Classifying topoi and the axiom of infinity

ANDREAS BLASS

In Memory of Evelyn Nelson

Abstract. Let $\mathcal{E}$ be an elementary topos. The axiom of infinity, asserting that $\mathcal{E}$ has a natural numbers object, is shown to be necessary—sufficiency has long been known—for the existence of an object-classifying topos over $\mathcal{E}$.

In the known constructions [1, 4, 6] of classifying topoi for geometric theories, it is always assumed that the base topos $\mathcal{E}$ satisfies the axiom of infinity. The purpose of this note is to show that this assumption is necessary. We show that, if a certain very simple geometric theory $T$ has a classifying topos over $\mathcal{E}$, then $\mathcal{E}$ has a natural numbers object. We also show, using well-known methods [1, 4, 6], that the existence of a classifying topos for $T$ follows from the existence of an object classifier, i.e., a classifying topos for the one-sorted geometric theory having no non-logical symbols and no axioms, the theory whose models in any category are just the objects of that category. Thus, our main result, which answers a question posed in [5], can be stated as follows.

THEOREM. Suppose that there exists an object classifier over the elementary topos $\mathcal{E}$. Then $\mathcal{E}$ has a natural numbers object.

Our notation and terminology will be standard [4].

Let $T$ be the one-sorted geometric theory having one constant symbol $o$, one unary function symbol $s$, no other non-logical symbols, and no axioms. Thus, a $T$-model in any topos (or in any category with a terminal object $1$) is simply a structure $\mathcal{A} = (A, o_{\mathcal{A}}, s_{\mathcal{A}})$ where $A$ is an object and $o_{\mathcal{A}} : 1 \to A$ and $s_{\mathcal{A}} : A \to A$. (We shall omit the subscripts on $o$ and $s$ whenever no confusion can arise). A homomorphism of $T$-models is a morphism between the underlying objects that commutes with the $o$ and $s$ morphisms. A natural numbers object is simply an initial $T$-model. A weakly initial $T$-model (in any topos), i.e., a $T$-model having

at least one homomorphism to every \( T \)-model, will be called a weak natural numbers object. The following lemma is based on ideas from [3, §5] (see also [4, §6.1]); the same idea, but using external rather than internal completeness, occurs in the proof of the adjoint functor theorem.

**Lemma.** If an elementary topos \( \mathcal{E} \) has a weak natural numbers object, then it has a natural numbers object.

**Proof.** Let \( \mathcal{A} = (A, o, s) \) be a weak natural numbers object. Consider the following formula \( \phi(z) \) in the Mitchell–Bénabou language [4, §5.4]:

\[
\forall x \in A \forall y \in A \forall z \in A \big( x \rightarrow s(y) \rightarrow x \rightarrow z \big),
\]

where \( x \) is of type \( \Omega^A \) and \( y \) and \( z \) are of type \( A \). (The formula "says" that \( z \) belongs to every sub-\( T \)-model of \( \mathcal{A} \).) The interpretation \( |\phi(z)| \) of this formula is a morphism \( A \to \Omega \), the characteristic morphism of a certain subobject \( N \) of \( A \). Easy calculations show that \( N \) is the universe of a sub-\( T \)-model \( N = (N, o_N, s_N) \) of \( \mathcal{A} \) and that the universe of every sub-\( T \)-model of \( \mathcal{A} \) includes \( N \). Since \( \mathcal{A} \) is a weak natural numbers object, so is every sub-\( T \)-model of \( \mathcal{A} \), in particular \( N \). To show that \( N \) is a natural numbers object, consider any two homomorphisms from \( N \) to any \( T \)-model. Their equalizer is a sub-\( T \)-model of \( N \), hence is all of \( N \) by the minimality property of \( N \), so the two homomorphisms are equal, as desired.

The following proposition, though rather easy, is the core of our argument, as it connects classifying toposes with natural numbers objects. In the statement of the proposition, \( T \) still denotes the theory defined above, although, as the referee pointed out, the lemma, the proposition, and Remark 2 at the end of the paper all remain correct, when suitably reformulated, for any algebraic (or even essentially algebraic) theory.

**Proposition.** Suppose \( T \) has a classifying topos \( \mathcal{E} \) over the topos \( \mathcal{F} \). Then \( \mathcal{F} \) has a natural numbers object.

**Proof.** By hypothesis, we have a (generic) \( T \)-model \( \mathcal{G} = (G, o, s) \) in \( \mathcal{E} \) and a geometric morphism \( p: \mathcal{E} \to \mathcal{F} \) such that, whenever \( f: \mathcal{F} \to \mathcal{F} \) is a geometric morphism and \( \mathcal{A} \) is a \( T \)-model in \( \mathcal{F} \), then there is a geometric morphism \( g: \mathcal{F} \to \mathcal{E} \) such that \( g^*(\mathcal{G}) \) is isomorphic to \( \mathcal{A} \) and \( pg \) is naturally isomorphic to \( f \). (The hypothesis actually contains more information than this, but this is all that we shall need.)

Since both the direct and inverse image functors of a geometric morphism preserve 1, they send \( T \)-models to \( T \)-models. In particular, we have \( T \)-models