Classifying Toposes for First Order Theories

Carsten Butz* and Peter Johnstone†

4 June 1997

Abstract

By a classifying topos for a first-order theory $\mathbb{T}$, we mean a topos $\mathcal{E}$ such that, for any topos $\mathcal{F}$, models of $\mathbb{T}$ in $\mathcal{F}$ correspond exactly to open geometric morphisms $\mathcal{F} \to \mathcal{E}$. We show that not every (infinitary) first-order theory has a classifying topos in this sense, but we characterize those which do by an appropriate 'smallness condition', and we show that every Grothendieck topos arises as the classifying topos of such a theory. We also show that every first-order theory has a conservative extension to one which possesses a classifying topos, and we obtain a Heyting-valued completeness theorem for infinitary first-order logic.

0 Introduction

It is by now well known that Grothendieck toposes are 'the same thing as' Morita-equivalence classes of geometric theories, in the following sense: for every geometric theory $\mathbb{T}$, there exists a topos $\mathcal{B}(\mathbb{T})$ (called the classifying topos of $\mathbb{T}$) containing a model $G_{\mathbb{T}}$ of $\mathbb{T}$ which is *generic* in the sense that, for any Grothendieck topos $\mathcal{F}$, the functor

$$\mathbf{Top}\,(\mathcal{F}, \mathcal{B}(\mathbb{T})) \longrightarrow \mathbb{T}\text{-Mod}(\mathcal{F}) \ ,$$

which sends a geometric morphism $f\colon \mathcal{F} \to \mathcal{B}(\mathbb{T})$ to the $\mathbb{T}$-model $f^*(G_{\mathbb{T}})$, is an equivalence of categories. Conversely, for every Grothendieck topos $\mathcal{E}$, there exists a geometric theory $\mathbb{T}$ such that $\mathcal{E} \simeq \mathcal{B}(\mathbb{T})$. This theory is not uniquely determined by $\mathcal{E}$, in that one can have many different presentations of 'the same' theory; but if we regard two theories as (Morita-)equivalent whenever they have equivalent categories of models in an arbitrary topos, then the correspondence between (categorical) equivalence classes of Grothendieck toposes and Morita-equivalence classes of geometric theories becomes one-to-one.

However, although a very large proportion of the first-order theories encountered in mathematical practice are (at least Morita-equivalent to) geometric ones, we do occasionally need to consider models of theories which are not geometric. (For example, as was first observed by Kock [14], it can often be profitable to consider the non-geometric first-order sentences satisfied by the generic model of a geometric theory.) For such theories, we cannot hope to have a classifying topos in the above sense, since it is well-known that geometric theories are exactly those whose models are preserved by inverse image functors. Nevertheless, we do know several characterizations of those geometric morphisms whose inverse image functors preserve arbitrary (infinitary) first-order logic: they are exactly the *open* geometric morphisms in the sense of [9] or [12]. So we might hope that, for at least some first-order theories $\mathbb{T}$ of interest, we could find a topos $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$ and a natural equivalence

$$\mathbf{Open}\,(\mathcal{F}, \mathcal{B}^{\mathrm{fo}}(\mathbb{T})) \simeq \mathbb{T}\text{-Mod}(\mathcal{F}) \ ,$$

where the left-hand side denotes the full subcategory of $\mathbf{Top}\,(\mathcal{F}, \mathcal{B}^{\mathrm{fo}}(\mathbb{T}))$ whose objects are open geometric morphisms. (Since every geometric theory may also be considered as a first-order theory, we

* *Current address: BRICS, Basic Research in Computer Science, Centre of the Danish National Research Foundation, Computer Science Department, Aarhus University, Ny Munkegade, DK–8000 Århus C, Denmark.* `butz@brics.dk`

† *Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, U.K.* `ptj@pmms.cam.ac.uk`

shall henceforth write $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$ for the classifying topos of such a theory as defined in the first paragraph, in order to distinguish it from the concept just introduced.)

Sadly, it is not possible to construct such a topos for an arbitrary first-order theory $\mathbb{T}$. The problem is, in a sense, that open geometric morphisms are 'too well-behaved': their inverse image functors preserve the interpretations, not only of finitary first-order formulae, but of arbitrary ($\mathcal{L}_{\infty\omega}$) infinitary ones as well. And there are simply too many of the latter, even if we restrict ourselves to propositional logic. It is well known ([6], [7], [21]) that the free complete Boolean algebra on a countable infinity of generators is a proper class; perhaps less familiar is the fact [11] that the corresponding problem for complete Heyting algebras occurs already with two generators—that is, there is a proper class of $\mathcal{L}_{\infty\omega}$ formulae in two propositional variables, no two of which are (constructively) provably equivalent. The model-theoretic version of this fact says that, for any cardinal κ, we can build a complete Heyting algebra H_κ of cardinality at least κ which is generated by two elements. Now let $\mathbb{T}$ be the empty theory in a language $\mathcal{L}$ with two primitive propositional symbols: we can regard the generators of H_κ as a model of $\mathbb{T}$ in the topos $\mathbf{Sh}(H_\kappa)$ of sheaves on H_κ. If the classifying topos $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$ existed, this model would be classified by an open geometric morphism $f\colon \mathbf{Sh}(H_\kappa) \to \mathcal{B}^{\mathrm{fo}}(\mathbb{T})$; since f^* preserves the interpretations of arbitrary first-order formulae, we see that it must be surjective on subobjects of 1, and hence (since this holds for all κ) that the lattice of subobjects of 1 in $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$ must be a proper class—which is impossible.

However, we shall show in this paper that the above 'cardinality problem' is the *only* obstruction to the existence of $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$. Specifically, we shall prove that, for an arbitrary $\mathbb{T}$, there exists a topos $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$ with the universal property described above if and only if, for every finite string $\vec{x}$ of variables in the language of $\mathbb{T}$, the complete Heyting algebra of $\mathbb{T}$-provable-equivalence classes of formulae with free variables in $\vec{x}$ is a set. Moreover, for any $\mathbb{T}$ and any cardinal κ such that $\mathbb{T}$ may be axiomatized by formulae in $\mathcal{L}_{\kappa\omega}$, there exists a canonical extension of $\mathbb{T}$ to a theory $\overline{\mathbb{T}}$ with the above smallness property, which is $\mathcal{L}_{\kappa\omega}$-conservative (i.e. every $\mathcal{L}_{\kappa\omega}$ sentence derivable from $\overline{\mathbb{T}}$ is derivable from $\mathbb{T}$). In particular, the generic $\overline{\mathbb{T}}$-model in $\mathcal{B}^{\mathrm{fo}}(\overline{\mathbb{T}})$ is an $\mathcal{L}_{\kappa\omega}$-conservative model of $\mathbb{T}$. Also, *every* Grothendieck topos $\mathcal{E}$ occurs, up to equivalence, as $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$ for some $\mathbb{T}$; thus we may, once again, identify Grothendieck toposes with Morita-equivalence classes of 'provably small' first-order theories.

The main ingredient in the proofs of the above statements is a characterization of the open geometric morphisms into a Grothendieck topos $\mathcal{E} = \mathbf{Sh}(\mathcal{C}, J)$ in terms of the flat and continuous functors on $\mathcal{C}$ which induce them. We prove this result in section 1 of the paper. Sections 2 and 3 are devoted to reviewing material from categorical logic, culminating in the 'Heyting-valued completeness theorem' for (intuitionistic) infinitary first-order logic. In section 4, we show how to construct the first-order theory classified by a given topos $\mathcal{E}$; in particular, if $\mathcal{E}$ is given as the classifying topos $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$ of a geometric theory $\mathbb{T}$, we show that its first-order theory may be taken to consist of all sentences satisfied in the generic model of $\mathbb{T}$. In section 5 we prove our main theorem characterizing those first-order theories which admit classifying toposes. Finally, section 6 is devoted to the discussion of some examples and applications.

Throughout the paper, for simplicity, we shall work in the 2-category of Grothendieck toposes (that is, toposes defined and bounded over the classical topos of sets $\mathbf{Set}$); the word 'topos', unless otherwise stated, will mean 'Grothendieck topos'. However, those readers who are accustomed to working over a base topos $\mathcal{S}$ will have little difficulty in extending our results to the latter more general context (for any topos $\mathcal{S}$ with a natural number object); of course, the cardinality bounds which we use have to be replaced by bounds expressed in terms of subobjects of a suitable object of $\mathcal{S}$.

Acknowledgements. The first author would like to thank Ieke Moerdijk and Steve Awodey for many helpful discussions on this topic. The second author would like to make a similar acknowledgement to Martin Hyland. Part of the first author's research was supported by a scholarship of the University of Utrecht.

1 Open maps into sheaf toposes

Let $\mathcal{E}$ be the topos of sheaves on a (small) site $(\mathcal{C}, J)$. Diaconescu's theorem [3] asserts that, for any topos $\mathcal{F}$, there is an equivalence between the category $\mathbf{Top}(\mathcal{F}, \mathcal{E})$ of geometric morphisms from $\mathcal{F}$ to $\mathcal{E}$, and the category of flat and continuous functors $\mathcal{C} \to \mathcal{F}$ (see also [16], section VII.7 or [8], Proposition

7.13). Our aim in this section is to characterize those functors $\mathcal{C} \to \mathcal{F}$ which correspond to open geometric morphisms, in the sense of [9] or [12].

Although, for most of this paper, our working definition of open maps will be that they are precisely those geometric morphisms $f: \mathcal{F} \to \mathcal{E}$ whose inverse image functors preserve full (infinitary) first-order logic, in the present section it will be convenient to use the characterization of them in terms of the map $\overline{\tau}: \Omega_{\mathcal{E}} \to f_*\Omega_{\mathcal{F}}$ which is the transpose of the characteristic function τ of the monomorphism $f^*(\top_{\mathcal{E}})$. We recall that f is open iff $\overline{\tau}$ has an internal left adjoint $\lambda: f_*\Omega_{\mathcal{F}} \to \Omega_{\mathcal{E}}$ (see [9], theorem 3.2 or [12], p. 56).

Now suppose that $\mathcal{E} = \mathbf{Sh}(\mathcal{C}, J)$, and that the geometric morphism f corresponds to a flat continuous functor $F: \mathcal{C} \to \mathcal{F}$. Let $\varepsilon: \mathcal{C} \to \mathcal{E}$ denote the composite of the Yoneda embedding $\mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ with the associated sheaf functor. Then we may identify $\Omega_{\mathcal{E}}$ with the sheaf which to an object A of $\mathcal{C}$ assigns the set of subobjects of $\varepsilon(A)$ in $\mathcal{E}$, or equivalently the set of J-closed sieves on A in $\mathcal{C}$ (recall that a sieve S is J-closed if, whenever we have a morphism $\alpha: B \to A$ and a J-covering sieve R on B such that $\alpha\beta \in S$ for all $\beta \in R$, then $\alpha \in S$). Similarly, we may identify $f_*(\Omega_{\mathcal{F}})$ with the sheaf whose value at A is the set of subobjects of $f^*\varepsilon(A) = F(A)$. In terms of these identifications, the map $\overline{\tau}$ may be identified with the action of f^* on subobjects of $\varepsilon(A)$; that is, $\overline{\tau}_A(S \rightarrowtail \varepsilon(A)) = (f^*S \rightarrowtail A)$. Note also that the action of a morphism $\alpha: A \to B$ of $\mathcal{C}$ on both $\Omega_{\mathcal{E}}$ and $f_*\Omega_{\mathcal{F}}$ is given by pullback of subobjects; since f^* preserves pullbacks, it is clear that the square

$$\begin{tikzcd}
\mathrm{Sub}(\varepsilon(B)) \arrow[r, "\overline{\tau}_B"] \arrow[d, "\varepsilon(\alpha)^*"'] & \mathrm{Sub}(F(B)) \arrow[d, "F(\alpha)^*"] \\
\mathrm{Sub}(\varepsilon(A)) \arrow[r, "\overline{\tau}_A"] & \mathrm{Sub}(F(A))
\end{tikzcd}$$

commutes.

Lemma 1.1 *The geometric morphism f is open if and only if*

(i) *for each object A in $\mathcal{C}$ and each family $\{S_i \mid i \in I\}$ of subobjects of $\varepsilon(A)$ we have $\bigwedge_{i \in I} \overline{\tau}_A(S_i) = \overline{\tau}_A(\bigwedge_{i \in I} S_i)$; and*

(ii) *for each morphism $\alpha: A \to B$ in $\mathcal{C}$ and each subobject S of $\varepsilon(A)$ we have $\forall_{F(\alpha)}\overline{\tau}_A(S) = \overline{\tau}_B\forall_{\varepsilon(\alpha)}(S)$.*

Here, as usual, $\forall_{F(\alpha)}$ and $\forall_{\varepsilon(\alpha)}$ denote the right adjoints of $F(\alpha)^*$ and $\varepsilon(\alpha)^*$ acting on subobjects. We note that, in both conditions, the equality sign could be replaced by an inclusion ($\leq$) of subobjects, since the reverse inclusion $\geq$ is automatic for any f.

Proof Since $\mathrm{Sub}(\varepsilon(A))$ and $\mathrm{Sub}(F(A))$ are complete lattices, the first condition of the lemma is equivalent to saying that each $\overline{\tau}_A$ has a left adjoint $\lambda_A : \Omega_{\mathcal{E}}(A) \to f_*\Omega_{\mathcal{F}}(A)$. Given the existence of these left adjoints, the second condition is equivalent to the commutativity of the diagram of left adjoints

$$\begin{tikzcd}
\mathrm{Sub}(F(B)) \arrow[r, "\lambda_B"] \arrow[d, "F(\alpha)^*"'] & \mathrm{Sub}(\varepsilon(B)) \arrow[d, "\varepsilon(\alpha)^*"] \\
\mathrm{Sub}(F(A)) \arrow[r, "\lambda_A"] & \mathrm{Sub}(\varepsilon(A))
\end{tikzcd}$$

for every $\alpha: A \to B$ in $\mathcal{C}$, i.e. to saying that the λ_A form a morphism of sheaves $\lambda: \Omega_{\mathcal{E}} \to f_*\Omega_{\mathcal{F}}$. So the two conditions together are equivalent to saying that $\overline{\tau}$ has an internal left adjoint in $\mathcal{E}$, i.e. that f is open. $\square$

We remark that Lemma 1.1 could also be deduced by combining [17], Lemma 6.1.8, with [9], Theorem 3.2(v).

Using the description of $\Omega_{\mathcal{E}}(A)$ as the set of closed sieves on A, we get the following translation of Lemma 1.1, which we state for future reference:

Corollary 1.2 *The geometric morphism f is open if and only if*

(i) *for each object A in $\mathcal{C}$ and each family $\{S_i \mid i \in I\}$ of closed sieves on A,*

$$\bigwedge_{i \in I} \bigvee \{\exists_{F(\beta)} F(B) \mid \beta : B \to A \in S_i\} = \bigvee \{\exists_{F(\beta)} F(B) \mid \beta : B \to A \in \bigcap_{i \in I} S_i\} \; ; \tag{1}$$

and

(ii) *for each $\alpha : A \to B$ in $\mathcal{C}$ and each closed sieve S on A,*

$$\forall_{F(\alpha)} \bigvee \{\exists_{F(\beta)} F(C) \mid \beta : C \to A \in S\} = \bigvee \{\exists_{F(\gamma)} F(D) \mid \gamma : D \to B \in \forall_\alpha S\} \; . \tag{2}$$

□

Here $\forall_\alpha S$ denotes the closed sieve

$$\{\gamma : D \to B \mid (\forall \beta : D \to A)((\alpha\beta = \gamma) \Rightarrow (\beta \in S))\} \; .$$

The map $\forall_\alpha : \Omega_{\mathcal{E}}(A) \to \Omega_{\mathcal{E}}(B)$ is the right adjoint of the action α^* of α on closed sieves.

We conclude this section with some more remarks about preservation of logical connectives. If $\{u_i : E_i \to E \mid i \in I\}$ is an epimorphic family of maps in $\mathcal{E}$, we have an embedding of complete Heyting algebras

$$\mathrm{Sub}(E) \xhookrightarrow{u^*} \prod_{i \in I} \mathrm{Sub}(E_i) \; .$$

Applying this to the epimorphic family $\{e : \varepsilon(A) \to E \mid e \in E(A), A \in \mathrm{ob}\, \mathcal{C}\}$ where E is any object of $\mathcal{E} = \mathbf{Sh}(\mathcal{C}, J)$, and to the image of this family under f^* for some geometric morphism $f : \mathcal{F} \to \mathcal{E}$, we obtain a commutative square of complete Heyting algebras

$$\begin{tikzcd} \mathrm{Sub}(E) \arrow[r, hook] \arrow[d] & \prod \mathrm{Sub}(\varepsilon(A)) \arrow[d] \\ \mathrm{Sub}(f^* E) \arrow[r, hook] & \prod \mathrm{Sub}(F(A)) \end{tikzcd} \tag{3}$$

where $F : \mathcal{C} \to \mathcal{F}$ is the flat continuous functor corresponding to f, and the vertical maps are induced by the action of f^* on subobjects.

Lemma 1.3 *Let $f : \mathcal{F} \to \mathcal{E} = \mathbf{Sh}(\mathcal{C}, J)$ be a geometric morphism. The functor f^* preserves meets of families of subobjects of cardinality less than κ (resp. all meets, implication, negation) if and only if it does so on subobjects of representables.*

Proof The nontrivial direction follows from the existence of the diagram (3) for each object E in $\mathcal{E}$. □

This allows us to translate properties of f into properties of F and vice versa, for example:

Corollary 1.4 *The geometric morphism f is sub-open in the sense of* [9] *(i.e. f^* preserves implication) if and only if, for each $A \in \mathrm{ob}\, \mathcal{C}$ and every two closed sieves S_1, S_2 on A, the image of the map*

$$\coprod_{\{\alpha : B \to A \mid \alpha^* S_1 = \alpha^* S_2\}} F(B) \longrightarrow F(A)$$

is the subobject $f^* S_1 \Leftrightarrow f^* S_2$ of $F(A)$

Proof Since $S_1 \Leftrightarrow S_2 = \{\alpha \mid \alpha^* S_1 = \alpha^* S_2\}$, the condition says that f^* preserves $\Leftrightarrow$ on the level of the site. But this is equivalent to saying that f^* preserves implication, because in any Heyting algebra $(h \Rightarrow h') \equiv (h \Leftrightarrow (h \wedge h'))$. □

2 First-order theories and syntactic categories

In this section we briefly recall the fundamentals of first-order categorical model theory; for more details we refer the reader to [17], [16], chapter X or [5], appendix B.

We shall work over a (possibly) many-sorted signature Σ specified by giving a set of base sorts (denoted $A, B, C, \ldots$), a set of primitive function-symbols $f, g, h, \ldots$ and a set of primitive relation-symbols $R, S, T, \ldots$; the latter will always include an equality predicate $=_A$ for each sort A. Each function- or relation-symbol has a specified type which is a finite string of sorts; we write $f\colon \vec{A} \to B$ or $R \subseteq \vec{A}$ to indicate the type of a primitive symbol. We identify constants of sort A with nullary function-symbols $[\,] \to A$ (where $[\,]$ denotes the empty string of sorts), and primitive propositions with nullary relation-symbols $R \subseteq [\,]$. A Σ-*structure* M in a category $\mathcal{C}$ with (at least) finite products assigns to each sort A an object MA, to each function-symbol $f\colon \vec{A} \to B$ a morphism $Mf\colon M\vec{A} \to MB$ (where $M\vec{A}$ denotes the product $MA_1 \times \cdots \times MA_n$, if $\vec{A} = (A_1, \ldots, A_n)$), and to each relation-symbol $R \subseteq \vec{A}$ a subobject $MR \rightarrowtail M\vec{A}$ (with the proviso that $M(=_A)$ is necessarily the diagonal subobject of $MA \times MA$).

Our formal language is the infinitary language $\mathcal{L}_{\kappa\omega}(\Sigma)$ (which we shall abbreviate to $\mathcal{L}_\kappa$ when there is no danger of confusion): that is, we allow the formation of infinite conjunctions $\bigwedge_{i \in I} \phi_i$ and disjunctions $\bigvee_{j \in J} \psi_j$ of formulae provided (a) the cardinalities of the index sets I and J are less than κ, and (b) the total number of free variables in any formula remains finite. (Note that we always allow the empty conjunction $\top$ and the empty disjunction $\bot$.) Here κ denotes an infinite regular cardinal (as it will throughout this paper); we allow the possibility that $\kappa = \infty$, the 'cardinality of the universe' (i.e. that we place no restrictions on the size of conjunctions and disjunctions). However, the restriction to finitely many free variables will remain in force throughout. By a *context*, we mean a finite string $\vec{x}$ of distinct variables; we shall write $\phi(\vec{x})$ to indicate that $\vec{x}$ is a *suitable context* for the formula ϕ, i.e. that all the free variables of ϕ appear in it (though it may contain variables which do not appear in ϕ). The *type* of a context $\vec{x}$ is the string $\vec{A}$ of (not necessarily distinct!) sorts of the variables in $\vec{x}$.

By a *first-order theory* $\mathbb{T}$ over Σ, we mean (informally) a set of sentences in $\mathcal{L}_{\infty\omega}(\Sigma)$; however, we shall identify two theories $\mathbb{T}$ and $\mathbb{T}'$ if each member of the first set is deducible from the second, and vice versa. (Here it should be emphasized that our notion of deducibility is relative to an intuitionistic deduction-system such as that in [2]—suitably extended to handle the infinitary connectives, cf. [19] or [13]—since we wish to study models of our theories in categories (such as non-Boolean toposes) where the rules of classical logic are not valid.) One might therefore be tempted to identify each theory with its deductive closure (i.e. the class of all sentences deducible from it); however, it is important for our purposes that the axioms of a theory form a set rather than a proper class. (Note that this implies, in particular, that for each theory $\mathbb{T}$ we can find a cardinal κ such that $\mathbb{T}$ is actually contained in $\mathcal{L}_{\kappa\omega}(\Sigma)$.)

We recall that a formula ϕ is said to be κ-*geometric* if it can be built up from atomic formulae by means of finite conjunctions, disjunctions of cardinality less than κ, and existential quantifications. (If $\kappa = \infty$, we simply say ϕ is *geometric*; if $\kappa = \omega$, we say ϕ is *coherent*. A (κ-)*geometric axiom* is a sentence of the form $(\forall \vec{x})(\phi \Rightarrow \psi)$ where ϕ and ψ are (κ-)geometric formulae and $\vec{x}$ is a suitable context for them; a (κ-)*geometric theory* is one specified by a set $\mathbb{T}$ of (κ-)geometric axioms.

In order to interpret compound formulae in a Σ-structure M, we need to impose further restrictions on the category $\mathcal{C}$ in which M lives. We recall that a *regular category* is one with finite limits and (regular epi/mono) factorizations, the latter being stable under pullback; by a κ-*geometric category* we mean a regular category having unions for arbitrary families of subobjects of cardinality less than κ, the unions being additionally stable under pullback. (Makkai and Reyes [17] call these κ-*logical categories*, and (in the case $\kappa = \omega$) Freyd and Scedrov [5] call them *pre-logoi*.) By a κ-*Heyting category*, we mean a regular category $\mathcal{C}$ having unions and intersections for families of subobjects of cardinality less than κ, and such that the operation of pulling back subobjects of B along a fixed morphism $f\colon A \to B$ of $\mathcal{C}$ has a right adjoint $\forall_f$ (it follows that the unions, as well as the intersections, are automatically stable under pullback; it is also easy to show that in such a category we automatically have an implication operation making each subobject lattice into a Heyting algebra, and that implication is also stable under pullback). Again, this concept has been given different names elsewhere in the literature; in particular, Freyd and Scedrov [5] call it a *logos* in the case $\kappa = \omega$. It is well known that any (Grothendieck) topos is an ∞-Heyting category.

It is clear that, given a Σ-structure M in a κ-geometric (resp. κ-Heyting) category $\mathcal{C}$, we may proceed to interpret any κ-geometric (resp. arbitrary $\mathcal{L}_\kappa$) formula-in-context $\phi(\vec{x})$ over Σ as a subobject $[\phi(\vec{x})]_M$ of $MA_1 \times \cdots \times MA_n$, where $\vec{A} = (A_1, \ldots, A_n)$ is the type of $\vec{x}$. Moreover, this interpretation is *sound* for the logical calculus mentioned above, in the sense that if a sentence σ is deducible from a theory $\mathbb{T}$, then it is satisfied in all $\mathbb{T}$-models in appropriate categories. (As usual, we say that σ is *satisfied* in a structure M if $[\sigma([])]_M \rightarrowtail M[] = 1$ is an isomorphism, and that M is a $\mathbb{T}$-*model* if it satisfies all the sentences in $\mathbb{T}$. For the geometric case, we need to modify the definition of satisfaction: a geometric axiom $\sigma = (\forall \vec{x})(\phi \Rightarrow \psi)$ is satisfied in M if $[\phi(\vec{x})]_M \leq [\psi(\vec{x})]_M$ in $\mathrm{Sub}(M\vec{A})$.)

In order to form a category of $\mathbb{T}$-models in $\mathcal{C}$, where $\mathbb{T}$ is not necessarily geometric, we need to exercise some care in our choice of morphisms. Clearly, given Σ-structures M and N, we should expect a morphism $f \colon M \to N$ to provide a $\mathcal{C}$-morphism $f_A \colon MA \to NA$ for each sort A of Σ. Given this, for each formula-in-context $\phi(\vec{x})$ we shall have a diagram

$$\begin{tikzcd}
{[\phi(\vec{x})]_M} \arrow[d, tail] & & {[\phi(\vec{x})]_N} \arrow[d, tail] \\
M\vec{A} \arrow[rr, "f_{\vec{A}}"] & & N\vec{A}
\end{tikzcd} \tag{4}$$

where $f_{\vec{A}}$ is the product of the f_{A_i}. The usual definition of a (homo)morphism of Σ-structures is equivalent to saying that we have $[\phi(\vec{x})]_M \leq f^*_{\vec{A}}[\phi(\vec{x})]_N$ (i.e. that the above diagram can be completed to a commutative square) for all atomic formulae ϕ; an easy induction then shows that the same condition holds for all geometric formulae ϕ, but it is well known that the induction breaks down for formulae involving implication or universal quantification.

A stronger condition, commonly known by the name *embedding* of Σ-structures, is to demand an equality $[\phi(\vec{x})]_M = f^*_{\vec{A}}[\phi(\vec{x})]_N$ for all atomic ϕ, i.e. to demand that the diagram (4) should be completable to a pullback square. Given that our language always includes the predicate $=_A$ whose interpretation in any M is the diagonal subobject of $MA \times MA$, this condition clearly forces each f_A to be monic. Once again, it is easy to show that this condition is inherited by all quantifier-free formulae ϕ, but it may fail for formulae involving quantifiers.

Classically, a morphism f is called an *elementary embedding* if the diagram (4) is completable to a pullback square for *all* $\mathcal{L}_\omega$ formulae-in-context $\phi(\vec{x})$; it seems reasonable to use the term κ-*elementary embedding* if (4) is a pullback for all $\mathcal{L}_\kappa$ formulae-in-context. (We shall give an example below to demonstrate the dependence of this concept on κ.) However, the fourth corner of the parallelogram, namely the condition that (4) should commute for all ($\mathcal{L}_\kappa$) formulae, is not commonly considered by classical model-theorists. Of course, the reason is that, if we restrict our attention to structures in a Boolean topos such as **Set**, then the commutativity of (4) for the complementary subobjects $[\phi(\vec{x})]$ and $[\neg\phi(\vec{x})]$ forces both the squares to be pullbacks; but this is far from being the case in a non-Boolean topos (again, we shall give an example below), and it turns out that this fourth condition is the one of greatest importance for us.

Morphisms satisfying this condition seem to have been first explicitly considered (in the case $\kappa = \omega$) in [18], where they were called 'weak elementary embeddings'; but this name seems unsatisfactory since such morphisms do not have to be monic. Accordingly, we introduce the name κ-*elementary morphism* of Σ-structures for a family of morphisms f_A such that we have $[\phi(\vec{x})]_M \leq f^*_{\vec{A}}[\phi(\vec{x})]_N$ for all $\mathcal{L}_\kappa$ formulae-in-context $\phi(\vec{x})$ over Σ. We write $\Sigma\text{-Str}(\mathcal{C})_\kappa$ for the category of all Σ-structures in $\mathcal{C}$ and κ-elementary morphisms between them; and, if $\mathbb{T}$ is a theory in $\mathcal{L}_\kappa$, we write $\mathbb{T}\text{-Mod}(\mathcal{C})_\kappa$ for the full subcategory of $\Sigma\text{-Str}(\mathcal{C})_\kappa$ whose objects are $\mathbb{T}$-models. (Note that the subscripts cannot be omitted even in the case $\kappa = \omega$; we use the unsubscripted notations for the categories whose morphisms are arbitrary homomorphisms of Σ-structures.)

Example 2.1 Let $\mathcal{E}$ be the Sierpiński topos whose objects are diagrams $A = (A(0) \to A(1))$ in **Set** and whose morphisms are commutative squares. Let Σ be the one-sorted signature whose only primitive symbol is equality (so that Σ-structures are simply objects). Let N be the terminal object of $\mathcal{E}$, and let M be any object such that $M(1)$ is a singleton. An easy induction shows that the interpretation in N of any $\mathcal{L}_\infty$ formula-in-context over Σ is either the whole of $N^n \cong 1$ or its zero subobject (that is, the intermediate subobject $(0 \to 1)$ cannot occur); so, in order to prove that the unique morphism $M \to N$ is ∞-elementary, it suffices to show that whenever $[\phi(\vec{x})]_N = 0$ we also have $[\phi(\vec{x})]_M = 0$. But if $[\phi(\vec{x})]_N = 0$, then since the functor 'evaluate at 1' is the inverse image of an open geometric morphism $\mathbf{Set} \to \mathcal{E}$ we see that the interpretation of $\phi(\vec{x})$ in the terminal object of $\mathbf{Set}$ is also 0, and hence also $[\phi(\vec{x})]_M(1) = 0$. And this forces $[\phi(\vec{x})]_M(0) = 0$; thus we have shown that $M \to N$ is an ∞-elementary morphism. But it is clearly not an embedding if $M(0)$ has more than one element.

Example 2.2 To show the dependence of κ-elementarity on the cardinal κ, it suffices to consider models in $\mathbf{Set}$. Let Σ be the single-sorted signature with a countable infinity of constant symbols c_i $(i \in \omega)$; let M be the set of natural numbers with c_i interpreted as i for each i, and let N be any uncountable set with each c_i interpreted as a different element of N. It is easy to see that the obvious embedding of Σ-structures $M \to N$ is ω-elementary, since M and N have isomorphic ultrapowers (indeed, for a suitable choice of its cardinality, N is isomorphic to an ultrapower of M); but it is not ω_1-elementary since the $\mathcal{L}_{\omega_1}$ sentence $(\forall x)\bigvee_{i\in\omega}(x = c_i)$ is satisfied in M but not in N.

We next recall briefly the construction of the *syntactic category* $\mathbf{Syn}^{\mathrm{g}}(\mathbb{T})$ of a geometric theory $\mathbb{T}$: for a detailed exposition, see [17], chapter 8 or [16], chapter X. The objects of $\mathbf{Syn}^{\mathrm{g}}(\mathbb{T})$ are equivalence classes $[\phi(\vec{x})]$ of geometric formulae-in-context in the language of $\mathbb{T}$, where the equivalence relation identifies $\phi(\vec{x})$ with $\psi(\vec{y})$ if the contexts $\vec{x}$ and $\vec{y}$ have the same type, and ψ is obtained from ϕ by substituting y_i for x_i, for each i. In defining morphisms $[\phi(\vec{x})] \to [\psi(\vec{y})]$ in $\mathbf{Syn}^{\mathrm{g}}(\mathbb{T})$, we may therefore assume (and it is convenient to do so) that the contexts $\vec{x}$ and $\vec{y}$ have no variables in common; such a morphism is then defined to be a $\mathbb{T}$-provable-equivalence class of geometric formulae in the concatenated context $\vec{x}, \vec{y}$ which are '$\mathbb{T}$-provably functional from ϕ to ψ', in the sense that the appropriate geometric axioms are derivable in $\mathbb{T}$. The composite of two morphisms

$$[\phi(\vec{x})] \xrightarrow{[\theta(\vec{x},\vec{y})]} [\psi(\vec{y})] \xrightarrow{[\gamma(\vec{y},\vec{z})]} [\chi(\vec{z})]$$

is defined to be the provable-equivalence class of the formula $(\exists\vec{y})(\theta \wedge \gamma)$. The verification that $\mathbf{Syn}^{\mathrm{g}}(\mathbb{T})$ is a (geometric) category is tedious but straightforward. Moreover, it contains a structure $U_{\mathbb{T}}$ for the signature of $\mathbb{T}$ obtained by setting $U_{\mathbb{T}}A = [\top(x)]$ (where x is a variable of sort A), $U_{\mathbb{T}}(R \subseteq \vec{A}) = [R(\vec{x})] \rightarrowtail [\top(\vec{x})]$, and $U_{\mathbb{T}}(f\colon \vec{A} \to B) = [(f(\vec{x}) = y)]$. A straightforward induction shows that, for every geometric formula-in-context $\phi(\vec{x})$, the interpretation $[\phi(\vec{x})]_{U_{\mathbb{T}}}$ is simply $[\phi(\vec{x})]$, and hence we have:

Proposition 2.3 *The Σ-structure $U_{\mathbb{T}}$ in $\mathbf{Syn}^{\mathrm{g}}(\mathbb{T})$ is a conservative model of $\mathbb{T}$; that is, it satisfies exactly those geometric axioms which are derivable from $\mathbb{T}$. Moreover, it is universal amongst $\mathbb{T}$-models in geometric categories, in the sense that for any geometric category $\mathcal{C}$ the functor*

$$\mathbf{Geom}\,(\mathbf{Syn}^{\mathrm{g}}(\mathbb{T}), \mathcal{C}) \longrightarrow \mathbb{T}\text{-Mod}(\mathcal{C})$$

which sends a geometric functor $F\colon \mathbf{Syn}^{\mathrm{g}}(\mathbb{T}) \to \mathcal{C}$ to the $\mathbb{T}$-model $F(U_{\mathbb{T}})$, is an equivalence of categories. □

If the theory $\mathbb{T}$ is κ-geometric (for some $\kappa < \infty$), then we may restrict ourselves to κ-geometric formulae in defining the objects and morphisms of the syntactic category, and thus obtain a (small) κ-geometric category $\mathbf{Syn}^{\mathrm{g}}_\kappa(\mathbb{T})$ with the corresponding universal property. However, there is no real need to do so. Although $\mathbf{Syn}^{\mathrm{g}}(\mathbb{T})$ is a large category, there is a canonical form theorem for geometric formulae which says that every such formula is provably equivalent to one of the form $\bigvee_{i\in I}\phi_i$ where the ϕ_i are regular formulae (that is, constructed using only finite conjunction and existential quantification): since we may clearly omit duplicate copies of the same formula from any such disjunction, it follows that we have, up to provable equivalence, only a set of distinct geometric formulae in any context. Since $\mathbb{T}$-provably equivalent formulae-in-context define isomorphic objects of $\mathbf{Syn}^{\mathrm{g}}(\mathbb{T})$, it follows that the latter is (not only locally small but) essentially small, i.e. equivalent to a small category—indeed, it is equivalent to $\mathbf{Syn}^{\mathrm{g}}_\kappa(\mathbb{T})$ for sufficiently large κ.

When working with full first-order logic, we do not have any such canonical form theorem (indeed, as we noted in the Introduction, it is possible to generate a proper class of inequivalent propositional formulae from just two primitive proposition symbols), and so we shall have to be more careful about the size of the formulae which we admit in the construction of our syntactic category. Nevertheless, the purely formal parts of the construction go through without any essential change. If $\mathbb{T}$ is a theory in $\mathcal{L}_\kappa$ (for some $\kappa \leq \infty$), we define the syntactic category $\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T})$, exactly as in the geometric case, but admitting arbitrary formulae of $\mathcal{L}_\kappa$ in the definition of both the objects and the morphisms of the category. Just as before, we may verify

Proposition 2.4 *The category $\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T})$ is a κ-Heyting category (small if $\kappa < \infty$), and contains a conservative $\mathbb{T}$-model $U_\mathbb{T}$ which is universal amongst $\mathbb{T}$-models in κ-Heyting categories, in the sense that for any such category $\mathcal{C}$ the functor*

$$\kappa\text{-}\mathbf{Heyt}\,(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}), \mathcal{C}) \longrightarrow \mathbb{T}\text{-}\mathrm{Mod}(\mathcal{C})_\kappa$$

which sends a κ-Heyting functor $F\colon \mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}) \to \mathcal{C}$ to the $\mathbb{T}$-model $F(U_\mathbb{T})$, is an equivalence of categories. $\square$

For a detailed proof of 2.4 in the case $\kappa = \omega$, see [5], B.314. Note, incidentally, the change in the codomain of the displayed functor of 2.4 as compared with 2.3: we are required to cut down to the κ-elementary morphisms between $\mathbb{T}$-models, in order to ensure that such morphisms correspond to natural transformations between functors defined on the syntactic category.

3 Syntactic sites and the completeness theorem

If $\mathcal{C}$ is a small κ-geometric category (for example, the syntactic category of a κ-geometric theory), we may equip it with the *κ-covering* Grothendieck topology J_κ: a sieve R on an object A of $\mathcal{C}$ belongs to J_κ iff it contains a family of maps $(\alpha_i : B_i \to A \mid i \in I)$ of cardinality less than κ, such that the union of the images of the α_i is the whole of A. It is well known that this topology is subcanonical (that is, all the representable functors $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ are sheaves): moreover, a functor $F\colon \mathcal{C} \to \mathcal{E}$ (where $\mathcal{E}$ is a topos) is flat and J_κ-continuous iff it preserves finite limits, images and unions of families of subobjects of cardinality less than κ—that is, iff it is a κ-geometric functor. Thus, on combining 2.3 with Diaconescu's theorem, we immediately obtain the standard proof of the existence of classifying toposes for κ-geometric theories: given such a theory $\mathbb{T}$, we may take $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$ to be the topos $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{g}}_\kappa(\mathbb{T}), J_\kappa)$. Note also that, since the Yoneda embedding $y\colon \mathbf{Syn}^{\mathrm{g}}_\kappa(\mathbb{T}) \to \mathcal{B}^{\mathrm{g}}(\mathbb{T})$ is full and faithful, the generic $\mathbb{T}$-model $G_\mathbb{T} = y(U_\mathbb{T})$ satisfies only those geometric sentences which are satisfied in $U_\mathbb{T}$ (equivalently, derivable from $\mathbb{T}$).

In passing, we remark that the choice of the cardinal κ is not crucial, provided it is large enough for all the axioms of $\mathbb{T}$ to be expressible in $\mathcal{L}_\kappa$. The canonical form theorem for geometric formulae, mentioned in the last section, ensures that if $\kappa < \lambda$ then every object of $\mathbf{Syn}^{\mathrm{g}}_\lambda(\mathbb{T})$ is the union of a family of subobjects of cardinality less than λ, all of which lie in $\mathbf{Syn}^{\mathrm{g}}_\kappa(\mathbb{T})$. Hence it is easy to verify that the inclusion

$$(\mathbf{Syn}^{\mathrm{g}}_\kappa(\mathbb{T}), J_\kappa) \longrightarrow (\mathbf{Syn}^{\mathrm{g}}_\lambda(\mathbb{T}), J_\lambda)$$

satisfies the hypotheses of the Grothendieck Comparison Lemma (cf. [16], p. 588) and so gives rise to an equivalence of sheaf toposes. Once again, the corresponding result is very far from being true when we replace the 'geometric' versions of the syntactic categories by the 'full first-order' ones.

The key result of this section is the following lemma which, although it has a claim to be considered 'folklore', appears never to have been noticed in print before.

Lemma 3.1 *Let $\mathcal{C}$ be a small Heyting category, and let J be a subcanonical Grothendieck topology on $\mathcal{C}$. Then the Yoneda embedding $y\colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J)$ preserves universal quantification.*

Proof Let $f : A \to B$ be a morphism of $\mathcal{C}$, $A' \rightarrowtail A$ a subobject of A, and write $B' \rightarrowtail B$ for $\forall_f(A' \rightarrowtail A)$. Let $R \rightarrowtail yB$ be any subobject of yB in $\mathbf{Sh}(\mathcal{C}, J)$ such that $(yf)^*(R) \leq yA'$ in $\mathrm{Sub}(yA)$; we must show that $R \leq yB'$, i.e. that every morphism in R factors through $B' \rightarrowtail B$. Let $g : C \to B$ be such a morphism, and let $h\colon D \to A$ be its pullback along f; then $h \in f^*(R)$, and so h factors through $A' \rightarrowtail A$; that is, the image $A'' \rightarrowtail A$ of h satisfies $A'' \leq A'$. Since image factorizations in $\mathcal{C}$ are stable under pullback, it follows that the image $B'' \rightarrowtail B$ of g satisfies $B'' \leq B'$; so we have our required factorization. $\square$

The above proof may be compared with that of the result (which is undoubtedly 'folklore') that the Yoneda embedding of a small cartesian closed category preserves exponentials. Indeed, in the particular case when $\mathcal{C}$ is locally cartesian closed as well as a Heyting category, it follows from the latter, since the functors $\forall_f$ are simply the restrictions of the functors Π_f (the right adjoints to pullback between slice categories) to subobjects. However, the significance of 3.1 itself is contained in the following two immediate corollaries.

Corollary 3.2 *If $\mathcal{C}$ is a small κ-Heyting category and J_κ denotes the κ-covering topology on $\mathcal{C}$, then the Yoneda embedding $\mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J_\kappa)$ is a κ-Heyting functor.*

Proof We have already observed that y is a κ-geometric functor (i.e. that it preserves finite intersections, unions of cardinality less than κ, and images). By the Lemma, it preserves universal quantification (and hence also implication in the subobject lattices); and it clearly preserves all intersections of subobjects which exist in $\mathcal{C}$. $\square$

Corollary 3.3 (Completeness theorem for $\mathcal{L}_{\kappa\omega}$) *Let $\mathbb{T}$ be an arbitrary theory in $\mathcal{L}_{\kappa\omega}(\Sigma)$. Then there exists a Grothendieck topos containing a conservative model of $\mathbb{T}$, i.e. one satisfying only those sentences of $\mathcal{L}_{\kappa\omega}(\Sigma)$ which are derivable from $\mathbb{T}$; moreover, this topos may be taken to be localic over* **Set**. *In particular, any sentence of $\mathcal{L}_{\kappa\omega}(\Sigma)$ which is satisfied in all $\mathbb{T}$-models in localic toposes is derivable from $\mathbb{T}$.*

Proof Apply 3.2 to the syntactic category $\mathbf{Syn}_\kappa^{\mathrm{fo}}(\mathbb{T})$. We have seen that the universal model of $\mathbb{T}$ in this category is conservative, and since the Yoneda embedding is full and faithful it reflects isomorphisms; so $y(U_\mathbb{T})$ is also a conservative model of $\mathbb{T}$. To obtain such a model in a localic topos, we combine this observation with the result ([12], p. 59; [9], Proposition 4.4) that for any Grothendieck topos $\mathcal{E}$ there exists an open surjection $\mathcal{F} \to \mathcal{E}$ with $\mathcal{F}$ localic. $\square$

Since the completeness theorem holds for all $\kappa < \infty$, it follows easily that its weak form holds for $\kappa = \infty$ as well: that is,

Corollary 3.4 *If $\mathbb{T}$ is a first-order theory and σ is any $\mathcal{L}_\infty$ sentence satisfied in all models of $\mathbb{T}$ in localic toposes, then σ is derivable from $\mathbb{T}$.*

Proof Given $\mathbb{T}$ and σ, we can clearly choose κ such that $\mathbb{T} \cup \{\sigma\} \subseteq \mathcal{L}_\kappa$. So the result follows from 3.3. $\square$

On the other hand, the strong form of the theorem (the existence of a single $\mathbb{T}$-model which is $\mathcal{L}_\kappa$-conservative) clearly fails for $\kappa = \infty$; for if $\mathbb{T}$ is the empty theory over a signature Σ with two primitive proposition symbols, then the result of [11] mentioned in the Introduction implies that a topos containing an $\mathcal{L}_\infty$-conservative model of $\mathbb{T}$ would have to have a proper class of subobjects of 1.

In the case $\kappa = \omega$, 3.3 was proved by P. Freyd, around 1975; his proof may be found in [5] under the name 'Stone representation theorem for logoi' (and see also [4], Theorem 5.22, where the theorem is stated in terms of models in categories of Heyting-valued sets). However, Freyd's proof (which admittedly yields a more precise result) is much less direct than the above. The observation that $y(U_\mathbb{T})$ yields a conservative model of $\mathbb{T}$ was made independently by E. Palmgren [20] and the first author.

We note also that $y(U_\mathbb{T})$ is a minimal model of $\mathbb{T}$ in the sense of Moerdijk and Palmgren [18]. Since any κ-Heyting functor is a κ-geometric functor, the equivalence of 2.4 yields a full and faithful functor

$$\mathbb{T}\text{-Mod}(\mathcal{F})_\kappa \longrightarrow \mathbf{Top}\,(\mathcal{F}, \mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(\mathbb{T}), J_\kappa)) \tag{5}$$

for any topos $\mathcal{F}$, sending a $\mathbb{T}$-model M in $\mathcal{F}$ to the geometric morphism whose inverse image extends the κ-Heyting functor $F\colon \mathbf{Syn}_\kappa^{\mathrm{fo}}(\mathbb{T}) \to \mathcal{F}$ corresponding to M. In particular, for any such $\mathbb{T}$-model M, we have a geometric morphism f such that $f^*(y(U_\mathbb{T})) \cong M$. It is easy to see that every open geometric morphism $\mathcal{F} \to \mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(\mathbb{T}))$ is in the image of (5) (since if f is open then the composite $f^* \circ y$ is a κ-Heyting functor $\mathbf{Syn}_\kappa^{\mathrm{fo}}(\mathbb{T}) \to \mathcal{F}$, and so must be the functor which corresponds to the $\mathbb{T}$-model $f^*y(U_\mathbb{T})$). However, it is not the case in general that every $\mathbb{T}$-model in $\mathcal{F}$ can be 'classified' in this way by an open geometric morphism. Nor is it the case that an arbitrary geometric morphism f into $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_{\kappa}(\mathbb{T}))$ is determined up to isomorphism by the Σ-structure $f^* y(U_{\mathbb{T}})$ (recall that a Σ-structure, according to our definition, consists of the objects and morphisms interpreting the primitive symbols of our language), even when the latter is a model of $\mathbb{T}$. We shall investigate this problem further in 4.4–4.6 below.

4 The first-order theory of a topos

Let $\mathcal{E} = \mathbf{Sh}(\mathcal{C}, J)$ be a topos; for convenience, we shall assume in this section that the category $\mathcal{C}$ (is small and) has finite limits, so that flat functors on $\mathcal{C}$ are just finite-limit-preserving ones. The standard proof that $\mathcal{E}$ occurs as a geometric classifying topos $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$ takes $\mathbb{T}$ to be the theory of flat continuous functors on $(\mathcal{C}, J)$, which we may present as follows. The sorts of our signature Σ are the objects of $\mathcal{C}$, and each morphism $f\colon A \to B$ of $\mathcal{C}$ is regarded as a unary function-symbol of Σ; there are no primitive predicate symbols except equality. The axioms of $\mathbb{T}$ fall into three groups: the first group, consisting of all sentences $(\forall x)(1_A(x) = x)$ where 1_A is an identity morphism of $\mathcal{C}$, and $(\forall x)(f(g(x)) = h(x))$ whenever $fg = h$ in $\mathcal{C}$, ensures that $\mathbb{T}$-models in a category $\mathcal{F}$ are in particular functors $\mathcal{C} \to \mathcal{F}$. Then we have a group of axioms to ensure that the functors preserve finite limits: specifically, we take the sentences $(\exists x)\top$ and $(\forall x, y)(x = y)$ where x and y are variables of sort 1 (the terminal object of $\mathcal{C}$), and for each pullback square

$$\begin{array}{ccc} P & \xrightarrow{p} & A \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle f} \\ B & \xrightarrow{g} & C \end{array}$$

in $\mathcal{C}$ we take the axioms $(\forall x, y)((f(x) = g(y)) \Rightarrow (\exists z)((p(z) = x) \wedge (q(z) = y)))$ and $(\forall z, z')(((p(z) = p(z')) \wedge (q(z) = q(z'))) \Rightarrow (z = z'))$, where the sorts of the variables x, y, z are A, B, P respectively. Finally, to ensure that models correspond to continuous (i.e. cover-preserving) functors, we adopt the axiom

$$(\forall x)(\top \Rightarrow \bigvee_{i \in I} (\exists y_i)(f_i(y_i) = x)) \ ,$$

for each family $(f_i\colon B_i \to A \mid i \in I)$ generating a J-covering sieve. (If J can be generated by covering families of cardinalities less than κ, we note that the resulting theory $\mathbb{T}$ is κ-geometric.)

The conditions on a flat continuous F for the corresponding geometric morphism to be open, which we established in 1.2, can readily be translated into (infinitary) first-order sentences in this language. Condition (1) becomes the sentence

$$(\forall x)\Big(\bigwedge_{i \in I}\Big(\bigvee_{j \in J_i} (\exists y_j)(\beta_j(y_j) = x)\Big) \Leftrightarrow \bigvee_{k \in K} (\exists y_k)(\beta_k(y_k) = x)\Big) \tag{6}$$

where $\{\beta_j\colon B_j \to A \mid j \in J_i\}$ is a family of morphisms generating the closed sieve S_i, and $\{\beta_k\colon B_k \to A \mid k \in K\}$ similarly generates $\bigcap_{i \in I} S_i$. And (2) becomes

$$(\forall y)\Big((\forall x)\Big((\alpha(x) = y) \Rightarrow \bigvee_{i \in I} (\exists z_i)(\beta_i(z_i) = x)\Big) \Leftrightarrow \bigvee_{j \in J} (\exists w_j)(\gamma_j(w_j) = y)\Big) \tag{7}$$

where $\{\beta_i\colon C_i \to A \mid i \in I\}$ and $\{\gamma_j\colon D_j \to B \mid j \in J\}$ are generating families for S and $\forall_\alpha S$ respectively. (As we observed after the statement of 1.1, the bi-implications which appear in these two sentences could be replaced by unidirectional implications ($\Rightarrow$), since the reverse implications are automatic.)

We have thus established:

Proposition 4.1 *For any Grothendieck topos $\mathcal{E}$, there is an infinitary first-order theory $\mathbb{T}$ such that we have a natural equivalence*

$$\mathbf{Open}\,(\mathcal{F}, \mathcal{E}) \simeq \mathbb{T}\text{-Mod}(\mathcal{F})$$

for all Grothendieck toposes $\mathcal{F}$. □

Note that, on the right-hand side of the above equivalence, we have the category whose morphisms are arbitrary homomorphisms between $\mathbb{T}$-models, and not just the elementary morphisms: this is because by definition $\mathbf{Open}(\mathcal{F}, \mathcal{E})$ is a full subcategory of $\mathbf{Top}(\mathcal{F}, \mathcal{E})$, and we know that morphisms in the latter correspond to arbitrary homomorphisms of Σ-structures. However, we shall see in a moment that it is at least possible to choose $\mathbb{T}$ in such a way that every homomorphism of $\mathbb{T}$-models is ∞-elementary.

To do this, we consider the case when $\mathcal{E}$ is given as the (geometric) classifying topos of a geometric theory $\mathbb{T}$, i.e. when $(\mathcal{C}, J)$ has the form $(\mathbf{Syn}^{\mathrm{g}}_{\kappa}(\mathbb{T}), J_{\kappa})$. In this case, it would clearly be advantageous to express the first-order theory ($\overline{\mathbb{T}}$, say) corresponding to $\mathcal{E}$ over the same signature as $\mathbb{T}$, instead of passing to the larger signature required to describe functors on the syntactic category. In particular, we should hope that by taking $\overline{\mathbb{T}}$ to consist of *all* first-order sentences which are satisfied in the generic model $G_{\mathbb{T}}$, we might obtain a theory adequate to characterize the open geometric morphisms into $\mathcal{E}$. (Of course, this $\overline{\mathbb{T}}$ is a proper class, and therefore not a theory as we defined it in section 2. However, it is clearly sufficient to choose a cardinal λ which strictly exceeds the sizes of all the subobject lattices $\mathrm{Sub}(G_{\mathbb{T}}(\vec{A}))$, for $\vec{A}$ a finite string of sorts of Σ, and then to take $\overline{\mathbb{T}}$ to consist of all the $\mathcal{L}_{\lambda}$ sentences satisfied by $G_{\mathbb{T}}$. Note that λ may well be strictly larger than the cardinal κ required to express $\mathbb{T}$ as a theory in $\mathcal{L}_{\kappa}$.)

Lemma 4.2 *For the first-order theory $\overline{\mathbb{T}}$ just described, every first-order formula-in-context is provably equivalent to a geometric formula-in-context; that is, given any formula-in-context $\phi(\vec{x})$, we can find a geometric formula ψ in the same context such that $(\forall \vec{x})(\phi \Leftrightarrow \psi)$ is derivable in $\overline{\mathbb{T}}$.*

Proof The interpretation of $\phi(\vec{x})$ in the generic $\mathbb{T}$-model $G_{\mathbb{T}}$ will be some subobject of $G_{\mathbb{T}}(\vec{A})$, where $\vec{A}$ is the type of $\vec{x}$; but every such subobject also occurs as the interpretation of some geometric (indeed, λ-geometric, where λ is as above) formula-in-context $\psi(\vec{x})$. So the sentence in the statement of the lemma is satisfied in $G_{\mathbb{T}}$, and hence derivable in $\overline{\mathbb{T}}$. □

We shall call a theory *geometrically saturated* if it has the property described in the statement of Lemma 4.2. We note that, if $\mathbb{T}$ is geometrically saturated, then any homomorphism of $\mathbb{T}$-models is automatically ∞-elementary, since we already observed that (4) is always completable to a commutative square for geometric formulae. (However, it seems unlikely that this property characterizes geometrically saturated theories, since it also holds when every formula-in-context is provably equivalent to a positive-existential formula—that is, one constructed using infinitary conjunction as well as the 'geometric' connectives.)

Proposition 4.3 *Let $\mathbb{T}$ be a geometric theory. Then, for any topos $\mathcal{F}$, we have*

$$\mathbf{Open}(\mathcal{F}, \mathcal{B}^{\mathrm{g}}(\mathbb{T})) \simeq \overline{\mathbb{T}}\text{-Mod}(\mathcal{F}) = \overline{\mathbb{T}}\text{-Mod}(\mathcal{F})_{\infty} ,$$

where $\overline{\mathbb{T}}$ is the full first-order theory of the generic $\mathbb{T}$-model in $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$.

Proof The equality $\overline{\mathbb{T}}\text{-Mod}(\mathcal{F}) = \overline{\mathbb{T}}\text{-Mod}(\mathcal{F})_{\infty}$ follows from the fact that $\overline{\mathbb{T}}$ is geometrically saturated, as we observed after 4.2. The equivalence with $\mathbf{Open}(\mathcal{F}, \mathcal{B}^{\mathrm{g}}(\mathbb{T}))$ follows from the fact that, since every object and morphism of $\mathbf{Syn}^{\mathrm{g}}(\mathbb{T})$ is (the interpretation in $G_{\mathbb{T}}$ of) a formula-in-context of $\mathcal{L}$, we can rewrite the conditions (6) and (7) as first-order sentences in this language. (The other first-order sentences above, which we used to axiomatize the theory of flat continuous functors on a site, can be similarly rewritten; but there is no need to do so, since they translate into sentences which are derivable from $\mathbb{T}$.) For example, if $\beta\colon B \to A$ is the morphism

$$[\psi(\vec{y})] \xrightarrow{[\theta(\vec{x}, \vec{y})]} [\phi(\vec{x})]$$

of $\mathbf{Syn}^{\mathrm{g}}(\mathbb{T})$, then the clause $(\exists y)(\beta(y) = x)$ (where x and y are variables of sorts A and B) becomes the formula-in-context $(\exists \vec{y})\theta(\vec{x})$. Thus we see that if a $\mathbb{T}$-model in $\mathcal{F}$ satisfies all the first-order sentences satisfied by $G_{\mathbb{T}}$, its classifying map must be open; the converse is immediate. □

We note that, in addition to being geometrically saturated, the theory $\overline{\mathbb{T}}$ of 4.3 is a geometrically conservative extension of $\mathbb{T}$—that is, any geometric axiom derivable from $\overline{\mathbb{T}}$ is already derivable from $\mathbb{T}$—since $G_{\mathbb{T}}$ is a (geometrically) conservative model of $\mathbb{T}$. We shall see in 5.4 below that these two properties suffice to characterize $\overline{\mathbb{T}}$ amongst extensions of $\mathbb{T}$; and in 6.4 we shall give an example of how an explicit description of the classifying topos $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$ may be used to give an explicit axiomatization of the theory $\overline{\mathbb{T}}$.

We also note that the two properties 'geometrically saturated' and 'geometrically conservative' imply that $\overline{\mathbb{T}}$ is maximal amongst geometrically conservative extensions of $\mathbb{T}$; that is, any proper extension of it will prove geometric axioms not derivable from $\mathbb{T}$. However, this property does not suffice to characterize $\overline{\mathbb{T}}$; see 6.2 below.

Finally in this section, we consider what can be said about the first-order theory of a topos which is given to be of the form $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_{\kappa}(\mathbb{T}), J_{\kappa})$ for some first-order theory $\mathbb{T}$ in $\mathcal{L}_{\kappa\omega}(\Sigma)$. In this case, we may again translate conditions (6) and (7) into first-order sentences in $\mathcal{L}_{\lambda\omega}(\Sigma)$ for sufficiently large λ, and thus we obtain:

Proposition 4.4 *Let $\mathbb{T}$ be a theory in $\mathcal{L}_{\kappa\omega}(\Sigma)$. Then there exists a cardinal $\lambda > \kappa$, and a theory $\overline{\mathbb{T}}$ in $\mathcal{L}_{\lambda\omega}(\Sigma)$, such that we have a natural equivalence*

$$\mathbf{Open}\,(\mathcal{F}, \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_{\kappa}(\mathbb{T}), J_{\kappa})) \simeq \overline{\mathbb{T}}\text{-}\mathrm{Mod}(\mathcal{F})_{\infty}$$

for any topos $\mathcal{F}$. Moreover, $\overline{\mathbb{T}}$ is an $\mathcal{L}_{\kappa}$-conservative extension of $\mathbb{T}$, and has the property that every $\mathcal{L}_{\infty\omega}(\Sigma)$ formula-in-context is $\overline{\mathbb{T}}$-provably equivalent to a disjunction of $\mathcal{L}_{\kappa\omega}$ formulae-in-context.

Proof Let γ be the cardinality of the set of morphisms of $\mathbf{Syn}^{\mathrm{fo}}_{\kappa}(\mathbb{T})$. Then it is clear that any closed sieve on an object of this category can be generated by at most γ morphisms, and that there are at most 2^{γ} closed sieves on any such object. Thus the index set of the conjunction appearing in (6) has cardinality at most 2^{γ}, and the disjunctions in both this condition and (7) are indexed by sets of cardinality at most γ. So, if we take λ to be $(2^{\gamma})^{+}$, and take $\overline{\mathbb{T}}$ to consist of all $\mathcal{L}_{\lambda}$ sentences satisfied by $y(U_{\mathbb{T}})$, it will certainly include all instances of these two conditions. Thus we see that, given a $\mathbb{T}$-model M in $\mathcal{F}$, the geometric morphism $f\colon \mathcal{F} \to \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_{\kappa}(\mathbb{T}))$ which corresponds to it under (5) will be open iff M is actually a model of $\overline{\mathbb{T}}$. It follows that any $\mathcal{L}_{\infty}$ sentence satisfied by $y(U_{\mathbb{T}})$ is satisfied in all $\overline{\mathbb{T}}$-models in toposes, and hence (by the $\mathcal{L}_{\infty}$ version of the completeness theorem) derivable from $\overline{\mathbb{T}}$; so the latter really is the full first-order theory of $y(U_{\mathbb{T}})$. The conservativity of $\overline{\mathbb{T}}$ over $\mathbb{T}$ follows from the fact, already noted in 3.3, that $y(U_{\mathbb{T}})$ is an $\mathcal{L}_{\kappa}$-conservative model of $\mathbb{T}$. And the final assertion follows from the fact that every $\mathcal{L}_{\infty}$ formula-in-context over Σ, when interpreted in $y(U_{\mathbb{T}})$, yields a subobject of some object of the site, which can be expressed as a closed sieve on that object and hence as the union of at most γ subobjects corresponding to $\mathcal{L}_{\kappa}$ formulae. (The same argument shows that every κ-elementary morphism between models of $\overline{\mathbb{T}}$ is ∞-elementary, and hence justifies the subscript ∞ on the right-hand side of the displayed equivalence in the statement of the Proposition.) $\square$

Corollary 4.5 *Let $\mathbb{T}$ be a first-order theory. If, for some cardinal κ such that $\mathbb{T}$ is expressible in $\mathcal{L}_{\kappa}$, the theory $\overline{\mathbb{T}}$ of Proposition 4.4 is equivalent to $\mathbb{T}$, then $\mathbb{T}$ has a classifying topos in the sense defined in the Introduction.*

Proof If $\overline{\mathbb{T}}$ is equivalent to $\mathbb{T}$, then every $\mathbb{T}$-model is classified by an open geometric morphism into $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_{\kappa}(\mathbb{T}))$ (that is, the image of (5) consists precisely of the open geometric morphisms); also, the argument at the end of the proof of 4.4 shows that every κ-elementary morphism of $\mathbb{T}$-models is ∞-elementary. So we may take $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_{\kappa}(\mathbb{T}))$ to be $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$. $\square$

The criterion provided by 4.5 for existence of $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$ is not a very practical one, since it is in general very hard even to determine the axioms of $\overline{\mathbb{T}}$, let alone to determine whether they are derivable from those of $\mathbb{T}$. In the next section, we shall provide a criterion (local smallness) which at least has the appearance of being more manageable, and which incidentally enables us to show that the converse of 4.5 is also true.

If $\mathbb{T}$ fails to satisfy the hypothesis of 4.5, it is still of interest to consider what can be said about the relationships between the toposes $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_{\kappa}(\mathbb{T}))$ for different values of κ. For the moment, let us fix a pair of regular cardinals $\kappa < \lambda$ such that $\mathbb{T}$ is expressible in $\mathcal{L}_\kappa$. The inclusion functor $\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}) \to \mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T})$ is clearly a κ-Heyting functor, and hence continuous (i.e. cover-preserving) when these two categories are equipped with the topologies J_κ and J_λ respectively; so it induces a geometric morphism

$$\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T}), J_\lambda) \xrightarrow{r^\lambda_\kappa} \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}), J_\kappa)$$

whose inverse image sends the minimal $\mathbb{T}$-model in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}))$ to that in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T}))$.

Moreover, it is easy to see that the diagram

$$\begin{array}{ccc} \mathbb{T}\text{-Mod}(\mathcal{F})_\lambda & \longrightarrow & \mathbf{Top}\,(\mathcal{F}, \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T}))) \\ \downarrow & & \downarrow \\ \mathbb{T}\text{-Mod}(\mathcal{F})_\kappa & \longrightarrow & \mathbf{Top}\,(\mathcal{F}, \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}))) \end{array} \tag{8}$$

(where the left vertical arrow is the inclusion, the right vertical arrow is induced by composition with r^λ_κ, and the horizontal arrows are instances of (5)) commutes up to natural isomorphism for any $\mathcal{F}$, since both ways round send a $\mathbb{T}$-model in $\mathcal{F}$ to the geometric morphism induced by the κ-Heyting functor $\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}) \to \mathcal{F}$ which corresponds to it. Taking $\mathcal{F} = \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}))$, we see that the geometric morphism $i^\lambda_\kappa \colon \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T})) \to \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T}))$ which corresponds to the minimal $\mathbb{T}$-model in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}))$ satisfies $r^\lambda_\kappa \circ i^\lambda_\kappa \cong 1$; that is, $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}))$ is a retract of $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T}))$. However, composition with i^λ_κ does not in general induce a commutative diagram similar to (8).

Corollary 4.6 *For a first-order theory $\mathbb{T}$ in $\mathcal{L}_\kappa$, the following are equivalent:*

(i) $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}))$ *has the universal property of a classifying topos $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$; that is, for any topos $\mathcal{F}$, the image of the functor (5) is exactly the category of open geometric morphisms $\mathcal{F} \to \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}))$.*

(ii) *For all $\lambda > \kappa$, the geometric morphisms r^λ_κ and i^λ_κ defined above are equivalences.*

(iii) *For all $\lambda > \kappa$, the morphism r^λ_κ is open.*

Proof (i) $\Rightarrow$ (iii) by the definition of $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$.

(iii) $\Rightarrow$ (ii): If r^λ_κ is open, then the composite

$$\mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T}) \longrightarrow \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T})) \xrightarrow{(r^\lambda_\kappa)^*} \mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T}))$$

is a λ-Heyting functor, where the first factor corresponds to the minimal $\mathbb{T}$-model in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}))$; and the composite clearly corresponds to the minimal $\mathbb{T}$-model in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T}))$. It follows that the composite $i^\lambda_\kappa \circ r^\lambda_\kappa$ is isomorphic to the identity; so r^λ_κ and i^λ_κ are inverse equivalences in **Top**.

(ii) $\Rightarrow$ (i): If (ii) holds, then the minimal $\mathbb{T}$-model in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(\mathbb{T}))$ is an $\mathcal{L}_\lambda$-conservative model of $\mathbb{T}$, for all $\lambda > \kappa$. In particular, taking λ to be the cardinal defined in the proof of 4.4, we deduce that $\overline{\mathbb{T}}$ is equivalent to $\mathbb{T}$; so the result follows from 4.5. □

Given a theory $\mathbb{T}$ which has no classifying topos (cf. 6.6 below), we can therefore expect the sequence of toposes $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\lambda(\mathbb{T}))$ to 'grow unboundedly' with the cardinal λ. It is tempting to speculate that it might be possible to embed **Top** in some larger 2-category in which one could compute the 'limit as $\lambda \to \infty$' of these toposes, and hence obtain an object which might support a truly generic $\mathbb{T}$-model, even in the cases when the sequence does not 'stabilize' for sufficiently large λ. But we shall not pursue these speculations here.

5 Locally small theories

Definition 5.1 Let $\mathbb{T}$ be an infinitary first-order theory over a signature Σ. We say $\mathbb{T}$ is *locally small* in a context $\vec{x}$ if there is a set $S_{\vec{x}}$ of formulae in the context $\vec{x}$, such that every $\mathcal{L}_\infty$ formula in this context is $\mathbb{T}$-provably equivalent to a member of $S_{\vec{x}}$. We say $\mathbb{T}$ is *locally small* if it is locally small in every context over Σ.

Note that $\mathbb{T}$ is locally small iff there exists a cardinal κ such that every $\mathcal{L}_\infty$ formula-in-context over Σ is $\mathbb{T}$-provably equivalent to an $\mathcal{L}_\kappa$ formula. Note also that any geometrically saturated theory, as defined after 4.2, is locally small, by the canonical form theorem for geometric formulae. Conversely, we have:

Lemma 5.2 *Let $\mathbb{T}$ be a locally small theory over a signature Σ. Then there exists a signature Σ' extending Σ, and a geometrically saturated theory $\mathbb{T}'$ over Σ', such that $\mathbb{T}'$ is 'Morita-equivalent' to $\mathbb{T}$ in the sense that for every topos $\mathcal{F}$ we have an equivalence*

$$\mathbb{T}\text{-Mod}(\mathcal{F})_\infty \simeq \mathbb{T}'\text{-Mod}(\mathcal{F})_\infty \ ,$$

and these equivalences are natural with respect to open geometric morphisms.

Proof We use (a version of) the technique known to model-theorists as 'Morleyization': we extend the signature Σ by adding, for each context $\vec{x}$ and each member ϕ of $S_{\vec{x}}$, a new relation symbol $R_\phi \rightarrowtail \vec{A}$ (where $\vec{A}$ is the type of $\vec{x}$). For the axioms of $\mathbb{T}'$, we take those of $\mathbb{T}$ together with all sentences of the form

$$(\forall \vec{x})(R_\phi(\vec{x}) \Leftrightarrow \phi) \ .$$

A straightforward induction then shows that every formula-in-context over Σ' is $\mathbb{T}'$-provably equivalent to a formula over Σ, and hence to an atomic formula over Σ', so $\mathbb{T}'$ is geometrically saturated (indeed 'atomically saturated'). On the other hand, given any $\mathbb{T}$-model M in a topos $\mathcal{F}$, there is a unique way of interpreting the additional primitive symbols of Σ' which makes M into a $\mathbb{T}'$-model, and similar remarks apply to ∞-elementary morphisms between such models; so we have the required equivalence $\mathbb{T}\text{-Mod}(\mathcal{F})_\infty \simeq \mathbb{T}'\text{-Mod}(\mathcal{F})_\infty$. □

We are now ready for our main theorem:

Theorem 5.3 *For an infinitary first-order theory $\mathbb{T}$, the following conditions are equivalent:*

(i) $\mathbb{T}$ *is locally small.*

(ii) $\mathbb{T}$ *is Morita-equivalent to a geometrically saturated theory.*

(iii) $\mathbb{T}$ *has a first-order classifying topos; i.e. there exists a topos $\mathcal{B}^{\text{fo}}(\mathbb{T})$ containing a $\mathbb{T}$-model $G_{\mathbb{T}}$ such that, for every topos $\mathcal{F}$, the functor*

$$\mathbf{Open}\,(\mathcal{F}, \mathcal{B}^{\text{fo}}(\mathbb{T})) \longrightarrow \mathbb{T}\text{-Mod}(\mathcal{F})_\infty$$

sending an open geometric morphism f to $f^*(G_{\mathbb{T}})$ is an equivalence of categories.

Proof (i) $\Rightarrow$ (ii) is Lemma 5.2.

(ii) $\Rightarrow$ (iii): It is clear that, if two theories are Morita-equivalent in the sense of 5.2, then one has a first-order classifying topos iff the other does; so we may as well assume that $\mathbb{T}$ itself is geometrically saturated. But, for a geometrically saturated theory $\mathbb{T}$, the inclusion functor $\mathbf{Syn}^{\text{g}}_\infty(\mathbb{T}^{\text{g}}) \to \mathbf{Syn}^{\text{fo}}_\infty(\mathbb{T})$ is an equivalence, where $\mathbb{T}^{\text{g}}$ denotes the geometric part of $\mathbb{T}$, i.e. the set of all geometric axioms (strictly, of κ-geometric axioms for a sufficiently large κ) derivable from $\mathbb{T}$. In particular, for such a κ, the inclusions $\mathbf{Syn}^{\text{fo}}_\kappa(\mathbb{T}) \to \mathbf{Syn}^{\text{fo}}_\lambda(\mathbb{T})$ are equivalences for all $\lambda > \kappa$ (since both categories are equivalent to $\mathbf{Syn}^{\text{g}}_\kappa(\mathbb{T}^{\text{g}})$); so condition (ii) of 4.6 is trivially satisfied.

(iii) $\Rightarrow$ (i): Suppose $\mathcal{B}^{\text{fo}}(\mathbb{T})$ exists. It follows from the $\mathcal{L}_\infty$ version of the completeness theorem (3.4) that the generic $\mathbb{T}$-model $G_{\mathbb{T}}$ in $\mathcal{B}^{\text{fo}}(\mathbb{T})$ must be $\mathcal{L}_\infty$-conservative. Thus the lattice of $\mathbb{T}$-provable-equivalence classes of formulae in a context $\vec{x}$ maps injectively to the subobject lattice $\text{Sub}(G_{\mathbb{T}}(\vec{A}))$, where $\vec{A}$ is the type of $\vec{x}$; hence it is a set. □

Note that it would be possible to give a direct proof of (i) $\Rightarrow$ (iii) by exploiting the fact that, for a locally small theory $\mathbb{T}$, the inclusion $\mathbf{Syn}^{\mathrm{fo}}_{\kappa}(\mathbb{T}) \to \mathbf{Syn}^{\mathrm{fo}}_{\infty}(\mathbb{T})$ is an equivalence for sufficiently large κ. However, the detour via geometrically saturated theories is of interest in its own right (cf. Example 6.3 below). Another benefit of the detour is the following corollary, promised in the last section:

Corollary 5.4 *For any geometric theory $\mathbb{T}$, there is (up to equivalence) exactly one geometrically conservative extension of $\mathbb{T}$ which is geometrically saturated, namely the full first-order theory of the generic model in $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$.*

Proof We saw after 4.3 that the full first-order theory $\overline{\mathbb{T}}$ of the generic $\mathbb{T}$-model has these two properties. Conversely, suppose $\mathbb{S}$ is any extension of $\mathbb{T}$ with these properties. Then $\mathbb{S}$ is locally small, and so has a classifying topos $\mathcal{B}^{\mathrm{fo}}(\mathbb{S})$; and by the proof of (ii) $\Rightarrow$ (iii) in 5.3 we see that in fact $\mathcal{B}^{\mathrm{fo}}(\mathbb{S}) \simeq \mathcal{B}^{\mathrm{g}}(\mathbb{S}^{\mathrm{g}})$, and that this equivalence identifies the generic model of $\mathbb{S}$ with the generic model of $\mathbb{S}^{\mathrm{g}}$. But $\mathbb{S}^{\mathrm{g}}$ is equivalent to $\mathbb{T}$ by conservativity; so all the sentences in $\overline{\mathbb{T}}$ must be derivable from $\mathbb{S}$, and conversely. $\square$

6 Examples and applications

Example 6.1 If $\mathbb{T}$ is a propositional theory (that is, if the signature Σ has no sorts—so that the only primitive symbols are propositions, and the only context is the empty one), then the syntactic category $\mathbf{Syn}^{\mathrm{fo}}_{\infty}(\mathbb{T})$ is simply the complete Heyting algebra generated by the primitive propositions, modulo the filter generated by the propositions in $\mathbb{T}$. Provided this algebra is small, the classifying topos $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$ is simply the topos of sheaves on it (for its canonical topology). For example, it is well known that the free complete Heyting algebra on one generator (unlike that on two generators, as previously mentioned) is a set: it has just one more element than the free Heyting algebra H on one generator, which is pictured on page 35 of [10] (the extra element comes immediately below the top element of H, and is the supremum of all the other elements of H). Equivalently, we may describe the free complete Heyting algebra as the algebra of ideals of H (since all but one of these ideals are principal); so if $\mathbb{T}$ is the empty theory over the propositional signature with a single primitive proposition, we may identify $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$ with the topos of sheaves on H for the finite cover topology. (Note, incidentally, that this particular $\mathbb{T}$ is not itself geometrically saturated.)

Example 6.2 If we consider the empty theory on one propositional generator p as a geometric theory, then its classifying topos $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$ is well known to be the Sierpiński topos $[\mathbf{2}, \mathbf{Set}]$ (cf. Example 2.1), or equivalently the topos of sheaves on the Sierpiński space $\mathbb{S} = \{0, 1\}$ with just one open point. It is easy to see that the full first-order theory of this topos is axiomatized by the single proposition $\neg\neg p$; geometrically, this corresponds to the fact that a continuous map $X \to \mathbb{S}$ is open iff the corresponding open subset of X is dense. From the explicit description of the free complete Heyting algebra on one generator mentioned in the previous example, it may be seen that $\mathbb{T}$ has just two maximal geometrically conservative extensions; one is axiomatized by $\neg\neg p$ and the other by $(p \vee \neg p)$. But the latter is not geometrically saturated, since $\neg p$ is not provably equivalent to a geometric formula.

Example 6.3 Let $\mathbb{T}$ be a geometric theory whose geometric classifying topos $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$ is Boolean. Since every geometric morphism into a Boolean topos is open ([9], 3.5), it follows easily that $\mathcal{B}^{\mathrm{fo}}(\mathbb{T})$ (exists and) coincides with $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$. Conversely, if a geometric theory $\mathbb{T}$ is equivalent to its geometrically saturated extension $\overline{\mathbb{T}}$, as defined in 4.3, then every geometric morphism into $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$ is open, and so the latter must be Boolean. (Thus the phenomenon first observed by Kock [14], that the generic model of a geometric theory may satisfy first-order sentences not derivable from that theory, is typical of all such theories having non-Boolean classifying toposes.)

We recall that in [1] Blass and Scedrov characterized those coherent (that is, ω-geometric) theories whose classifying toposes are Boolean: their characterization involved the conjunction of two conditions, of which the first was a finiteness condition and the second was (the finitary version of) what we have called 'geometric saturatedness'. (Actually, in defining the latter, they used a notion of '$\mathbb{T}$-provable equivalence' relative to a classical logical calculus, rather than the (much more restrictive) constructive provable equivalence which we require. However, if it is already known that $\mathcal{B}^{\mathrm{g}}(\mathbb{T})$ is Boolean, then the two notions of equivalence agree, since each is equivalent to saying that the interpretations of two formulae-in-context coincide in the generic model.) Examples of such theories include the theories of infinite decidable sets, and of dense (trichotomous) linear orderings without endpoints.

Example 6.4 As an example of the explicit computation of the full first-order theory of the generic model of a geometric theory, we consider the theory $\mathbb{O}$ of objects: that is, the empty theory over the single-sorted signature with no primitive symbols except equality. It is well known ([8], 6.33; [16], VIII 4.3) that the object classifier $\mathcal{B}^{\mathrm{g}}(\mathbb{O})$ may be identified with the functor category $[\mathbf{Set}_f, \mathbf{Set}]$, where $\mathbf{Set}_f$ is (a small skeleton of) the category of finite sets, the generic object $G_{\mathbb{O}}$ being the inclusion functor $\mathbf{Set}_f \to \mathbf{Set}$. If one considers what first-order sentences are satisfied by $G_{\mathbb{O}}$, it is not hard for anyone familiar with Kripke–Joyal semantics to verify that they include $(\forall x, y)\neg\neg(x = y)$, $\neg\neg(\exists x)\top$ and

$$(\forall x_1, \ldots, x_n)\neg(\forall y) \bigvee_{i=1}^{n} (y = x_i) \tag{9}$$

for each $n > 0$. (We can think of the first two sentences as saying that $G_{\mathbb{O}}$ is 'not not a singleton', and of (9) as saying that it is 'not not infinite'. Note also that the second axiom may be considered as the case $n = 0$ of (9).)

Many years ago, the second author conjectured that the above sentences might suffice to axiomatize the full first-order theory of $G_{\mathbb{O}}$; but at the time he lacked any means of verifying or disproving this conjecture. With the machinery provided by section 1 of this paper, we are now able to show that the conjecture is false, but that it is in a sense 'very close to the truth'. In the first place, we note that since any monomorphism with nonempty domain in $\mathbf{Set}_f$ is split, a cosieve on a nonempty object of $\mathbf{Set}_f$ is determined by the epimorphisms which it contains; and since there are (up to isomorphism) only finitely many epimorphisms with a given domain in $\mathbf{Set}_f$, it follows that the subobject lattice of each representable functor $\mathbf{Set}_f \to \mathbf{Set}$ is finite. (In the notation introduced before 4.2, we can take the cardinal λ to be ω.) In particular, in determining what it means for a flat functor $\mathbf{Set}_f^{\mathrm{op}} \to \mathcal{F}$ to induce an open geometric morphism, we do not have to worry about condition (1) of 1.2, on infinite intersections of subobjects; all the work is done by condition (2), on universal quantification. Further, since $\mathbf{Set}_f$ is generated as a category by the surjections $(n + 1) \to n$ which identify a single pair of elements and the injections $n \to (n+1)$ which add a single new element, it suffices to consider condition (2) for morphisms of these two forms.

Of course, given an object A of a topos $\mathcal{F}$, the corresponding flat functor $F\colon \mathbf{Set}_f^{\mathrm{op}} \to \mathcal{F}$ is simply the functor $n \mapsto A^n$; equivalently, it sends n to the interpretation $[\top(x_1, \ldots, x_n)]_A$. A cosieve R on n in $\mathbf{Set}_f$, as we observed above, is generated by a finite set of surjections with domain n: each of these is in turn specified by a finite list of pairs of elements of n to be identified, and so we can think of R as the interpretation in $G_{\mathbb{O}}$ of a finite disjunction of finite conjunctions of equations $(x_i = x_j)$; that is, of a positive quantifier-free formula ϕ in the context $(x_1, \ldots, x_n)$.

Now let $\alpha\colon n + 1 \to n$ be the surjection which identifies the last two elements of $n + 1$, and suppose given a cosieve R on n, corresponding to a positive quantifier-free formula ϕ. The cosieve $\forall_\alpha(R)$ on $n + 1$ consists of all those morphisms $n + 1 \to p$ whose pushouts along α lie in R; but we may restrict our attention to surjections, which correspond to finite conjunctions of equations in $(x_1, \ldots, x_n, x_{n+1})$, and the condition for such a conjunction ψ to determine a morphism in $\forall_\alpha(R)$ is simply that

$$(\forall x_1, \ldots, x_{n+1})((\psi \wedge (x_n = x_{n+1})) \Rightarrow \phi) \tag{10}$$

should be provable. On the other hand, if $F\colon \mathbf{Set}_f^{\mathrm{op}} \to \mathcal{F}$ corresponds to an object A of $\mathcal{F}$, then $\forall_{F(\alpha)}(F(R))$ is the largest subobject of A^{n+1} whose intersection with $F(\alpha)\colon A^n \rightarrowtail A^{n+1}$ is contained in $F(R)$; that is, it is the interpretation in A of the formula $((x_n = x_{n+1}) \Rightarrow \phi)$. So in this case condition (2) becomes the validity in A of the sentence

$$(\forall x_1, \ldots, x_{n+1})\Big(((x_n = x_{n+1}) \Rightarrow \phi) \Rightarrow \bigvee \psi\Big)\ , \tag{11}$$

where the disjunction on the right is over all finite conjunctions of atomic formulae ψ satisfying (10). However, since ϕ does not involve x_{n+1}, from $((x_n = x_{n+1}) \Rightarrow \phi)$ we may deduce $((x_n = x_n) \Rightarrow \phi)$

by substitution, and provided ϕ actually involves a variable this reduces to ϕ; so (11) is automatically satisfied except possibly in the cases when ϕ is either $\top$ or $\bot$. For $\phi = \top$ it is again automatic; and for $\phi = \bot$ it simply reduces to $(\forall x_1, \ldots, x_{n+1})(((x_n = x_{n+1}) \Rightarrow \bot) \Rightarrow \bot)$, which is (essentially) the first of the three sentences listed above.

For the injections $n \to n+1$, it is necessary to consider the case $n = 0$ separately, and we deal with this first. There are only two cosieves on 1 in $\mathbf{Set}_f$, and the validity of (2) for the top one is again automatic, so we need only consider the empty cosieve. The universal quantification of this cosieve along $0 \to 1$ is again empty, whereas if we perform the quantification in $\mathcal{F}$ we obtain $[(\forall x)\bot]_A = [\neg(\exists x)\top]_A$, so the condition we obtain is simply the sentence $\neg\neg(\exists x)\top$.

Finally, consider the inclusion $\beta\colon n \to n+1$ $(n > 0)$ and a cosieve R on $n+1$, corresponding to a positive quantifier-free formula ϕ in $(x_1, \ldots, x_{n+1})$. We find that a surjection from n, corresponding to a conjunction of atomic formulae ψ in $(x_1, \ldots, x_n)$, belongs to $\forall_\beta(R)$ iff $(\forall x_1, \ldots, x_{n+1})(\psi \Rightarrow \phi)$ is provable. So condition (2) in this case becomes the sentence $(\forall x_1, \ldots, x_n)((\forall x_{n+1})\phi \Rightarrow \bigvee \psi)$, where the disjunction is over all ψ as above. If ϕ itself does not mention x_{n+1}, this is trivially satisfied; if it does, then from ϕ we may deduce $(\bigvee_{i=1}^{n}(x_{n+1} = x_i) \vee \phi')$, where ϕ' is the disjunction of all those conjunctions of equations in ϕ that do not mention x_{n+1}. So the condition we require is

$$(\forall x_1, \ldots, x_n)\Big((\forall x_{n+1})\Big(\bigvee_{i=1}^{n}(x_{n+1} = x_i) \vee \phi\Big) \Rightarrow \phi\Big) \;, \tag{12}$$

where ϕ is any positive quantifier-free formula in $(x_1, \ldots, x_n)$. Of course, (9) is just the special case of this when $\phi = \bot$; conversely, we could deduce (12) from (9) if we had the (constructively invalid) 'dual Frobenius rule' that $(\forall x)(\chi \vee \phi)$ entails $((\forall x)\chi \vee \phi)$ provided x is not free in ϕ. Thus our axiomatization of the full first-order theory of $G_{\mathbb{O}}$ is exactly that conjectured above, except for the replacement of the scheme (9) by the stronger (12).

We remark that (12) is a genuine strengthening of (9). Let $\mathcal{F} = [\mathbf{N}, \mathbf{Set}]$ where $\mathbf{N}$ is the ordered set of natural numbers; thus objects of $\mathcal{F}$ are diagrams of the form

$$A(0) \xrightarrow{\alpha_0} A(1) \xrightarrow{\alpha_1} A(2) \longrightarrow \cdots .$$

Let A be the object defined by $A(i) = \{0, 1\}$ for all i, and $\alpha_i(0) = \alpha_i(1) = 0$ for all i. Then A satisfies (9) (as well as the other two conditions in our axiomatization), but fails to satisfy (12) for $n = 2$ with ϕ taken to be $(x_1 = x_2)$. Note also that the classifying map $\mathcal{F} \to \mathcal{B}^{\mathrm{g}}(\mathbb{O})$ of A is induced by a functor $\mathbf{N} \to \mathbf{Set}_f$ (namely A itself!); it may be verified that this functor satisfies (the dual of) condition (a) of ([9], 2.5) but not condition (b) or (c), and so the classifying map of A is sub-open but not open. On the other hand, if we modify A by taking each $A(i)$ to be a three-element set, then its classifying map is not even sub-open.

Similar techniques may be used to axiomatize the full first-order theories of other familiar toposes such as the topos of simplicial sets.

Example 6.5 In [18], Moerdijk and Palmgren construct a topos containing a minimal model of Heyting arithmetic (HA). It is easy to see that their construction is exactly that of the topos $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_{\omega}(\mathrm{HA}))$, together with the minimal model $y(U_{\mathrm{HA}})$. On the other hand, the minimal model which they construct for the theory HAI (Heyting arithmetic extended by an external induction scheme for standard natural numbers and an overspill principle), is not simply $y(U_{\mathrm{HAI}})$, since it is not conservative (see [18], 4.10).

Example 6.6 The fact (already mentioned in the Introduction) that the free complete Heyting algebra on two generators is a proper class [11] enables us to show that 'most' familiar theories are not locally small. Given a theory $\mathbb{T}$ and a context $\vec{x}$, let us say that two formulae ϕ and ψ in the context $\vec{x}$ form a *free pair* if, given any topos $\mathcal{E}$ and any two subobjects U, V of the terminal object of $\mathcal{E}$, we can find a $\mathbb{T}$-model M in $\mathcal{E}$, together with an assignment of elements $c_i\colon 1 \to MA_i$ to the variables x_i in $\vec{x}$, such that the interpretations $[\phi(\vec{c})]_M$ and $[\psi(\vec{c})]_M$ are U and V respectively. (For example, if $\mathbb{T}$ is the theory of objects (as in 6.4), then the formulae $(x = y)$ and $(y = z)$ are a free pair in the context (x, y, z), since we can take M to be the colimit of the diagram

$$1 \longleftarrow\!\!\!< U >\!\!\!\longrightarrow 1 \longleftarrow\!\!\!< V >\!\!\!\longrightarrow 1 \;,$$

with the obvious assignment of values to the three variables. The same pair of formulae will be free for any (single-sorted) theory $\mathbb{T}$ with the property that any object of a topos can occur as a subobject of the underlying object of a $\mathbb{T}$-model.) If $\mathbb{T}$ has a free pair of formulae in the context $\vec{x}$, then it cannot be locally small in that context, since by taking $\mathcal{E} = \mathbf{Sh}(H_\kappa)$ (where H_κ is a two-generator complete Heyting algebra of cardinality at least κ, as in the Introduction) and taking U and V to be the generators of H_κ, we see that the propositional combinations of the free pair of formulae will yield at least κ inequivalent formulae in the context $\vec{x}$.

References

[1] A. Blass and A. Scedrov: Boolean classifying topoi, *J. Pure Appl. Alg.* 28 (1983), 15–30.

[2] A. Boileau and A. Joyal: La logique des topos, *J. Symbolic Logic* 46 (1981), 6–16.

[3] R. Diaconescu: Change of base for toposes with generators, *J. Pure Appl. Alg.* 6 (1975), 191–218.

[4] M.P. Fourman and D.S. Scott: Sheaves and logic, in M.P. Fourman, C.J. Mulvey and D.S. Scott, editors, *Applications of Sheaves*, Lecture Notes in Math. 753 (Springer–Verlag, 1979), 302–401.

[5] P.J. Freyd and A. Scedrov: *Categories, Allegories* (North–Holland, 1990).

[6] H. Gaifman: Infinite Boolean polynomials, I, *Fund. Math.* 54 (1964), 229–250.

[7] A.W. Hales: On the non-existence of free complete Boolean algebras, *Fund. Math.* 54 (1964), 45–66.

[8] P.T. Johnstone: *Topos Theory*, L.M.S. Monographs no. 10 (Academic Press, 1977).

[9] P.T. Johnstone: Open maps of toposes, *Manuscripta Math.* 31 (1980), 217–247.

[10] P.T. Johnstone: *Stone Spaces*, Cambridge Studies in Advanced Math. 3 (Cambridge University Press, 1982).

[11] D. de Jongh: A class of intuitionistic connectives, In J. Barwise, H.J. Keisler and K. Kunen, editors, *The Kleene Symposium* (North–Holland, 1980), 103–111.

[12] A. Joyal and M. Tierney: *An Extension of the Galois Theory of Grothendieck*, Mem. Amer. Math. Soc. 309 (American Mathematical Society, 1984).

[13] C. Kalicki: Infinitary propositional intuitionistic logic, *Notre Dame J. Formal Logic* 21 (1980), 216–228.

[14] A. Kock: Universal projective geometry via topos theory, *J. Pure Appl. Alg.* 9 (1976), 1–24.

[15] J. Lambek and P.J. Scott: *Introduction to Higher–Order Categorical Logic*, Cambridge Studies in Advanced Math. no. 7 (Cambridge University Press, 1983).

[16] S. Mac Lane and I. Moerdijk: *Sheaves in Geometry and Logic* (Springer–Verlag, 1992).

[17] M. Makkai and G.E. Reyes: *First Order Categorical Logic*, Lecture Notes in Math. 611 (Springer–Verlag, 1977).

[18] I. Moerdijk and E. Palmgren: Minimal models of Heyting arithmetic, *J. Symbolic Logic*, to appear.

[19] M.E. Nadel: Infinitary intuitionistic logic from a classical point of view, *Ann. Math. Logic* 14 (1978), 159–191.

[20] E. Palmgren: Constructive sheaf semantics, *Math. Logic Quarterly*, to appear.

[21] R.M. Solovay: New proof of a theorem of Gaifmann and Hales, *Bull. Amer. Math. Soc.* 72 (1966), 282–284.