= \mathrm{inl}(a) \\ r_{\mathcal{P}_{\mathrm{gr}}\, G}(\mathrm{inr}(d)) &= \mathrm{inr}(\{r_G(x) \mid x \in d\}) \ . \end{aligned}$$

Lemma 11.10 *Given graphs G, H, then for any $p\colon |\mathcal{P}_{gr} G|$ and $q\colon |\mathcal{P}_{gr} H|$,*

$$\begin{aligned} p \sim_{\mathcal{P}_{gr} G, \mathcal{P}_{gr} H} q \ \textit{iff} \ & (\exists\, a\colon A_G,\, b\colon A_H.\ p = \mathrm{inl}(a) \ \wedge\ q = \mathrm{inl}(b) \ \wedge\ a =_{A_G, A_H} b) \\ & \vee\ (\exists\, d\colon \mathcal{P}|G|,\, e\colon \mathcal{P}|H|.\ p = \mathrm{inr}(d) \ \wedge\ q = \mathrm{inr}(e) \ \wedge \\ & \qquad (\forall x' \in d.\, \exists y' \in e.\ x' \sim_{G,H} y') \ \wedge \\ & \qquad (\forall y' \in e.\, \exists x' \in d.\ x' \sim_{G,H} y')\ ) \ . \end{aligned}$$

PROOF. Straightforward. $\square$

Given an object (D, m, G) in $\mathcal{E}_{\mathrm{nwf}}$, we specify its powerobject $\mathcal{P}_{\mathrm{nwf}}(D, m, G)$ to be $(\mathcal{P}D,\ \mathrm{inr} \circ \mathcal{P}(m),\ \mathcal{P}_{\mathrm{gr}}\, G)$. By Lemma 11.10, one sees that $\mathcal{P}D/\sim_{\mathcal{P}_{\mathrm{gr}}\, G} \cong \mathcal{P}(D/\sim_G)$, thus $\mathcal{P}_{\mathrm{nwf}}(D, m, G)$ is indeed a carrier for the powerobject on $\mathcal{E}_{\mathrm{nwf}}$. To specify the membership relation, one first needs to define a binary product in $\mathcal{E}_{\mathrm{nwf}}$. The easiest way to do this is to go via the equivalence with $\mathcal{E}$ of Proposition 11.7. Thus $(D, G) \times (E, H)$ is defined to be the discrete object over the product $(D/\sim_G) \times (E/\sim_H)$ in $\mathcal{E}$. The membership relation between (D, m, G) and $\mathcal{P}_{\mathrm{nwf}}(D, m, G)$ is then given by:

$$\ni_{(D/\sim_G)} \rightarrowtail \mathcal{P}(D/\sim_G) \times (D/\sim_G) \ \cong\ (\mathcal{P}D/\sim_{\mathcal{P}_{\mathrm{gr}}\, G}) \times (D/\sim_G) \ .$$

Although it may seem unnatural to use the "discrete" product above, doing so avoids having to use the any of the alternative more "set-theoretic" products on membership graphs, such as one based on Kuratowski pairing, all of which are more complex. Although, ultimately, we also shall need such a set-theoretic product in order for inclusions to be structural, we use Lemma 3.9.to produce it for us, and thus do not need to consider it explicitly.

Proposition 11.11 *With powerobjects specified as above, $\mathcal{I}_{nwf}$ satisfies property* (ssi4) *of Definition 3.8.*

PROOF. Given arbitrary $F\colon (D,G) \longrightarrow (E,H)$ in $\mathcal{E}_{\mathrm{nwf}}$, the action of the covariant powerobject functor produces that $\mathcal{P}_{\mathrm{nwf}}F\colon \mathcal{P}_{\mathrm{nwf}}(D,G) \longrightarrow \mathcal{P}_{\mathrm{nwf}}(E,H)$, defined as follows.

$$(\mathcal{P}_{\mathrm{nwf}}F)(X,Y) \text{ iff } \forall d \in X.\, \exists e \in Y.\, F(d,e) \ \wedge\ \forall e \in Y.\, \exists x \in X.\, F(d,e)\ .$$

Suppose now that F is an inclusion and $(\mathcal{P}_{\mathrm{nwf}}F)(X,Y)$. Then $\forall d \in X.\, \exists e \in Y.\ d \sim_{G,H} e$ and $\forall e \in Y.\, \exists x \in X.\ d \sim_{G,H} e$, because F is an inclusion. Thus $X \sim_{\mathcal{P}_{\mathrm{gr}}G,\mathcal{P}_{\mathrm{gr}}H} Y$, by Lemma 11.10, This shows that $\mathcal{P}_{\mathrm{nwf}}F$ is indeed an inclusion, as required. □

Corollary 11.12 *It is possible to specify topos structure on $\mathcal{E}_{nwf}$ such that the directed system of inclusions $\mathcal{I}_{nwf}$ is structural.*

PROOF. Apply Lemma 3.9. □

The above corollary finally completes the proof of Proposition 11.5.

The topos $\mathcal{E}_{\mathrm{nwf}}$ represents sets using membership graphs without any well-foundedness assumption. As one might expect, when the forcing interpretation is considered over $\mathcal{E}_{\mathrm{nwf}}$, a non-well-founded set theory results; see Section 11.4 below for details. We end this section with a straightforward variation on the construction of $\mathcal{E}_{\mathrm{nwf}}$ that instead gives rise to a set theory of well-founded sets.

A membership graphs G is said to be *well-founded* if, internally in $\mathcal{E}$,

$$\begin{aligned}\forall X\colon \mathcal{P}|G|.\ [\ (\forall x\colon |G|.\ ((\exists\, a\colon A_G.\ r_G(x) = \mathrm{inl}(a)) \ \to\ x \in X) \wedge \\ ((\exists\, Y\colon \mathcal{P}|G|.\ r_G(x) = \mathrm{inr}(Y)\ \wedge\ Y \subseteq X) \to x \in X)\ ] \\ \to X = |G|\ .\end{aligned}$$

The topos $\mathcal{E}_{\mathrm{wf}}$ is defined to be the full subcategory of $\mathcal{E}_{\mathrm{nwf}}$ of objects (D,m,G) where G is well-founded. Similarly, define $\mathcal{I}_{\mathrm{wf}}$ to be the restriction of $\mathcal{I}_{\mathrm{nwf}}$ to objects of $\mathcal{E}_{\mathrm{wf}}$.

Proposition 11.13

(1) The equivalence between $\mathcal{E}$ and $\mathcal{E}_{nwf}$ cuts down to an equivalence between $\mathcal{E}$ and $\mathcal{E}_{wf}$.
(2) The specified powerobject structure on $\mathcal{E}_{nwf}$ restricts to $\mathcal{E}_{wf}$.
(3) $\mathcal{I}_{wf}$ is a directed structural system of inclusions on $\mathcal{E}_{wf}$.
(4) If $\mathcal{I}$ is superdirected and $\mathcal{E}$ is cocomplete then $\mathcal{I}_{wf}$ is superdirected.

We omit the proof, which is a routine verification that the various constructions all preserve well-foundedness.

11.3 Proofs of Theorems 3.10 and 3.18

Theorems 3.10 and 3.18 are finally proved by a simple combination of the results of the previous two sections.

The following proposition is the sharper version of Theorems 3.10 referred to below the statement of the theorem.

Proposition 11.14 *Given a topos $\mathcal{E}$ that interprets equality of objects, there exists an equivalent category $\mathcal{E}'$ carrying a dssi $\mathcal{I}'$ relative to specified topos structure on $\mathcal{E}'$.*

PROOF. Combine Propositions 11.1 and 11.5. □

The proposition below simply restates Theorem 3.18.

Proposition 11.15 *For any cocomplete topos $\mathcal{E}$, there is an equivalent category $\mathcal{E}'$ carrying an sdssi $\mathcal{I}'$ relative to specified topos structure on $\mathcal{E}'$.*

PROOF. Combine Propositions 11.4 and 11.5. □

11.4 The set theories BIZFA *and* BINWFA

As discussed in Section 2, the axioms comprising BIST formalize the constructions on sets that are useful in everyday mathematical practice. Nevertheless, there are many standard set-theoretic principles not present in BIST. In this section, we define two further set theories: BIZFA (Basic Intuitionistic Zermelo-Fraenkel set theory with Atoms) and BINWFA (Basic Intuitionistic Non-Well-Founded set theory with Atoms), which are obtained by extending BIST with just such principles. The rationale for introducing these theories at

$$
\begin{array}{rl}
\text{DS} & \mathsf{S}(x) \vee \neg\mathsf{S}(x) \\
\text{TC} & \exists y.\, x \in y \ \wedge\ \forall z \in y.\, \forall w \in z.\, w \in y \\
\text{R}{\in}\text{-Ind} & (\forall x.!\phi) \ \rightarrow\ (\forall x.\, (\forall y \in x.\, \phi[y]) \ \rightarrow\ \phi[x]) \ \rightarrow\ \forall x.\, \phi[x] \\
\text{MC} & \mathsf{S}(x) \ \wedge\ r \in U^x \ \wedge\ (\forall y \in x.\, \mathsf{S}(r(y)) \ \rightarrow\ r(y) \subseteq x) \wedge \\
& (\forall x' \subseteq x.\, (\forall y \in x.\, (\mathsf{S}(r(y)) \rightarrow \forall v \in r(y).\, v \in x') \rightarrow y \in x') \ \rightarrow\ x' = x) \\
& \quad \rightarrow\ \exists c \in U^x.\, \forall y \in x.\, (\neg\mathsf{S}(r(y)) \ \rightarrow\ c(y) = r(y)) \wedge \\
& \qquad\qquad\qquad\qquad (\mathsf{S}(r(y)) \ \rightarrow\ c(y) = \{c(v) \mid v \in r(y)\}) \\
\text{AFA} & \mathsf{S}(x) \ \wedge\ r \in U^x \ \wedge\ (\forall y \in x.\, \mathsf{S}(r(y)) \ \rightarrow\ r(y) \subseteq x) \\
& \quad \rightarrow\ \exists! c \in U^x.\, \forall y \in x.\, (\neg\mathsf{S}(r(y)) \ \rightarrow\ c(y) = r(y)) \wedge \\
& \qquad\qquad\qquad\qquad (\mathsf{S}(r(y)) \ \rightarrow\ c(y) = \{c(v) \mid v \in r(y)\})
\end{array}
$$

Fig. 7. Additional axioms for BIZFA and BINWFA

this point is that BIZFA and BINWFA are validated by the forcing interpretation in the categories $\mathcal{E}_{\mathrm{wf}}$ and $\mathcal{E}_{\mathrm{nwf}}$, constructed in Section 11.2. respectively. Thus both set theories are compatible with the internal logic of every elementary topos, and hence conservative over HAH.

The theories BIZFA and BINWFA are defined as follows

$$
\begin{array}{rlrl}
\text{BIZFA}^- = & \text{BIST}^- + \text{DS} + \text{TC} + \text{R}{\in}\text{-Ind} + \text{MC} & \text{BIZFA} = & \text{BIZFA}^- + \text{Inf} \\
\text{BINWFA}^- = & \text{BIST}^- + \text{DS} + \text{TC} + \text{AFA} & \text{BINWFA} = & \text{BINWFA}^- + \text{Inf}\,,
\end{array}
$$

where the new axioms are listed in Figure 7. We now examine these axioms in more detail.

The axiom DS (Decidable Sethood) was makes a clean division of the universe into sets and atoms (i.e. non-sets). By Lemma 2.2 and Corollary 2.6, bounded Separation, bSep, is derivable in BIZFA^-+ DS, hence in both BIZFA^- and BINWFA^-.

The axiom TC (Transitive Containment) simply states, in the obvious way, that every element of the universe (whether a set or not) is a member of a transitive set.

The schema R∈-Ind (Restricted Membership Induction) is an intuitionistically acceptable formulation of the axiom of Foundation. [24] The axiom is formulated for restricted properties only. As a special case, one obtains membership induction for bounded formulas. Also of interest is the full membership induction

[24] As has been frequently observed [REF], many classical formulations of Foundation imply unwanted cases of excluded middle.

principle:

$$\in\text{-Ind} \qquad (\forall x.\, (\forall y \in x.\, \phi[y]) \;\rightarrow\; \phi[x]) \;\rightarrow\; \forall x.\, \phi[x] \ .$$

Proposition 11.16 $\mathrm{BIZFA}^{-} + \mathrm{Sep} \vdash \in$-*Ind*.

PROOF. Immediate. □

The final axiom of BIZFA is MC (Mostowski Collapse). In this axiom, the function r represents a directed graph structure on a set x of vertices. For each vertex $y \in x$, either $r(y)$ is an atom (i.e. $\neg\mathsf{S}(y)$), in which case y is a leaf vertex and $r(y)$ is its labelling; or, if $r(y)$ is a set, then $r(y) \subseteq x$ gives the adjacency set of y. The second line of the axiom imposes, by way of stating the appropriate induction principle, the requirement that, as a relation, r is co-well-founded. The axiom then states that any vertex in such a co-well-founded graph collapses to an element of the universe by way of a function c that preserves the atoms in x and maps the graph relation on x to the membership relation. One can prove in BIZFA that c is unique.

In spite of its complexity, MC is a natural axiom to consider together with DS, TC and R$\in$-Ind. Indeed, TC and R$\in$-Ind imply that the membership relation on the transitive closure of any set x yields a co-well-founded graph, as above. MC is a kind of ontological completeness axiom, expressing that every such graph is represented by a set. Note also that DS is used implicitly in the formulation of MC, where the requirements on the function c make a case distinction on the basis of whether $r(y)$ is a set or not.

The set theory BINWFA replaces the two axioms R$\in$-Ind and MC with AFA, which is a straightforward adaptation, to a universe with atoms, of Honsell and Forti's Anti Foundation Axiom, as popularized by Aczel [**?**]. Formally, the axiom AFA is simply a strengthening of MC with the well-foundedness assumption on the membership graph dropped. Because well-foundedness is no longer assumed, in the case of AFA it is necessary to assert the uniqueness of the collapse function c.

As stated earlier, the reason for introducing the set theories BIZFA and BINWFA is because they are validated by forcing interpretations into the toposes with systems of inclusions constructed in Section 11.2. We assert this formally here, but omit the lengthy (though routine) verification entirely. At any rate, the result should not present much of a surprise, since the correctness of the new axioms is quite apparent from the construction of the toposes using membership graphs.

Proposition 11.17 *Let $\mathcal{E}_{nwf}, \mathcal{I}_{nwf}, \mathcal{E}_{wf}$ and $\mathcal{I}_{wf}$ be as constructed in Section 11.2. Then:*

(1) $(\mathcal{E}_{wf}, \mathcal{I}_{wf}) \models \mathrm{BIZFA}^-$.
(2) $(\mathcal{E}_{nwf}, \mathcal{I}_{nwf}) \models \mathrm{BINWFA}^-$.

Since, by Sections 11.1 and 11.2, every elementary topos (interpreting equality of objects) is equivalent both to toposes of the form $\mathcal{E}_{\mathrm{wf}}$ and to toposes of the form $\mathcal{E}_{\mathrm{nwf}}$, it follows that any topos can be construed both as a model of BIZFA^- and as a model of BINWFA^-. It seems plausible that the theeories BIZFA^- and BINWFA^- are actually complete relative to interpretations in toposes with sdsi's of the form $(\mathcal{E}_{\mathrm{wf}}, \mathcal{I}_{\mathrm{wf}})$ and $(\mathcal{E}_{\mathrm{nwf}}, \mathcal{I}_{\mathrm{nwf}})$ respectively. However, we have not investigated this possibility in detail.

We end the section with some simple observations about $\mathrm{BIZFA}^-$and BINWFA^-.

Lemma 11.18

(1) BIZFA $\vdash$ vN-Inf .
(2) BINWFA $\vdash$ vN-Inf .

PROOF. For statement 1, one verifies that, for each $n \in N$, the class $\{x \in N \mid x < n\}$ is a set. Define $r \in U^N$ by $r(n) = \{x \in N \mid x < n\}$. One verifies that this relation is co-well-founded (this amounts to deriving "course of values" induction for restricted properties on N). Thus MC gives a unique $c \in U^N$ satisfying the specified conditions. One verifies that the image of c, which is a set by Replacement, satisfies the properties required by vN-Inf.

Statement 2 is proved identically, using AFA rather than MC, thus there is no need to verify that the relation is co-well-founded. $\square$

The above is a typical, though very simple, application of collapsing a relation to a set. In a proof-theoretic study relating classical set theories and toposes [REF], Mathias makes a strong case for the general usefulness of such constructions. Indeed, the axiom MC (there formulated as a classically equivalent axiom called H) plays a central role in his paper.

We end our discussion by showing that the restrictedness condition on the membership induction axiom of BIZFA is essential. As happens also for the Separation, Induction and Excluded Middle axioms, if the restrictedness condition on membership induction is dropped, then the proof-theoretic strength of the set theory goes beyond that compatible with every elementary topos.

Proposition 11.19 BIZFA+ $\in$-*Ind* $\vdash$ Ind .

PROOF. Working in BIZFA, by Lemma 11.18, let N_{vN} be the smallest set containing $\emptyset$ and closed under the operation $y \mapsto y \cup \{y\}$. One easily shows that N_{vN} is isomorphic to the N already constructed. Now assume $\in$-Ind. By the above, it suffices to prove Ind for N_{vN} rather than N. For any formula $\phi[x]$, consider $\psi[x]$-$\in$-Ind, where ψ is the formula $x \in N_{\mathrm{vN}} \rightarrow \phi[x]$. Then $\psi[x]$-$\in$-Ind simplifies to

$$(\forall x \in N_{\mathrm{vN}}.\, (\forall y \in x.\, \phi[y]) \rightarrow \phi[x]) \rightarrow \forall x \in N_{\mathrm{vN}}.\, \phi[x]\,.$$

But the membership relation on the (transitive) set N_{vN} agrees with the arithmetic relation $<$. So the above states

$$(\forall x \in N_{\mathrm{vN}}.\, (\forall y < x.\, \phi[y]) \rightarrow \phi[x]) \rightarrow \forall x \in N_{\mathrm{vN}}.\, \phi[x]\,.$$

This is "course of values" induction, which directly implies the simple induction of $\phi[x]$-Ind. $\square$

Corollary 11.20 BIZFA+ $\in$-*Ind* $\vdash$ Con(HAH) .

PROOF. By Proposition 2.18.

12 Realizability toposes

Acknowledgements

This work began in summer 2000 after discussions between the first and third authors at Category Theory 2000 in Como, Italy. Various stages of this work have been presented in lectures since October 2001, in particular at: PSSL, Copenhagen, 2002 (Simpson and Butz); Logic Colloquium, Germany, 2002 (Streicher); MFPS, Pittsburgh, 2004 (Awodey); and in graduate courses at RIMS, Kyoto, autumn 2002 (Simpson) and CMU, Pittsburgh, spring 2003 (Awodey). We have especially benefited from discussions with Peter Aczel, Nicola Gambino, Kentaro Sato and Dana Scott. We acknowledge the use of Paul Taylor's diagram macros.

References

[1] P. Aczel. The type theoretic interpretation of constructive set theory. In *Logic Colloquium '77*, number 96 in Stud. Logic Foundations Math., pages 55–66. North Holland, 1978.

[2] P. Aczel and M. Rathjen. Notes on constructive set theory. Technical Report 40, 2000/2001, Mittag-Leffler Institute, Sweden, 2001.

[3] C. Butz. Bernays-Gödel type theory. *J. Pure Appl. Algebra*, 178:1–23, 2003.

[4] M.P. Fourman. Sheaf models for set theory. *J. Pure Appl. Algebra*, 19:91–101, 1980.

[5] H. Friedman. Set theoretic foundations for constructive analysis. *Ann. of Math.*, 105:1–28., 1977.

[6] P.T. Johnstone. *Sketches of an elephant: a topos theory compendium. Vols. 1–2.* Number 43–44 in Oxford Logic Guides. Oxford University Press, 2002.

[7] A. Joyal and I. Moerdijk. *Algebraic set theory.* Number 220 in London Mathematical Society Lecture Note Series. Cambridge University Press, 1995.

[8] J. Lambek and P.J. Scott. *Introduction to higher order categorical logic.* Number 7 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1986.

[9] S. Mac Lane and I. Moerdijk. *Sheaves in Geometry and Logic.* Springer-Verlag, 1992.

[10] J. Myhill. Constructive set theory. *J. Symbolic Logic*, 40:347–382, 1975.

[11] A.K. Simpson. Elementary axioms for categories of classes. In *Proc. 14th Annual IEEE Symposium on Logic in Computer Science*, pages 77–85, 1999.

[12] A. Ščedrov. Intuitionistic set theory. In *Harvey Friedman's Research on the Foundations of Mathematics*, number 117 in Stud. Logic Foundations Math., pages 257–284. North Holland, 1985.