Utrecht University

Master's Thesis

Classifying Topoi and Model Theory

Author:
Mark Kamsma

Supervisor:
Dr. J. van Oosten
Second reader:
Prof. dr. I. Moerdijk

June 15, 2018

Utrecht University

Acknowledgements

I would like to thank my supervisor and tutor, dr. Jaap van Oosten. He has provided me with a lot of invaluable feedback on this thesis. Additionally, he has been a great tutor throughout my graduate studies.

Next I would like to thank my parents for continuously motivating and encouraging me. Not only during this thesis, but throughout my entire study (including primary and secondary school).

My fellow students, Tom de Jong, Jetze Zoethout, Menno de Boer and Sven Bosman, have been a source of inspiration and many interesting conversations. I have really enjoyed discussing the topics of our theses in our seminar.

Finally, I would like to thank my friends for the much needed periods of relaxation, in particular: Tom, Melvin, Lotte, Tom, Martijn and Ivor.

Abstract

Every geometric theory has a classifying topos, but when trying to extend this to full first-order theories one may run into trouble. Such a first-order classifying topos for a first-order theory T, is a topos $\mathcal{F}$ such that for every topos $\mathcal{E}$, the models of T in $\mathcal{E}$ correspond to open geometric morphisms $\mathcal{E} \to \mathcal{F}$. The trouble is that not every first-order theory may have such a first-order classifying topos, as was pointed out by Carsten Butz and Peter Johnstone in [BJ98]. They characterized which theories do admit such a first-order classifying topos, and show how to construct such a first-order classifying topos.

The work of Butz and Johnstone is the main subject of this thesis. The construction of a classifying topos for both geometric theories and first-order theories is worked out in detail. We will also study the characterization of which theories admit a first-order classifying topos. In doing so, we obtain certain completeness results that are interesting in their own right. These are completeness results for deduction-systems for various kinds of infinitary logic, with respect to models in topoi. Building on top of those results, we also obtain a completeness result for classical infinitary logic, with respect to Boolean topoi.

One of the first goals of this thesis was to form a link between Topos Theory and Model Theory, via the first-order classifying topos. In order to bridge the gap between the intuitionistic logic of topoi and the classical logic of Model Theory, we introduce the concept of a Boolean classifying topos. We provide a characterization of which first-order theories admit such a Boolean classifying topos, much like the one for first-order classifying topoi. Then we give a simple example of how to link Boolean classifying topoi to Model Theory, by characterizing complete theories in terms of their Boolean classifying topos.

Contents

Chapter 1

Introduction

In topology there is the notion of a classifying space. For example, the classifying space for cohomology is a space $K(A, n)$ such that for any abelian group A and paracompact space X there is an isomorphism

$$H^n(X, A) \cong [X, K(A, n)],$$

where $[X, K(X, n)]$ denotes the set of homotopy classes of continuous maps $X \to K(A, n)$. In other words, n-dimensional cohomology classes correspond to continuous maps $X \to K(A, n)$, up to homotopy equivalence (this example and some others appear in [MLM92, Section VIII.1]).

In Topos Theory this idea was vastly generalized to the notion of a classifying topos. Besides viewing a topos as a generalized topological space, we can view it as a mathematical universe. That means that internally, a topos contains all kinds of mathematical structures. For example, one can consider the internal groups of a topos. It turns out, that there is a topos $\mathcal{F}$, such that for a suitable notion of morphism of topoi, every internal group of any topos $\mathcal{E}$ corresponds to such a morphism $\mathcal{E} \to \mathcal{F}$, up to isomorphism. This $\mathcal{F}$ is then the classifying topos for groups.

Mathematical structures in general are studied in Model Theory, where we use first-order logic to describe certain structures of interest. This is done by choosing a language and then collecting the necessary axioms in a theory. The structures of interest are then the models of such a theory. Many naturally arising theories (e.g. algebraic theories, like the theory of groups) are so-called geometric theories. That roughly means that their axioms only use conjunction, disjunction, existential quantification and at most one implication symbol. It is well-known that every such geometric theory has a classifying topos. However, problems arise when considering the full first-order setting.

Carsten Butz and Peter Johnstone published a result [BJ98] in 1998 where they defined a suitable notion of first-order classifying topos, and characterized those theories that have such a first-order classifying topos. This thesis is based on that result. To be able to form a link between Model Theory and Topos Theory via classifying topoi, we have the problem that the logic of a topos is generally infinitary and intuitionistic. In order to bridge this gap we develop the notion of a Boolean classifying topos, and we provide a characterization theorem of when such a Boolean classifying topos exists, as well as different constructions of how to obtain such a Boolean classifying topos. At the end we give a simple example of how a model-theoretic property can be characterized in terms of the Boolean classifying topos, and that some model-theoretic properties are not suitable for such a characterization. In particular we will look at the property of being complete and at quantifier elimination, where the former is our example and the latter is the non-example.

The reader is assumed to be familiar with Category Theory and the basics of Topos Theory, Model Theory and Lattice Theory. For completeness and to establish certain conventions, some of the basic notions from those fields will be treated in this thesis, but generally not very extensively. Throughout this entire thesis we assume, and freely use, the Axiom of Choice.

The structure of this thesis is roughly as follows. In chapters 2 and 3 we will treat some basic knowledge concerning Topos Theory, infinitary logic and the internal logic in a category. These are mostly well-known results and constructions, and can be found in for example [MLM92], [Oos16] or [Joh77] (also [Joh02a] and [Joh02b] were used as reference for parts of the material in these chapters).

Then in chapter 4 we will see the precise definition of a classifying topos and of a first-order classifying topos. After that, we show how to perform the usual construction of a classifying topos for a geometric theory in chapters 5 and 6. The machinery we develop in those chapters will prove to be very useful in constructing a first-order classifying topos, which we start with in chapter 7. Then in chapter 9 this culminates to the main theorem of [BJ98], namely Theorem 9.0.7, which characterizes those first-order theories that have a first-order classifying topos. In the proof of this theorem it is shown how to construct such a first-order classifying topos. The first half of this material appears in [MLM92], but our treatment has been slightly changed to be more in line with the rest of the material, which is mainly based on [BJ98].

The contents of chapter 8 are used in the construction of the first-order classifying topos, but they are already interesting in their own right because they are completeness results for various deduction-systems. The completeness results for the systems of geometric logic and intuitionistic logic appear in both [Joh02b] and [BJ98].

Then in chapter 10 we introduce a new kind of classifying topos, that of the Boolean classifying topos, and we characterize which theories have such a Boolean classifying topos. We also provide different ways to construct such a Boolean classifying topos: either starting from the theory itself or by starting from a first-order classifying topos, if it exists. We also generalize a theorem of Andreas Blass and Andrej Ščedrov ([BŠ83, Theorem 1]) that gives necessary and sufficient conditions on coherent theories for their classifying topos to be Boolean. The proof of this generalization heavily realies on the technical tools we developed earlier in that chapter for Boolean classifying topoi.

Then in chapter 11 we establish a simple connection between Model Theory and Topos Theory. To be more precise, we give a characterization of when a certain kind of theory is complete, in terms of the Boolean classifying topos. Also an example is given of why we cannot hope that every model-theoretic property, in this case quantifier elimination, can be characterized in such a way.

In section 8.1 we obtain completeness results for the classical deduction-system for infinitary logic, with respect to Boolean topoi. Although there are a few completeness results for infinitary classical logic, I have not been able to find the results as they appear in section 8.1 anywhere else in literature. Additionally, the results in chapters 10 and 11 do also not seem to appear anywhere else in literature.

Chapter 2

Some Topos Theory

This chapter is meant as a summary of some of the basics of Topos Theory, that we will need later on. The reader is assumed to be familiar with the material in this chapter, so we will not go in depth anywhere. For more elaborate texts on this subject, one can look at any of [MLM92], [Oos16] or [Joh77]. There are also the books [Joh02a] and [Joh02b], but these are probably more useful as reference books than actual introductory texts.

2.1 Sheaves

Throughout this entire text, when we speak of a "topos" (plural: "topoi") we will mean a Grothendieck topos over **Set**. In this section we will make precise what that means. For starters, **Set** denotes the category of sets, as usual. The following definition is the first step towards defining a Grothendieck topos.

Definition 2.1.1. Let $\mathcal{C}$ be any category, then by $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$ we denote the *category of presheaves*. The objects of this category are contravariant functors from $\mathcal{C}$ to **Set**. We call these objects *presheaves*. A morphism of presheaves is then just a natural transformation.

From now on we fix an arbitrary small category $\mathcal{C}$. The category $\mathcal{C}$ can actually be embedded in $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$. Given an object C in $\mathcal{C}$, one can define a presheaf $\mathcal{C}(-, C)$ that assigns to each object D in $\mathcal{C}$ the set $\mathcal{C}(D, C)$ of arrows from D to C. To an arrow $f : D \to D'$ it assigns the operation $\mathcal{C}(f, C) : \mathcal{C}(D', C) \to \mathcal{C}(D, C)$, which sends $g : D' \to C$ to $gf : D \to C$. This construction is used a lot, and has the following name.

Definition 2.1.2. The operation $C \mapsto \mathcal{C}(-, C)$ can be made into a functor $\mathcal{C} \to \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$, by sending an arrow $f : C \to C'$ to the operation $\mathcal{C}(-, f)$ (which is postcomposition with f). This functor is called the *Yoneda embedding*. We denote this functor by y, so $yC = \mathcal{C}(-, C)$. Any presheaf of the form yC for some C is called a *representable presheaf*.

The following is well-known, and proofs of it can be found in practically every book on Category Theory (for example [Oos16, Proposition 2.2]).

Lemma 2.1.3 (Yoneda lemma). *For every presheaf X and object C in $\mathcal{C}$ there is a bijection*

$$\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}(yC, X) \cong X(C),$$

and this is natural in X and C.

The following corollary justifies the name Yoneda *embedding*.

Corollary 2.1.4. *The Yoneda embedding is full and faithful.*

A Grothendieck topos is a special kind of subcategory of the presheaf category on $\mathcal{C}$, namely the category of "sheaves". Which presheaves are considered to be sheaves is determined by a topology on the presheaf category. There are a few equivalent constructions that give such a topology. We follow here the definitions as they appear in [Oos16, Section 10].

Definition 2.1.5. We recall that a *sieve* on an object C of $\mathcal{C}$ is a set of arrows in $\mathcal{C}$ with codomain C, that is closed under precomposition. A *Grothendieck topology* specifies a set of *covering sieves* on C, for each C in $\mathcal{C}$, such that for each object C:

1. the maximal sieve $\max(C)$ is covering;
2. for any covering sieve R on C, and any $f : C' \to C$, $f^*(R) = \{g : C'' \to C' : fg \in R\}$ is a covering sieve on C';
3. let R be any sieve on C and let S be covering for C such that for every $f : C' \to C$ in S, $f^*(R)$ is covering for C', then R is covering for C.

In any category, the subobjects of an object X form a poset. We denote this poset by $\mathrm{Sub}(X)$. If the category has pullbacks, then given any arrow $f : X \to Y$ we obtain an operation $f^* : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$ by pulling a subobject of Y back along f. This f^* is order-preserving, so we can actually view it as a functor when we consider $\mathrm{Sub}(Y)$ and $\mathrm{Sub}(X)$ as categories, and we call it a *pullback functor*. This poset structure allows for the following definition.

Definition 2.1.6. A *universal closure operation* specifies an operation $\bar{(\cdot)} : \mathrm{Sub}(X) \to \mathrm{Sub}(X)$ for every presheaf X, such that for any two subobjects A, B of X:

1. $A \leq \bar{A}$;
2. $\bar{A} = \bar{\bar{A}}$;
3. if $A \leq B$, then $\bar{A} \leq \bar{B}$;
4. for any $f : Y \to X$, we have $f^*(\bar{A}) = \overline{f^*(A)}$.

We will denote the subobject classifier by Ω, and the truth arrow by true. We recall that in $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$, the presheaf Ω assigns to an object C in $\mathcal{C}$ the set of sieves on C and true picks the maximal sieve on C at component C. We will sometimes use the notation $\Omega_{\mathcal{E}}$ and $\mathsf{true}_{\mathcal{E}}$ to make explicit that we mean the subobject classifier in a certain category $\mathcal{E}$.

Definition 2.1.7. A *Lawvere-Tierney topology* is an arrow $J : \Omega \to \Omega$, such that for any C in $\mathcal{C}$ and any two sieves R and S on C:

1. $R \subseteq J_C(R)$;
2. $J_C(R \cap S) = J_C(R) \cap J_C(S)$;
3. $J_C(J_C(R)) = J_C(R)$.

Definition 2.1.8. Let $\mathcal{E}$ be a full subcategory of $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$, and let $i : \mathcal{E} \hookrightarrow \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$ denote the inclusion functor. Then if i has a left adjoint a that preserves finite limits, we call a a *sheafification functor*.

We have seen four different notions now: a Grothendieck topology, a universal closure operation, a Lawvere-Tierney topology and a sheafification functor. It turns out that these are all essentially the same data. That is, we have the following theorem.

Theorem 2.1.9. *Each of the following structures uniquely determines the other three.*

(i) *A Grothendieck topology.*

(ii) *A universal closure operation.*

(iii) *A Lawvere-Tierney topology.*

(iv) *A sheafification functor.*

Proof. The equivalence of the first three appears as [Oos16, Theorem 10.5]. For the equivalence with the sheafification functor one can look at [Oos16, page 107], where it is described how to find the sheafification functor from a Grothendieck topology and how to construct a universal closure operation given a sheafification functor. □

Due to the equivalence of the different structures, we may sometimes just refer to a topology J without explicitly referring to what kind of topology we mean.

Each of the structures in Theorem 2.1.9 gives a way of defining which presheaves are considered to be sheaves. Since they are all equivalent, we will for now just consider the definition for a Grothendieck topology.

Definition 2.1.10. Let X be a presheaf, C be an object in $\mathcal{C}$ and S a sieve on C. A *compatible family* for S is then a family $\{x_f : f \in S\}$ such that $x_f \in X(\mathrm{dom}(f))$, and such that for each $f : C' \to C$ in S and $g : C'' \to C'$ in $\mathcal{C}$ one has $x_{fg} = X(g)(x_f)$.

An *amalgamation* for such a compatible family is an element $x \in X(C)$ such that $x_f = X(f)(x)$ for all $f \in S$. We call X a *sheaf* if every compatible family for a covering sieve has exactly one amalgamation.

Note that the the definition of a compatible family and an amalgamation works for any sieve. The Grothendieck topology comes into play when defining the sheaves, because there we only consider the covering sieves. So to define sheaves we need two pieces of data: the category $\mathcal{C}$ and a Grothendieck topology. This then finally allows us to define what a Grothendieck topos is.

Definition 2.1.11. A pair, $(\mathcal{C}, J)$ where $\mathcal{C}$ is a small category and J is a Grothendieck topology, is called a *site*. A sheaf in $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$ for J is then also called a *sheaf on the site* $(\mathcal{C}, J)$. We denote by $\mathbf{Sh}(\mathcal{C}, J)$ the full subcategory of sheaves on the site $(\mathcal{C}, J)$. A *Grothendieck topos* is then a category of the form $\mathbf{Sh}(\mathcal{C}, J)$, for some site $(\mathcal{C}, J)$.

From now on we will drop the “Grothendieck” from “Grothendieck topos” and just talk about a “topos”. We will use the notation $\mathcal{E}$ and $\mathcal{F}$ for arbitrary topoi.

By definition a site uniquely determines a topos, but this does not work the other way around. Even though a topos $\mathcal{E}$ must be of the form $\mathbf{Sh}(\mathcal{C}, J)$ by definition, the site $(\mathcal{C}, J)$ is by no means unique. There are many sites that give equivalent topoi. In fact, when considering an arbitrary topos $\mathcal{E}$, there are so many different sites that yield $\mathcal{E}$, that we may assume certain properties of such a site. That is, we have the following proposition, which is a consquence of Giraud's theorem and appears for example in [MLM92, Appendix, Corollary 4.2] (in fact, that statement is much stronger, but the statement here is all we need).

Proposition 2.1.12. *Let $\mathcal{E}$ be an arbitrary topos. Then there is a small category $\mathcal{C}$ which has finite limits, with a topology J such that yC is a sheaf for every object C in $\mathcal{C}$. The site $(\mathcal{C}, J)$ is then such that $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, J)$.*

We will not go into further detail on how to define sheaves in terms of a universal closure operation or a Lawvere-Tierney topology. It will however be useful to recall how sheaves and the sheafification functor are related. Consider a site $(\mathcal{C}, J)$, then we have an inclusion $i : \mathbf{Sh}(\mathcal{C}, J) \hookrightarrow \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$. This inclusion has a left adjoint $a : \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}} \to \mathbf{Sh}(\mathcal{C}, J)$ that preserves finite limits and is thus a sheafification functor. This is then exactly the sheafification functor corresponding to J in Theorem 2.1.9. Again, for the details of this construction we refer to [Oos16, page 107].

Suppose we have two inclusions $\mathcal{F} \overset{j}{\hookrightarrow} \mathcal{E} \overset{i}{\hookrightarrow} \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$ such that they both have a sheafification functor as a left adjoint (say: $b \dashv j$ and $a \dashv i$). We can compose the inclusions to an inclusion $ij : \mathcal{F} \hookrightarrow \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$ with corresponding sheafification functor ba. In other words, we obtain a topology such that $\mathcal{F}$ as a topos corresponds to that topology. It therefore makes sense to talk about the topology on an arbitrary topos, instead of just on presheaf category. In fact, the definitions of a Lawvere-Tierney topology and a universal closure operation also make sense in an arbitrary topos instead of just in some presheaf category.

Definition 2.1.13. If a topos $\mathcal{F}$ is equivalent to a topos $\mathcal{F}'$, such that there is an inclusion $\mathcal{F}' \hookrightarrow \mathcal{E}$ with a sheafification functor, we say that $\mathcal{F}$ is a *subtopos* of $\mathcal{E}$.

The sheafification functor also has a special interaction with subobjects, in particular with what are called closed subobjects.

Definition 2.1.14. Let a universal closure operation $\bar{(\cdot)}$ be given. We call a subobject A of X *closed* if $\bar{A} = A$.

The following proposition states that, to find the subsheaves (i.e. subobjects in the category of sheaves) of the sheafification of some object X, we only need to consider its closed subobjects.

Proposition 2.1.15. *The sheafification functor a induces an isomorphism between the closed subobjects of some object X and the subsheaves of $a(X)$, and this isomorphism is natural in X.*

Proof. This is exactly [MLM92, Corollary V.3.8]. □

For the sheafification of representable presheaves Proposition 2.1.15 yields a very explicit description of their subobjects.

Definition 2.1.16. Let $(\mathcal{C}, J)$ be a site, then a sieve S on some object C of $\mathcal{C}$ is called a *closed sieve* if for all $f : D \to C$ we have that:

$$f^*(S) \text{ is covering } \implies f \in S.$$

The terminology can be explained as follows. When we consider a sieve S as a subobject of yC, then S is a closed subobject of yC for a topology J precisely when S is a closed sieve. More details about closed sieves can be found in [MLM92, Section III.7]. In particular, we have the following corollary to Proposition 2.1.15.

Corollary 2.1.17. *The subsheaves of $a(yC)$, the sheafification of a representable presheaf, are precisely the closed sieves on C.*

There is one specific kind of topology that will be useful to us. In terms of a Grothendieck topology we define the following.

Definition 2.1.18. The $\neg\neg$*-topology*, or *double negation topology*, is defined as follows. A sieve S on C is covering if and only if for any $f : D \to C$ there is $g : E \to D$ such that $fg \in S$. We denote this topology by $J_{\neg\neg}$.

The universal closure operator that corresponds under Theorem 2.1.9 to the double negation topology will explain its name, so we will have a look at that. It is well-known that in a topos, the subobject poset of any object is actually a complete Heyting algebra. A proof of this will appear later in section 3.2, where we introduce Heyting categories (Definition 3.2.3), and prove that subobject posets in such categories are Heyting algebras (Proposition 3.2.4). Since any topos is in particular a Heyting category, Proposition 3.2.4 applies to topoi as well.

Proposition 2.1.19. *The operation $\neg\neg(\cdot) : \mathrm{Sub}(X) \to \mathrm{Sub}(X)$ in a topos is a universal closure operation and it is the universal closure operation corresponding under Theorem 2.1.9 to the double negation topology.*

Proof. The fact that this is a universal closure operator, is essentially [MLM92, Theorem VI.1.3], but we will include a proof here anyway. Fix an object X, and let A and B be subobjects of X. We check all four requirements. The numbering below refers to that of Definition 2.1.6.

1. We have $A \leq \neg\neg A$ if and only if $A \wedge \neg A \leq 0$, since $\neg A = A \to 0$. So because $A \wedge \neg A = 0$, we always have $A \leq \neg\neg A$.

2. This is direct since $\neg\neg A = \neg\neg\neg\neg A$.

3. If $A \leq B$, then $\neg B \leq \neg A$ and thus $\neg\neg A \leq \neg\neg B$.

4. Let $f : Y \to X$ be any arrow. We recall that the pullback functor f^* is a homomorphism of Heyting algebras (a proof of this fact also appears in Proposition 3.2.4), so $f^*(\neg\neg A) = \neg\neg f^*(A)$.

The fact that this universal closure operator coincides with the double negation Grothendieck topology from Definition 2.1.18 can be found as [MLM92, Corollary VI.1.5]. □

In particular, we can always look at the double negation sheaves in any topos. So given any topos $\mathcal{E}$, we denote by $\mathbf{Sh}_{\neg\neg}(\mathcal{E})$ the topos of double negation sheaves in $\mathcal{E}$.

The reason why this double negation construction is so interesting to us, is because it guarantees the resulting topos to be Boolean. That is, we have the following definition and proposition.

Definition 2.1.20. A topos $\mathcal{E}$ is called Boolean if all of its subobject posets are actually Boolean algebras.

Proposition 2.1.21. *Any topos of double negation sheaves is Boolean.*

Proof. See for example [MLM92, III.8(21)]. □

A lot more can be said about the double negation topology, see for example [Joh02a, Section A4.5]. We will not discuss these properties here, because we will not need any of them. However, the following proposition concerning Boolean topoi will be useful to us.

Proposition 2.1.22. *A topos $\mathcal{E}$ is Boolean if and only if every subtopos of $\mathcal{E}$ is Boolean.*

Proof. This is a part of [Joh02a, Proposition A4.5.22]. □

2.2 Geometric morphisms

As with any mathematical structure, we will need a notion of morphism of topoi. It turns out that the following definition gives a good notion of such a morphism. For a more detailed treatment and motivation [MLM92, Chapter VII] is recommended.

Definition 2.2.1. A functor is called *left exact* if it preserves all finite limits.

Definition 2.2.2. A *geometric morphism* $f : \mathcal{E} \to \mathcal{F}$ is a pair of adjoints $f^* \dashv f_*$ such that $f_* : \mathcal{E} \to \mathcal{F}$ and $f^* : \mathcal{F} \to \mathcal{E}$, where f^* is left exact. We call f_* the *direct image part* and f^* the *inverse image part*.

Since the composition of geometric morphisms is again a geometric morphism, we have a category of topoi and geometric morphisms.

Definition 2.2.3. The category **Topos** has as objects topoi and as arrows geometric morphisms. With $\mathbf{Topos}(\mathcal{E}, \mathcal{F})$ we will mean the category with as objects geometric morphisms $\mathcal{E} \to \mathcal{F}$ and the arrows $f \to g$ are natural transformations $f^* \to g^*$.

Note that in the above definition we could as well have defined an arrow $f \to g$ as a natural transformation $g_* \to f_*$, since there is a bijective correspondence between these two. However, for us the definition as it is above will be more useful.

Definition 2.2.4. A geometric morphism $f : \mathcal{E} \to \mathcal{F}$ is called an *embedding* if f_* is full and faithful. It is called a *surjection* if f^* is faithful.

So the definition of a sheafification functor (Definition 2.1.8) can also be stated as "a geometric embedding into $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$". Actually, we can just require "a geometric embedding into some topos", because the composition of geometric embeddings is again a geometric embedding.

The following proposition is a well-known general fact about adjoints (see for example [ML71, Theorem IV.3.1]), but it provides a useful characterization of geometric morphisms that are an embedding or surjection.

Proposition 2.2.5. *Let $F : \mathcal{C} \rightleftarrows \mathcal{D} : G$ be adjoints ($F \dashv G$), then the following are true.*

(i) The right adjoint G is full and faithful if and only if the counit $\varepsilon : FG \to Id$ is an isomorphism.

(ii) The left adjoint F is faithful if and only if the unit $\eta : Id \to GF$ is a monomorphism at every object (i.e. $\eta_C : C \to GF(C)$ is a monomorphism for all objects C in $\mathcal{C}$).

Corollary 2.2.6. *Let $f : \mathcal{E} \to \mathcal{F}$ be a geometric morphism, then the following are true.*

(i) Let $\varepsilon : f^* f_* \to Id$ be the counit. Then f is an embedding if and only if ε is an isomorphism.

(ii) Let $\eta : Id \to f_*f^*$ be the unit. Then f is a surjection if and only if $\eta_X : X \to f_*f^*(X)$ is a monomorphism for all objects X in $\mathcal{F}$.

Later on we will see that the inverse image parts of geometric morphisms preserve geometric logic. In general they do not preserve first-order logic. We will be interested in the geometric morphisms that do preserve first-order logic. What it means to preserve geometric logic and first-order logic will be made precise later on. For now let us take a look at the class of open geometric morphisms, which are precisely those geometric morphisms that do preserve first-order logic, as we will see in Proposition 4.0.9. Open geometric morphisms are considered in detail in [Joh80].

We first note that given any left exact functor $F : \mathcal{F} \to \mathcal{E}$ (e.g. the inverse image part of a geometric morphism) we have a map

$$\tau_X : F(\Omega_{\mathcal{F}}^X) \to \Omega_{\mathcal{E}}^{F(X)}$$

for every object X of $\mathcal{F}$, given by the classifying map of the relation $F(\in_X) \rightarrowtail F(\Omega_{\mathcal{F}}^X \times X) \cong F(\Omega_{\mathcal{F}}^X) \times F(X)$.

Definition 2.2.7. We call a geometric morphism $f : \mathcal{E} \to \mathcal{F}$ an *open geometric morphism* if $\tau_X : f^*(\Omega_{\mathcal{F}}^X) \to \Omega_{\mathcal{E}}^{f^*(X)}$ is a monomorphism for each object X in $\mathcal{F}$. We call f *sub-open* if τ_1 is a monomorphism.

The argument from [Joh80, Lemma 1.3] shows us that the composition of open geometric morphisms is again open: let $g : \mathcal{G} \to \mathcal{F}$ and $f : \mathcal{F} \to \mathcal{E}$ be two composable open geometric morphisms. Then

$$\begin{tikzcd} g^*f^*(\Omega_{\mathcal{E}}^X) \arrow[r, "g^*(\tau_X^f)"] \arrow[rr, bend right, "\tau_X^{fg}"] & g^*(\Omega_{\mathcal{F}}^{f^*(X)}) \arrow[r, "\tau_X^g"] & \Omega_{\mathcal{G}}^{g^*f^*(X)} \end{tikzcd}$$

commutes, where the superscript in the τ denotes to which geometric morphism it belongs. Since g^* preserves monomorphisms, we have that τ_X^{fg} is a monomorphism if both τ_X^f and τ_X^g are. We denote by $\mathbf{Open}(\mathcal{E}, \mathcal{F})$ the full subcategory of $\mathbf{Topos}(\mathcal{E}, \mathcal{F})$ with as objects open geometric morphisms.

Some other very useful results about open geometric morphisms can be found in [Joh80]. We mention the following results concerning Boolean topoi, which will be useful later on.

Proposition 2.2.8. *The inclusion $i : \mathbf{Sh}_{\neg\neg}(\mathcal{E}) \to \mathcal{E}$ is sub-open. Furthermore, if $\mathcal{F}$ is any Boolean topos, then a geometric morphism $f : \mathcal{F} \to \mathcal{E}$ is sub-open if and only if it factors through i.*

Proof. Both claims can be found in [Joh80], with the first claim being Corollary 1.8 and the second claim being Proposition 3.6(i). □

Proposition 2.2.9. *The following conditions on a topos $\mathcal{E}$ are equivalent:*

(i) $\mathcal{E}$ is Boolean.

(ii) Every geometric morphism $\mathcal{F} \to \mathcal{E}$ is open.

(iii) Every geometric surjection $\mathcal{F} \to \mathcal{E}$ is sub-open.

Proof. These are exactly cases (i), (ii) and (iv) in [Joh80, Corollary 3.5]. □

Corollary 2.2.10. *Let $f : \mathcal{F} \to \mathcal{E}$ be a sub-open geometric morphism with $\mathcal{F}$ Boolean. Then f factors through the inclusion $i : \mathbf{Sh}_{\neg\neg}(\mathcal{E}) \to \mathcal{E}$, say as $f = ih$ for some $h : \mathcal{F} \to \mathbf{Sh}_{\neg\neg}(\mathcal{E})$, where h is an open geometric morphism.*

Proof. Since f is sub-open, it factors as

$$\mathcal{F} \xrightarrow{h} \mathbf{Sh}_{\neg\neg}(\mathcal{E}) \xrightarrow{i} \mathcal{E}$$

by Proposition 2.2.8. Since $\mathbf{Sh}_{\neg\neg}(\mathcal{E})$ is Boolean, we conclude by Proposition 2.2.9 that h is an open geometric morphism. □

2.3 Diaconescu's theorem

The main goal of this section will be to establish an equivalence between a certain class of functors $\mathcal{C} \to \mathcal{E}$ for a small category $\mathcal{C}$ and geometric morphisms $\mathcal{E} \to \mathbf{Sh}(\mathcal{C}, J)$. This equivalence is known as *Diaconescu's theorem*. A major part of [MLM92, Chapter VII] is devoted to proving this theorem. We will shortly discuss the necessary results here.

Definition 2.3.1. Let $\mathcal{C}$ be a small category and $A : \mathcal{C} \to \mathcal{E}$ be a functor, then we define the *tensor product functor* $- \otimes_{\mathcal{C}} A : \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}} \to \mathcal{E}$ as follows. Let $P : \mathcal{C}^{\mathrm{op}} \to \mathcal{E}$, then $P \otimes_{\mathcal{C}} A$ is defined as the coequalizer

$$\coprod_{\substack{u:C'\to C\\ p\in P(C)}} A(C') \overset{\sigma}{\underset{\theta}{\rightrightarrows}} \coprod_{\substack{C\in\mathcal{C}\\ p\in P(C)}} A(C) \longrightarrow P \otimes_{\mathcal{C}} A.$$

Here θ takes the summand $A(C')$ indexed by (u, p) to $A(C')$ indexed by $(C', P(u)(p))$ via $Id_{A(C')}$, and σ takes the same summand to $A(C)$ via $A(u) : A(C') \to A(C)$ indexed by (C, p).

Even though we will not use this explicit description anywhere, it may be good to see an example of what this actually is (these examples can also be found in [MLM92, page 356–357]).

Example 2.3.2. Suppose that $\mathcal{E} = \mathbf{Set}$ and that $\mathcal{C}$ is a group G. That is, $\mathcal{C}$ has one object and every arrow is an isomorphism. Then a functor $A : \mathcal{C} \to \mathbf{Set}$ is just a set with a left group action, and a presheaf $P : \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ is just a set with a right group action. The coequalizer in Definition 2.3.1 becomes

$$P \times G \times A \overset{\sigma}{\underset{\theta}{\rightrightarrows}} P \times A \longrightarrow P \otimes_G A,$$

where $\theta(p, g, a) = (pg, a)$ and $\sigma(p, g, a) = (p, ga)$. So the elements of $P \otimes_G A$ are equivalence classes of pairs (p, a), where the equivalence relation is generated by $(pg, a) \sim (p, ga)$. In other words: if we write $p \otimes a$ for the equivalence class of (p, a), we have that $P \otimes_G A$ is the set of elements $p \otimes a$ under the equality $pg \otimes a = p \otimes ga$.

This looks a lot like the tensor product of vector spaces, but without additive structure. Even more general, in the above we only make use of the monoid structure on G. So if we view G as the monoid corresponding to the multiplicative structure on a ring R, then this is just the tensor product of a right R-module and left R-module without conditions on the additive structure.

Example 2.3.3. We can generalize Example 2.3.2 slightly to arbitrary small $\mathcal{C}$ (so we are still in the case where $\mathcal{E} = \mathbf{Set}$). The coequalizer then becomes

$$\coprod_{C,C'} P(C) \times \mathcal{C}(C', C) \times A(C') \overset{\sigma}{\underset{\theta}{\rightrightarrows}} \coprod_C P(C) \times A(C) \longrightarrow P \otimes_{\mathcal{C}} A.$$

Now we have the following explicit descriptions for θ and σ: let $(p, u, a) \in P(C) \times \mathcal{C}(C', C) \times A(C')$, then $\theta(p, u, a) = (P(u)(p), a)$ and $\sigma(p, u, a) = (p, A(u)(a))$. For convenience, let us assume that $P(C)$ and $P(C')$ are disjoint for any C and C' in $\mathcal{C}$, and the same for A. Then the elements of $P \otimes_{\mathcal{C}} A$ are equivalence classes of pairs (p, a) where $p \in P(C)$ and $a \in A(C)$ for some C. The equivalence relation is generated by

$$(P(u)(p), a) \sim (p, A(u)(a)), \quad \text{for } u : C' \to C.$$

Theorem 2.3.4. *Let $\mathcal{C}$ be a small category and $A : \mathcal{C} \to \mathcal{E}$ be a functor. Then there are adjoints $L_A \dashv R_A$ with $L_A : \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}} \rightleftarrows \mathcal{E} : R_A$ defined by*

$$L_A(P) = P \otimes_{\mathcal{C}} A,$$
$$R_A(E)(C) = \mathcal{E}(A(C), E).$$

Proof. A proof can be found in [MLM92, Theorem VII.2.1], where its generalized version (as we stated it here) can be found in section 7 of that chapter. ☐

If the reader is familiar with Kan extensions, it may we worth noting that L_A in Theorem 2.3.4 is the left Kan extension of A along the Yoneda embedding, as is noted in [MLM92, page 380].

The following example shows us a familiar case of the adjunction described in Theorem 2.3.4. This example can also be found in [MLM92, page 356].

Example 2.3.5. Let us take $\mathcal{C}$ to be the category with one object and only an identity arrow, and $\mathcal{E} = \mathbf{Set}$. A functor $A : \mathcal{C} \to \mathbf{Set}$ is then simply a set and so is a presheaf $P : \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$. We thus obtain the following tensor product functor:

$$P \times A \overset{\sigma}{\underset{\theta}{\rightrightarrows}} P \times A \longrightarrow P \otimes_{\mathcal{C}} A.$$

Since there is only an identity arrow in $\mathcal{C}$, the maps θ and σ become trivial. That is, $\theta(p, a) = \sigma(p, a) = (p, a)$. So the equivalence relation they generate on $P \times A$ is the trivial one, and $P \otimes_{\mathcal{C}} A$ is just $P \times A$.

Applying this to Theorem 2.3.4, we see that $L_A(P) = P \times A$. Since in this case $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}} \cong \mathbf{Set}$, we may just view R_A as a functor $\mathbf{Set} \to \mathbf{Set}$. So we obtain $R_A(E) = \mathbf{Set}(A, E) = E^A$. So in this case, Theorem 2.3.4 just states that $- \times A \dashv (-)^A$.

With the construction in Theorem 2.3.4 we are already very close to turning a functor $A : \mathcal{C} \to \mathcal{E}$ into a geometric morphism $\mathcal{E} \to \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$. All that we need for this is the left adjoint, that is L_A in Theorem 2.3.4, to be left exact. For this we make the following definition.

Definition 2.3.6. We call a functor $A : \mathcal{C} \to \mathcal{E}$, with $\mathcal{C}$ a small category, *flat* if $- \otimes_{\mathcal{C}} A$ is left exact. We denote by $\mathbf{Flat}(\mathcal{C}, \mathcal{E})$ the category of all flat functors $\mathcal{C} \to \mathcal{E}$ and natural transformations between them.

With these definitions we can now state Diaconescu's theorem.

Theorem 2.3.7 (Diaconescu)**.** *Let $\mathcal{C}$ be a small category, then the geometric morphisms $\mathcal{E} \to \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$ correspond to flat functors $\mathcal{C} \to \mathcal{E}$. That is,*

$$\mathbf{Topos}(\mathcal{E}, \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}) \underset{\sigma}{\overset{\rho}{\rightleftarrows}} \mathbf{Flat}(\mathcal{C}, \mathcal{E})$$

is an equivalence of categories. The explicit description of σ is given by

$$\sigma(A)^* = - \otimes_{\mathcal{C}} A, \quad \sigma(A)_*(E) = \mathcal{E}(A(-), E).$$

The explicit description of ρ is given by $\rho(f) = f^* y$, where y is the Yoneda embedding. Moreover, this equivalence is natural in $\mathcal{E}$.

Proof. A proof can be found as [MLM92, Theorem VII.7.2]. □

If we define a Grothendieck topology J on $\mathcal{C}$, we can even restrict this correspondence to the geometric morphisms into $\mathbf{Sh}(C, J)$.

Definition 2.3.8. A flat functor $A : \mathcal{C} \to \mathcal{E}$ is said to be *continuous* for a Grothendieck topology J, if A sends covering sieves to epimorphic families in $\mathcal{E}$. We will denote by $\mathbf{FlatCon}((\mathcal{C}, J), \mathcal{E})$ the full subcategory of $\mathbf{Flat}(\mathcal{C}, \mathcal{E})$ whose objects are continuous functors on the site $(\mathcal{C}, J)$.

We note that by [MLM92, Lemma VII.7.3] an equivalent definition of a continuous functor is to require that A sends covering sieves to colimits in $\mathcal{E}$.

Corollary 2.3.9 (Diaconescu, continuous version)**.** *The equivalence in Theorem* 2.3.7 *restricts to an equivalence*

$$\mathbf{Topos}(\mathcal{E}, \mathbf{Sh}(\mathcal{C}, J)) \simeq \mathbf{FlatCon}((\mathcal{C}, J), \mathcal{E})$$

which is again natural in $\mathcal{E}$.

Proof. This is exactly [MLM92, Corollary VII.7.4]. □

It can be quite hard to determine whether or not a functor is flat. For this the following lemma will be useful.

Lemma 2.3.10. *Let $\mathcal{C}$ be a small category with finite limits. Then $A : \mathcal{C} \to \mathcal{E}$ is flat if and only if A is left exact.*

Proof. This appears as [MLM92, Corollary VII.9.3]. $\square$

It will be useful to characterize those flat continuous functors that will yield an open geometric morphism. For we this we have [BJ98, Lemma 1.1], which we will now present.

Let $(\mathcal{C}, J)$ be some site, and let $A : \mathcal{C} \to \mathcal{E}$ be a flat and continuous functor. Let $f : \mathcal{E} \to \mathbf{Sh}(\mathcal{C}, J)$ be the corresponding geometric morphism (using Corollary 2.3.9). Furthermore, we denote by

$$i : \mathbf{Sh}(\mathcal{C}, J) \rightleftarrows \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}} : a$$

the inclusion and the sheafification functor.

We recall that the subobject classifier $\Omega_{\mathbf{Sh}(\mathcal{C},J)}$ can be seen as the sheaf that associates to an object X of $\mathcal{C}$ the set of closed sieves on X. That is, $\Omega_{\mathbf{Sh}(\mathcal{C},J)}(X)$ is essentially the set of subobjects of $ay(X)$. Now considering $f_*(\Omega_{\mathcal{E}})$, we have

$$if_*(\Omega_{\mathcal{E}})(X) \cong \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}(yX, if_*(\Omega_{\mathcal{E}})) \cong \mathbf{Sh}(\mathcal{C}, J)(ay(X), f_*(\Omega_{\mathcal{E}})) \cong \mathcal{E}(f^*ay(X), \Omega_{\mathcal{E}}) \cong \mathrm{Sub}_{\mathcal{E}}(f^*ay(X)),$$

and since $A \cong f^*ay$ under the correspondence of Corollary 2.3.9 we can think of $f_*(\Omega_{\mathcal{E}})$ as the sheaf that associates with X the set of subobjects of $A(X)$ in $\mathcal{E}$.

Focusing now on the map $\tau_1 : f^*(\Omega_{\mathbf{Sh}(\mathcal{C},J)}) \to \Omega_{\mathcal{E}}$ from Definition 2.2.7, we let $\bar{\tau} : \Omega_{\mathbf{Sh}(\mathcal{C},J)} \to f_*(\Omega_{\mathcal{E}})$ denote its transpose. Then under the above identifications, $\bar{\tau}_X : \Omega_{\mathbf{Sh}(\mathcal{C},J)}(X) \to f_*(\Omega_{\mathcal{E}})(X)$ sends a subobject Y of $ay(X)$ to the subobject $f^*(Y)$ of $A(X)$. Furthermore, for any arrow $\alpha : X \to Z$ in $\mathcal{C}$ we have actions

$$\begin{aligned} ay(\alpha)^* &: \mathrm{Sub}(ay(Z)) \to \mathrm{Sub}(ay(X)), \\ A(\alpha)^* &: \mathrm{Sub}(A(Z)) \to \mathrm{Sub}(A(X)), \end{aligned}$$

by pulling back the subobjects. Since f^* preserves pullbacks, these fit into a commuting diagram:

$$\begin{tikzcd} \mathrm{Sub}(ay(Z)) \arrow[r, "\bar{\tau}_Z"] \arrow[d, "ay(\alpha)^*"'] & \mathrm{Sub}(A(Z)) \arrow[d, "A(\alpha)^*"] \\ \mathrm{Sub}(ay(X)) \arrow[r, "\bar{\tau}_X"'] & \mathrm{Sub}(A(X)) \end{tikzcd}$$

We also recall that in any topos, the right adjoint of any pullback functor exists. We denote the right adjoint of the pullback functor $A(\alpha)^*$ by $\forall_{A(\alpha)}$ and similar for $ay(\alpha)^*$.

We can now state the lemma characterizing when f is open.

Lemma 2.3.11. *The geometric morphism f is open if and only if the following two properties hold:*

(i) for each object X in $\mathcal{C}$ and each family $\{S_i : i \in I\}$ of subobjects of $ay(X)$ we have

$$\bigwedge_{i \in I} \bar{\tau}_X(S_i) = \bar{\tau}_X(\bigwedge_{i \in I} S_i),$$

(ii) for each arrow $\alpha : X \to Z$ in $\mathcal{C}$ and each subobject S of $ay(X)$ we have

$$\forall_{A(\alpha)}(\bar{\tau}_X(S)) = \bar{\tau}_Z(\forall_{ay(\alpha)}(S)).$$

Proof. By [Joh80, Theorem 3.2(iv)] we have that f is open if and only if $\bar{\tau}$ has an internal left adjoint, so we will show that (i) and (ii) are equivalent to $\bar{\tau}$ having an internal left adjoint. The rest of this proof is from [BJ98, Lemma 1.1].

Fix some object X in $\mathcal{C}$. Condition (i) just states that $\bar{\tau}_X : \mathrm{Sub}(ay(X)) \to \mathrm{Sub}(A(X))$ preserves all limits. Since $\mathrm{Sub}(ay(X))$ and $\mathrm{Sub}(A(X))$ are both complete lattices, we see that condition (i) is equivalent to saying that $\bar{\tau}_X$ has a left adjoint $\lambda_X : \mathrm{Sub}(A(X)) \to \mathrm{Sub}(ay(X))$.

Condition (ii) says that for every arrow $\alpha : X \to Z$ in $\mathcal{C}$ we have a commutative diagram

$$\begin{tikzcd}
\mathrm{Sub}(ay(X)) \arrow[r, "\bar{\tau}_X"] \arrow[d, "\forall_{ay(\alpha)}"'] & \mathrm{Sub}(A(X)) \arrow[d, "\forall_{A(\alpha)}"] \\
\mathrm{Sub}(ay(Z)) \arrow[r, "\bar{\tau}_Z"'] & \mathrm{Sub}(A(Z))
\end{tikzcd}$$

Note that all the arrows here are right adjoints, so moving over to corresponding commuting diagram of left adjoints we find

$$\begin{tikzcd}
\mathrm{Sub}(A(Z)) \arrow[r, "\lambda_Z"] \arrow[d, "A(\alpha)^*"'] & \mathrm{Sub}(ay(Z)) \arrow[d, "ay(\alpha)^*"] \\
\mathrm{Sub}(A(X)) \arrow[r, "\lambda_X"'] & \mathrm{Sub}(ay(X))
\end{tikzcd}$$

So condition (ii) is equivalent to saying that λ is a natural transformation, and thus a morphism of sheaves.

We conclude that conditions (i) and (ii) together are equivalent to $\bar{\tau}$ having an internal left adjoint, which concludes the proof. $\square$

Chapter 3

Infinitary logic

In this chapter we will look at different kinds of infinitary logic and their connection to Category Theory. Most of the definitions and constructions are pretty standard, but we treat them here anyway to be complete and to agree upon notation and conventions. The logic we consider will in general be multi-sorted. Let us start with the definition of a language.

Definition 3.0.12. A *language* $\mathscr{L}$ consists of the following information.

- A set of *sorts*, usually denoted by $S, S_1, S_2, \ldots$. A finite (possibly empty) list of sorts is called a *type*. We usually use X or Y to denote an arbitrary type.
- A set of *function symbols*, usually denoted $f : X \to S$. So each function symbol has a type as domain and a single sort as codomain. Note that we allow the empty type as domain, this will give us *constants*. However, we will never treat constants separately.
- A set of *relation symbols*, usually denoted by R. To each relation symbol a certain type is associated.

A language always has equality. That is, for every sort S there is a relation symbol $=$ of type SS. Furthermore, every sort has its own countably infinite list of variables.

We will be considering formulas in context. That is, we will keep track of which variables can appear in a formula (but they do not have to). For this, we make the following definition.

Definition 3.0.13. A *context* is a finite string x of distinct variables, possibly of different sorts and possibly empty. Thus every context has a type. For two contexts x and y, we denote the concatenation by x, y.

So, unless stated otherwise, x denotes a (possibly empty) list of variables and not just a single variable. It would of course be possible to make an explicit distinction between lists of variables and single variables by using a notation like $\bar{x}$ for lists, and just x for variables. However, we have very few cases where we want to consider a single variable. So in those cases we will just explicitly state that we are only considering a single variable, and keep a cleaner notation in the more general cases.

Definition 3.0.14. Let $\mathscr{L}$ be a language, then the *terms* of $\mathscr{L}$ are defined by induction. To each term we associate a context and a sort. The only requirement for the context is that it contains the free variables appearing in the term (but possibly more). The sort can be thought of as the 'output' of the term. The exact definition is as follows.

- Any variable is a term in any context (containing that variable). The sort of this term is the sort of the variable.
- Let $t_1, \ldots, t_n$ all be terms in the same context x, of sorts $S_1, \ldots, S_n$ respectively and let $f : S_1 \ldots S_n \to S$ be a function symbol, then $f(t_1, \ldots, t_n)$ is a term of sort S in context x.

As mentioned earlier, we will be considering infinitary logic. Usually one sees the notation $\mathscr{L}_{\kappa,\lambda}$ for infinite regular cardinals κ, λ with $\lambda \leq \kappa$. This is then taken to mean "allow disjunctions and conjunctions of size $< \kappa$, and quantifier strings of size $< \lambda$". We will not be interested in the full power of infinitary logic. In particular, we will only be looking at formulas with finite quantifier strings. In other words, we will be looking at $\mathscr{L}_{\kappa,\omega}$, and we will denote this by $\mathscr{L}_\kappa$. This is made precise in the following definition.

Definition 3.0.15. Let κ be a regular infinite cardinal, and let $\mathscr{L}$ be a language. Then the class of $\mathscr{L}_\kappa$-formulas is the smallest class such that the following hold.

- All *atomic formulas* are $\mathscr{L}_\kappa$-formulas. Atomic formulas are those of the following form:
  - $t_1 = t_2$, where t_1 and t_2 are terms of the same sort;
  - $R(t_1, \dots, t_n)$, where R is a relation symbol of type $S_1 \dots S_n$ and each t_i is a term of sort S_i;
  - $\top$ and $\bot$.
- If Φ is a set of $\mathscr{L}_\kappa$-formulas and $|\Phi| < \kappa$, then $\bigwedge \Phi$ and $\bigvee \Phi$ are $\mathscr{L}_\kappa$-formulas if they contain at most finitely many free variables.
- If φ and ψ are $\mathscr{L}_\kappa$-formulas, then $\varphi \to \psi$ is an $\mathscr{L}_\kappa$-formula.
- If φ is an $\mathscr{L}_\kappa$-formula and x is a (single) free variable in φ, then $\exists x \varphi$ and $\forall x \varphi$ are $\mathscr{L}_\kappa$-formulas.

We use the notation $\varphi(x)$ to mean that x is a suitable context for φ. That is, all free variables of φ appear in x. A *formula in context* is a pair consisting of a formula φ and a suitable context x. We denote this by $x.\varphi$.

By κ and λ we will always mean some infinite regular cardinal.

Given two formulas in context $x.\varphi$ and $y.\psi$, where x and y have the same type, we can consider $\psi' = \psi[x/y]$. That is, when we substitute x_i for y_i in ψ, we obtain ψ' which is a formula in context x. We will use the notation $\psi(x)$ for ψ'. Note that this does not violate our earlier convention for this notation, since x is a suitable context for ψ'. One should read it as "we force x to be a suitable context for ψ, even though technically ψ had context y".

If we wish to drop any restrictions on the size of our formulas, we obtain $\mathscr{L}_\infty$. The class of $\mathscr{L}_\infty$-formulas is defined analogous to Definition 3.0.15, but we drop the conditions on the size everywhere. As a convention, whenever we allow constructions of size $< \kappa$, we will allow the notation $\kappa = \infty$. This means that we drop all restrictions on size.

In Definition 3.0.15 we talked about defining a class of formulas. In the case of $\mathscr{L}_\infty$ this is indeed a proper class, because then we can make disjunctions (or conjunctions) of arbitrary length. However, note that when κ is just a cardinal, we actually obtain a set of $\mathscr{L}_\kappa$-formulas. A rigorous proof of this would require us to make precise how exactly we wish to represent our formulas, which can be a bit tedious and does not offer any useful insights for the rest of this text. One can think of it as follows: each formula is essentially a sequence of symbols in $\mathscr{L}$ plus a finite number of logical symbols, thus there is just a set of symbols. These sequences can never be longer than κ. So the $\mathscr{L}_\kappa$-formulas are contained in the collection of sequences of symbols of length $< \kappa$, which is a set.

3.1 Geometric logic

Definition 3.1.1. A *κ-geometric formula* is a formula that is built up using only atomic formulas, finite conjunction, disjunction of size $< \kappa$ and existential quantification. In particular, a κ-geometric formula is always an $\mathscr{L}_\kappa$-formula. If $\kappa = \infty$ we simply call such a formula a *geometric formula*, and if $\kappa = \omega$ we call such a formula a *coherent formula.*

As we see here, we used our convention of allowing $\kappa = \infty$. We even extended the convention in the sense that if we omit κ from the notation, we mean $\kappa = \infty$.

Definition 3.1.2. A *sequent* is an assertion of the form $\varphi \vdash_x \psi$, where φ and ψ are formulas in context x.

One should read a sequent $\varphi \vdash_x \psi$ as "from φ it follows that ψ, for all x". Sometimes we may repeat part of the context in a sequent, to clarify which variables can be expected to appear in that (part of a) formula. For example $\varphi \vdash_{x,y} \psi(x)$ means that ψ, even though considered as a formula in context x, y, will have all its free variables contained in x.

Definition 3.1.3. We call a sequent σ a *κ-geometric sequent* or *κ-geometric axiom* (or *coherent sequent* and *coherent axiom*, in case $\kappa = \omega$) if the formulas appearing on both sides of the sequent are κ-geometric (resp. coherent).

A *κ-geometric theory* is then a set of κ-geometric axioms. In the case $\kappa = \omega$ we obtain the notion of a *coherent theory*.

Any κ-geometric formula or its negation can also be seen as a geometric axiom. For let $\varphi(x)$ be a κ-geometric formula, then $\top \vdash_x \varphi$ is its corresponding κ-geometric axiom. The κ-geometric axiom corresponding to its negation is $\varphi \vdash_x \bot$.

Definition 3.1.4. If we allow any $\mathscr{L}_\kappa$-formula on both sides of a sequent we obtain the notion of a *κ-infinitary first-order axiom*. Again, a set of such κ-infinitary first-order axioms is then called a *κ-infinitary first-order theory*.

In the case for geometric logic we need the notion of a sequent, because we have no way to express "from $\varphi(x)$ it follows that $\psi(x)$, for all x" within the logic itself. First-order logic on the other hand, can do this:

$$\forall x(\varphi(x) \to \psi(x)),$$

or even using

$$\varphi(x) \to \psi(x).$$

The last form makes use of the idea that we can leave out the outermost universal quantifiers. So $\varphi(x)$ will mean the same thing as $\forall x \varphi(x)$, and it will turn out that one holds if and only if the other does, both semantically and proof-theoretically.

It will often be more convenient to think about a κ-infinitary first-order theory as just a set of formulas, instead of sequents. So we adopt the following convention: when we are talking about first-order logic, we may stop talking about sequents and we will just consider formulas.

Just to be clear how the technical translation should be done: a sequent $\varphi \vdash_x \psi$ corresponds to the first-order formula $\varphi \to \psi$, and a first-order formula $\varphi(x)$ corresponds to the sequent $\top \vdash_x \varphi$. From the definitions in sections 3.2 and 3.5 it is easily seen that this translation indeed changes nothing, both semantically and proof-theoretically.

3.2 Interpreting logic in a category

So far we have distinguished three types of logic of increasing strength: coherent logic, κ-geometric logic and full κ-infinitary first-order logic. Each of these logics can be interpreted in a category, if that category satisfies certain conditions. Naturally, we impose more conditions on such a category if we want to interpret stronger forms of logic. Let us first recall what a *regular category* is.

Definition 3.2.1. We call a category $\mathcal{C}$ *regular* if it satisfies the following conditions.

1. $\mathcal{C}$ has all finite limits.
2. The coequalizer of the kernel pair of any morphism $f : X \to Y$ exists. Recall that the *kernel pair* of f is a pair of arrows $p_0, p_1 : Z \to X$ such that

$$\begin{tikzcd} Z \arrow[r, "p_0"] \arrow[d, "p_1"'] & X \arrow[d, "f"] \\ X \arrow[r, "f"'] & Y \end{tikzcd}$$

is a pullback.

3. Regular epimorphisms are stable under pullback.

Let us recall a few properties of regular categories. First of all, the subobject poset $\mathrm{Sub}(X)$ of any object X is a meet-semilattice: the meet of two subobjects is given by pulling one back along the other.

Every arrow has a *regular epi-mono factorization*. That is, for any arrow $f : X \to Y$ there is a factorization

$$X \xrightarrow{e} \mathrm{Im}(f) \xrightarrow{m} Y$$

of f where e is a regular epimorphism, m a monomorphism and $\mathrm{Im}(f)$ is an object called the *image of* f. Moreover, this factorization is unique up to isomorphism.

Given $f : X \to Y$ one can also define $\exists_f : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$ by sending a subobject represented by $m : E \to X$ to $\mathrm{Im}(fm)$. This operation is also order preserving, and as a functor $\exists_f$ is left adjoint to the pullback functor f^*.

A regular category already allows the interpretation of so called regular logic, which only allows formulas constructed using finite conjunction and existential quantification. So this is even weaker than coherent logic. However, all we need for κ-geometric logic is something to interpret disjunction. For this, as we will see, we need joins of subobjects.

Definition 3.2.2. A *κ-geometric category* is a regular category in which the subobject posets have all joins of size $< \kappa$, and these joins are stable under pullback. A *κ-geometric functor* is a left exact functor that preserves regular epimorphisms and joins of size $< \kappa$. The κ-geometric categories and κ-geometric functors then form a category $\mathbf{Geom}_\kappa$.

As before, the case $\kappa = \omega$ gives us the corresponding notions of a *coherent category* and a *coherent functor*.

In the case of $\kappa = \omega$ we will not introduce separate notation for the category of coherent categories and coherent functors. This is just $\mathbf{Geom}_\omega$. In our treatment of interpreting logic in a category we will not separately consider coherent logic, as this is just a special case of our treatment of κ-geometric logic.

In a κ-geometric category the subobject posets also have joins, which means that each subobject poset is actually a lattice. In particular, it is a bounded lattice since the greatest subobject of an object X is just X itself and the empty join exists and thus gives a least element.

Finally, to be able to interpret full κ-infinitary first-order logic, we would also need to be able to interpret the universal quantifier, infinitary meets and implication. For this we make the following definition.

Definition 3.2.3. A *κ-Heyting category* is a κ-geometric category where every pullback functor $f^* : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$ has a right adjoint $\forall_f : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$, and the subobject lattices have meets of size $< \kappa$. A *κ-Heyting functor* is a κ-geometric functor that also preserves the right adjoints $\forall_f$ and the meets of size $< \kappa$. The κ-Heyting categories and κ-Heyting functors then form a category $\mathbf{Heyt}_\kappa$.

In some literature, a coherent category is called a *pre-logos* and an ω-Heyting category is called a *logos*. However, we will stick to the naming in our definitions as it matches the logic we can interpret in it. The reason for the name κ-Heyting category will also be justified by the following proposition.

Proposition 3.2.4. *For any object X in a κ-Heyting category, the subobject lattice $\mathrm{Sub}(X)$ is actually a Heyting algebra. Furthermore, joins and meets of size $< \kappa$ are stable under pullback and the implication operation is also stable under pullback. In particular we have that $f^* : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$ is a Heyting algebra homomorphism for any $f : X \to Y$.*

Proof. We already have that $\mathrm{Sub}(X)$ is a bounded lattice, so all we need is a implication operation. Let A and B be subobjects of X, and let a be a monomorphism into X representing A. We claim that the implication operation $A \to B$ is given by $\forall_a(A \wedge B)$ where $A \wedge B$ is considered as a subobject of A. Then for any subobject C of X we have that $C \leq \forall_a(A \wedge B)$ if and only if $A \wedge C = a^*(C) \leq A \wedge B$ if and only if $A \wedge C \leq B$.

To see that meets of size $< \kappa$ are stable under pullback, we note that the pullback along some $f : X \to Y$ of such a meet is just sending it through $f^* : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$. This operation is left adjoint to $\forall_f$ and thus preserves limits, in particular it preserves meets of size $< \kappa$. Before proving that the implication operation is also stable under pullback, we note that any regular category satisfies the Frobenius law. That is, for $f : X \to Y$, $X' \in \mathrm{Sub}(X)$ and $Y' \in \mathrm{Sub}(Y)$ we have

$$\exists_f(X') \wedge Y' = \exists_f(X' \wedge f^*(Y')).$$

Let now A and B be subobjects of Y, we will show that for any subobject C of X we have $C \leq f^*(A \to B)$ if and only if $C \wedge f^*(A) \leq f^*(B)$, because from that we can conclude that $f^*(A \to B) = f^*(A) \to f^*(B)$. By the construction earlier in this proof, we have that $A \to B = \forall_a(A \wedge B)$, where a is a monomorphism representing A. The proof is now by repeated application of the adjoints $\exists_f \dashv f^*$ and $a^* \dashv \forall_a$ and using the equalities we just discussed:

$$
\begin{array}{rcl}
C \wedge f^*(A) \leq f^*(B) & & \iff \\
C \wedge f^*(A) \leq f^*(A) \wedge f^*(B) = f^*(A \wedge B) & & \iff \\
a^*(\exists_f(C)) = A \wedge \exists_f(C) = \exists_f(f^*(A) \wedge C) \leq A \wedge B & & \iff \\
\exists_f(C) \leq \forall_a(A \wedge B) = A \to B & & \iff \\
C \leq f^*(A \to B) & &
\end{array}
$$

□

We note that any topos is in particular an ∞-Heyting category.

As hinted in the definitions above, a κ-geometric category allows for the interpretation of κ-geometric logic. We will now explain how this is done. Our treatment and notation is based largely on [MLM92, Section X.2]. After that we will explain how to extend this interpretation to full κ-infinitary first-order logic in κ-Heyting categories.

For the rest of this section, fix a language $\mathscr{L}$, a cardinal κ and a κ-geometric category $\mathcal{C}$. To interpret logical formulas in $\mathcal{C}$ we first need an $\mathscr{L}$-structure in $\mathcal{C}$.

Definition 3.2.5. An *$\mathscr{L}$-structure M* in $\mathcal{C}$ consists of the following data.

- For each sort S of $\mathscr{L}$ an object S^M in $\mathcal{C}$. Given a type $X = S_1, \ldots, S_n$, we let X^M be the product $S_1^M \times \ldots \times S_n^M$. In particular, we allow the interpretation of the empty type, which is the empty product and thus the terminal object 1.
- For each function symbol $f : X \to S$, an arrow $f^M : X^M \to S^M$ in $\mathcal{C}$ (where X is a type and S is a sort). Constants are considered function symbols with the empty type as domain. In particular, the interpretation of a constant c of sort S is thus an arrow $c^M : 1 \to S^M$. However, there is no added value in treating constants separately, so we will not mention them explicitly.
- For each relation symbol R of type X, a subobject R^M of X^M in $\mathcal{C}$.

This may sound very similar to the way we define an $\mathscr{L}$-structure in Model Theory. In fact, it is the same, if we take $\mathcal{C}$ to be **Set**. This is also the intuition the reader should keep in mind throughout this section. One can think of the interpretation of a term t of sort S as a function from the type of context of t to S^M. The interpretation of a formula in context x can be seen as those elements that satisfy the formula. Let us make this precise. The interpretation of a term is given by induction on its construction.

Definition 3.2.6. The interpretation t^M of a term t will be an arrow $X^M \to S^M$ in $\mathcal{C}$, where $X = S_1 \ldots S_n$ is the type of its context x and S is the sort of the term.

- If $t = x_i$ (just a single variable in the string x), then t^M is the projection $X^M \to S_i^M$.
- If $t = f(t_1, \ldots, t_k)$, then t^M is the composition $f^M \langle t_1^M, \ldots, t_k^M \rangle$. Here $\langle t_1^M, \ldots, t_k^M \rangle$ is the unique arrow $X^M \to T_1^M \times \ldots \times T_k^M$, determined by the arrows $t_i^M : X^M \to T_i^M$ (where T_i is the sort of the term t_i).

To interpret formulas we make use of the additional properties of a κ-geometric category (up until now, the definitions would work in any category with finite limits). As mentioned before, the interpretation of a formula can thought of as the subset of those elements that satisfy the formula. For this, it is of course necessary to specify where we should be looking for those 'elements' in the first place. This is why we will require the formula to have a context. In other words, we interpret formulas in context.

Definition 3.2.7. Given a κ-geometric formula in context $x.\varphi$ where X is the type of x. Its interpretation $\{x : \varphi(x)\}^M$ is a subobject of X^M, defined inductively on the construction of φ.

- If φ is an atomic formula of the form $t = s$, for terms t and s, then $\{x : t(x) = s(x)\}^M$ is the equalizer of t^M and s^M (which are both arrows $X^M \to S^M$ for some sort S).
- If φ is an atomic formula of the form $R(t_1, \ldots, t_n)$, where each term t_i is of sort T_i, then $\{x : R(t_1(x), \ldots, t_n(x))\}^M$ is the pullback of $R^M \to T_1^M \times \ldots \times T_n^M$ along $\langle t_1^M, \ldots, t_n^M \rangle$. Here $\langle t_1^M, \ldots, t_n^M \rangle$ is the unique arrow into $T_1^M \times \ldots \times T_n^M$, determined by the arrows $t_i^M : X^M \to T_i^M$.
- If φ the atomic formula $\top$, then $\{x : \top\}^M$ is just the maximal subobject of X^M (i.e. X^M itself).
- If φ the atomic formula $\bot$, then $\{x : \bot\}^M$ is just the minimal subobject of X^M.
- If φ is of the form $\psi \wedge \chi$, then $\{x : \psi(x) \wedge \chi(x)\}^M = \{x : \psi(x)\}^M \wedge \{x : \chi(x)\}^M$, where the $\wedge$ on the right side is the meet in the subobject lattice of X^M.
- If φ is of the form $\bigvee_{i \in I} \varphi_i$, then $\left\{x : \bigvee_{i \in I} \varphi_i(x)\right\}^M = \bigvee_{i \in I} \{x : \varphi_i(x)\}^M$, where the $\bigvee$ on the right side is the join in the subobject lattice of X^M.
- If φ is of the form $\exists y \psi(x, y)$, and Y is the type of y. Then there is a projection $\pi : X^M \times Y^M \to X^M$, which gives an operation $\exists_\pi : \mathrm{Sub}(X^M \times Y^M) \to \mathrm{Sub}(X^M)$. We define $\{x : \exists y \psi(x, y)\}^M = \exists_\pi(\{x, y : \psi(x, y)\}^M)$.

We want to emphasize that in general there is no such thing as an 'element' or 'subset', even though we can intuitively think in that way. The reader should keep in mind that the suggestive notation and explanation used here is to help with the intuition, but that at no point we presume the situation to be exactly as in **Set**.

Recall from Definition 3.1.3 that a κ-geometric theory was a set of κ-geometric axioms. We can now define what it means for an $\mathscr{L}$-structure to be a model of such a theory.

Definition 3.2.8. We say that a κ-geometric axiom $\varphi \vdash_x \psi$ is *valid* in an $\mathscr{L}$-structure M if $\{x : \varphi(x)\}^M \leq \{x : \psi(x)\}^M$. We call M a *model* of a κ-geometric theory T, if every sequent in T is valid in M.

As promised, we can also interpret full κ-infinitary first-order logic if we require $\mathcal{C}$ to be a κ-Heyting category. The definition of an $\mathscr{L}$-structure remains the same, and we extend Definition 3.2.7 to also interpret the connectives for implication, infinite conjunction and the universal quantifier. Here we make use of Proposition 3.2.4, namely that every subobject lattice in a κ-Heyting category is a Heyting algebra.

Definition 3.2.9. Given an $\mathscr{L}_\kappa$-formula in context $x.\varphi$ where X is the type of x. We extend Definition 3.2.7 for its interpretation $\{x : \varphi(x)\}^M$ by the following clauses.

- If φ is of the form $\psi \to \chi$, then $\{x : \psi(x) \to \chi(x)\}^M = \{x : \psi(x)\}^M \to \{x : \chi(x)\}^M$, where the $\to$ on the right side is the Heyting implication in the subobject lattice of X^M.
- If φ is of the form $\bigwedge_{i \in I} \varphi_i$, then $\left\{x : \bigwedge_{i \in I} \varphi_i(x)\right\}^M = \bigwedge_{i \in I} \{x : \varphi_i(x)\}^M$, where the $\bigwedge$ on the right side is the meet in the subobject lattice of X^M.
- If φ is of the form $\forall y \psi(x, y)$, and Y is the type of y. Then there is a projection $\pi : X^M \times Y^M \to X^M$, which gives an operation $\forall_\pi : \mathrm{Sub}(X^M \times Y^M) \to \mathrm{Sub}(X^M)$. We define $\{x : \forall y \psi(x, y)\}^M = \forall_\pi(\{x, y : \psi(x, y)\}^M)$.

It is clear how Definition 3.2.8 extends to any κ-infinitary first-order axiom. This then also defines when an $\mathscr{L}$-structure is a model of a κ-infinitary first-order theory.

3.3 The category of models

Now that we can speak of internal structures in a category $\mathcal{C}$, we would also like to describe what an internal homomorphism of such structures would be. For this we make the definition below, following [BJ98].

Definition 3.3.1. Let M and N be $\mathscr{L}$-structures in $\mathcal{C}$. Suppose we have a family $h_S : S^M \to S^N$ of arrows in $\mathcal{C}$ for each sort S in $\mathscr{L}$. For a type $X = S_1 \dots S_n$ we set $h_X = h_{S_1} \times \dots \times h_{S_n}$. We call this family a *homomorphism* from M to N if for any atomic formula $\varphi(x)$ the following diagram can be completed to a commutative square:

$$
\begin{tikzcd}
\{x : \varphi(x)\}^M \arrow[d, rightarrowtail] & \{x : \varphi(x)\}^N \arrow[d, rightarrowtail] \\
X^M \arrow[r, "h_X"] & X^N
\end{tikzcd}
$$

Note that Definition 3.3.1 is equivalent to saying that $\{x : \varphi(x)\}^M \leq h_X^*(\{x : \varphi(x)\}^N)$.

By induction we see that the condition $\{x : \varphi(x)\}^M \leq h_X^*(\{x : \varphi(x)\}^N)$ actually holds for any geometric formula $\varphi(x)$, if h is a homomorphism. However, this may break down for κ-infinitary first-order formulas. So, again following [BJ98], we introduce the following notion.

Definition 3.3.2. A homomorphism h is called a *κ-elementary morphism* if

$$\{x : \varphi(x)\}^M \leq h_X^*(\{x : \varphi(x)\}^N)$$

holds for all $\mathscr{L}_\kappa$-formulas $\varphi(x)$.

We will see κ-elementary morphisms pop up at some later point, when we see that they are related to natural transformations of open geometric morphisms.

Homomorphisms can be composed to a new homomorphism by composing their arrows for each sort. The same can be done for κ-elementary embeddings. In this way we obtain the notion of a category of $\mathscr{L}$-structures in $\mathcal{C}$.

Definition 3.3.3. We define the *category of structures* in $\mathcal{C}$ to be $\mathscr{L}$–$\mathbf{Str}(\mathcal{C})$, with as objects the $\mathscr{L}$-structures in $\mathcal{C}$ and as arrows the homomorphisms of $\mathscr{L}$-structures. We denote by $\mathscr{L}$–$\mathbf{Str}(\mathcal{C})_\kappa$ the subcategory of $\mathscr{L}$–$\mathbf{Str}(\mathcal{C})$ with as arrows κ-elementary morphisms.

The *category of models* in $\mathcal{C}$ of a theory T is the full subcategory T–$\mathbf{Mod}(\mathcal{C})$ of $\mathscr{L}$–$\mathbf{Str}(\mathcal{C})$ whose objects are models of T. Cutting down to κ-elementary morphisms we obtain the full subcategory T–$\mathbf{Mod}(\mathcal{C})_\kappa$ of $\mathscr{L}$–$\mathbf{Str}(\mathcal{C})_\kappa$.

Note that in this definition we have violated our convention a little bit: $\mathscr{L}$–$\mathbf{Str}(\mathcal{C})$ and $\mathscr{L}$–$\mathbf{Str}(\mathcal{C})_\infty$ are two different categories! The same thing goes for T–$\mathbf{Mod}(\mathcal{C})$ and T–$\mathbf{Mod}(\mathcal{C})_\infty$. So in this notation, no subscript will always mean "with all homomorphisms" and a subscript κ will always mean "only with κ-elementary morphisms".

From the definition of $\mathscr{L}$-structures internal in a category $\mathcal{C}$ (Definition 3.2.5) it follows directly that left exact functors send $\mathscr{L}$-structures to $\mathscr{L}$-structures. That is, given an $\mathscr{L}$-structure M in $\mathcal{C}$ and a left exact functor $F : \mathcal{C} \to \mathcal{D}$, then F sends M to an $\mathscr{L}$-structure $F(M)$ in $\mathcal{D}$. Since all κ-geometric functors and κ-Heyting functors are in particular left exact, we can now formulate the following two propositions.

Proposition 3.3.4. *All κ-geometric functors preserve κ-geometric logic, in the following sense. Let $\mathcal{C}$ and $\mathcal{D}$ be κ-geometric categories, and let $F : \mathcal{C} \to \mathcal{D}$ be a κ-geometric functor. If $\varphi(x)$ is a κ-geometric formula in some language $\mathscr{L}$ and M is an $\mathscr{L}$-structure in $\mathcal{C}$, then*

$$F(\{x : \varphi(x)\}^M) = \{x : \varphi(x)\}^{F(M)}.$$

Proof. The argument is by induction on the construction of $\varphi(x)$. For atomic formulas (except for $\perp$) and the connective $\wedge$ the statement is true because F is left exact. For disjunctions of size $< \kappa$ (so this includes the empty disjunction $\perp$) the statement is true since F preserves joins of size $< \kappa$.

All that remains is the existential quantifier. For this we recall that for an arrow $f : X \to Y$ and a subobject of X represented by $m : A \to X$, the subobject $\exists_f(A)$ is given by $\mathrm{Im}(fm)$. Since F preserves regular epimorphisms, we have that $F(A) \to F(\mathrm{Im}(fm)) \to F(Y)$ is an regular epi-mono factorization of $F(fm)$, so $\mathrm{Im}(F(fm)) \cong F(\mathrm{Im}(fm))$.

Suppose that $\varphi(x)$ is of the form $\exists y \psi(x, y)$. Denote by $\pi_X^{F(M)} : X^{F(M)} \times Y^{F(M)} \to X^{F(M)}$ and $\pi_X^M : X^M \times Y^M \to X^M$ the projections. Note that since F is left exact, we have that $F(\pi_X^M) = \pi_X^{F(M)}$.

If we slightly abuse notation and also denote $\{x, y : \psi(x, y)\}^M$ for the monomorphism representing the subobject $\{x, y : \psi(x, y)\}^M$, we indeed find

$$\begin{aligned}
\{x : \varphi(x)\}^{F(M)} &= \exists_{\pi_X^{F(M)}}(\{x, y : \psi(x, y)\}^{F(M)}) && \text{(definition for the interpretation of } \exists) \\
&= \exists_{F(\pi_X^M)}(F(\{x, y : \psi(x, y)\}^M)) && \text{(induction hypothesis)} \\
&= \mathrm{Im}(F(\pi_X^M)F(\{x, y : \psi(x, y)\}^M)) && \text{(definition of } \exists_{F(\pi_X^M)}) \\
&= \mathrm{Im}(F(\pi_X^M\{x, y : \psi(x, y)\}^M)) && \\
&= F(\mathrm{Im}(\pi_X^M\{x, y : \psi(x, y)\}^M)) && \text{(by the discussion above)} \\
&= F(\exists_{\pi_X^M}(\{x, y : \psi(x, y)\}^M)) && \text{(definition of } \exists_{\pi_X^M}) \\
&= F(\{x : \varphi(x)\}^M). && \text{(definition for the interpretation of } \exists)
\end{aligned}$$

□

Proposition 3.3.5. *All κ-Heyting functors preserve full κ-infinitary first-order logic, in the following sense. Let $\mathcal{C}$ and $\mathcal{D}$ be κ-Heyting categories, and let $F : \mathcal{C} \to \mathcal{D}$ be a κ-Heyting functor. If $\varphi(x)$ is an $\mathscr{L}_\kappa$-formula and M is an $\mathscr{L}$-structure in $\mathcal{C}$, then*

$$F(\{x : \varphi(x)\}^M) = \{x : \varphi(x)\}^{F(M)}.$$

Proof. Since every κ-Heyting functor is in particular κ-geometric we can continue the proof by induction from Proposition 3.3.4. Universal quantification and meets of size $< \kappa$ are preserved by definition. The implication operation $A \to B$ could be constructed as $\forall_a(A \wedge B)$, where a is a monomorphism representing A (see Proposition 3.2.4). Since both operations in this construction are preserved, the entire construction is preserved. □

Corollary 3.3.6. *Let T be a κ-geometric theory and let M be a model of T in some κ-geometric category $\mathcal{C}$. If $F : \mathcal{C} \to \mathcal{D}$ is a κ-geometric morphism (with $\mathcal{D}$ also being κ-geometric), then $F(M)$ is also a model of T.*

Proof. We have for every κ-geometric axiom $\varphi \vdash_x \psi$ in T that $\{x : \varphi(x)\}^M \leq \{x : \psi(x)\}^M$. So by Proposition 3.3.4 we conclude that

$$\{x : \varphi(x)\}^{F(M)} = F(\{x : \varphi(x)\}^M) \leq F(\{x : \psi(x)\}^M) = \{x : \psi(x)\}^{F(M)},$$

from which it follows that $F(M)$ is indeed a model for T. □

By similar reasoning we also have the following corollary.

Corollary 3.3.7. *Let T be a κ-infinitary first-order theory and let M be a model of T in some κ-Heyting category $\mathcal{C}$. If $F : \mathcal{C} \to \mathcal{D}$ is a κ-Heyting morphism (with $\mathcal{D}$ also being κ-Heyting), then $F(M)$ is also a model of T.*

3.4 Internal language of a category

One useful application of the internal logic of a category is that we can reason about the objects and arrows in a category similar to the way we usually reason in **Set**. For this we need a language associated to the category.

Definition 3.4.1. For a small κ-geometric category $\mathcal{C}$ we define the *internal language* $\mathscr{L}(\mathcal{C})$ as follows. For each object X of $\mathcal{C}$ we have a sort $X_{\mathscr{L}(\mathcal{C})}$ in $\mathscr{L}(\mathcal{C})$. Likewise, for every arrow $f : X \to Y$ in $\mathcal{C}$ we add a function symbol $f_{\mathscr{L}(\mathcal{C})} : X_{\mathscr{L}(\mathcal{C})} \to Y_{\mathscr{L}(\mathcal{C})}$ to $\mathscr{L}(\mathcal{C})$. There are no additional relation symbols.

Note that we require $\mathcal{C}$ to be κ-geometric in the definition. Actually, we can define the internal language for any small category and we can already reason in it when our category is regular. However, we will only be interested in the cases where our category is at least κ-geometric, so for the rest of this section $\mathcal{C}$ is assumed to be κ-geometric.

There is a canonical way of interpreting $\mathscr{L}(\mathcal{C})$ in $\mathcal{C}$, by interpreting each sort $X_{\mathscr{L}(\mathcal{C})}$ as X and each function symbol $f_{\mathscr{L}(\mathcal{C})} : X_{\mathscr{L}(\mathcal{C})} \to Y_{\mathscr{L}(\mathcal{C})}$ as $f : X \to Y$. For better readability, and because it should be clear from the context which symbol is meant, we will no longer write the subscript in the sorts and function symbols (e.g. the sort $X_{\mathscr{L}(\mathcal{C})}$ will just be denoted by X).

The following examples show how we can use the internal language of a category to reason about its objects and arrows. For a more detailed exposition we refer to [But98, Section 5].

Example 3.4.2. An object X of $\mathcal{C}$ is a terminal object in $\mathcal{C}$ if and only if the sequents

$$\top \vdash_{\emptyset} \exists x(x = x),$$
$$\top \vdash_{x_1,x_2} x_1 = x_2$$

are valid, where x, x_1, x_2 are variables of sort X.

Example 3.4.3. A diagram

$$\begin{array}{ccc} P & \xrightarrow{g'} & X \\ {\scriptstyle f'}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow[g]{} & Z \end{array}$$

is a pullback diagram if and only if the following sequents are valid

$$\top \vdash_{p} f(g'(x)) = g(f'(x)),$$
$$f(x) = g(y) \vdash_{x,y} \exists p(g'(p) = x \wedge f'(p) = y),$$
$$g'(p) = g'(p') \wedge f'(p) = f'(p') \vdash_{p,p'} p = p',$$

where each variable is of the appropriate sort.

Let now $\mathcal{D}$ be some κ-geometric category $\mathcal{D}$ and M be some $\mathscr{L}(\mathcal{C})$-structure in $\mathcal{D}$. For every object X in $\mathcal{C}$ we then have an object X^M in $\mathcal{D}$, and for every arrow f in $\mathcal{C}$ we have an arrow f^M in $\mathcal{D}$. We can hope that M defines a functor $\mathcal{C} \to \mathcal{D}$. This is generally not true, because identity arrows and composition need not be preserved in general.

This can be fixed by defining a theory T in $\mathscr{L}(\mathcal{C})$, and requiring M to be a model of T. The following example illustrates this, and does even more: it produces a theory whose models are left exact functors $\mathcal{C} \to \mathcal{D}$.

Example 3.4.4. We define a theory T in $\mathscr{L}(\mathcal{C})$ of left exact functors on $\mathcal{C}$ as follows. First we add an axiom

$$\top \vdash_{x} Id_X(x) = x$$

for every object X in $\mathcal{C}$. This ensures that the identity arrows will be sent to identity arrows. Next we want to respect the composition operation, so we add an axiom

$$\top \vdash_{x} f(g(x)) = h(x),$$

where h denotes the composition fg, for each two composable arrows f and g.

Up until this point, models of T in $\mathcal{D}$ will be the same thing as functors $\mathcal{D} \to \mathcal{C}$. Now we only need to make sure that these functors are left exact. We do this by requiring that terminal objects and pullbacks are preserved. We thus add the sequents mentioned in Example 3.4.2 and Example 3.4.3 as an axiom, for every terminal object and pullback square in $\mathcal{C}$.

Essentially we have now established the equivalence

$$\mathbf{Lex}(\mathcal{C}, \mathcal{D}) \simeq T\text{–}\mathbf{Mod}(D),$$

natural in $\mathcal{D}$. Here $\mathbf{Lex}(\mathcal{C}, \mathcal{D})$ denotes the category of left exact functors from $\mathcal{C}$ to $\mathcal{D}$, with natural transformations between them.

3.5 Deduction-systems

Each of the logics we discussed comes with its own deduction-system. We are interested in the internal logic of categories, so we consider intuitionistic deduction-systems. As is desirable for deduction-systems, the deduction-systems that are described in this section will be sound and complete for their corresponding categorical semantics.

Our treatment of the deduction-systems is based on [Joh02b, section D1.3], but we will only be interested in the systems for κ-geometric logic and full κ-infinitary first-order logic (again, coherent logic is just the special case of κ-geometric with $\kappa = \omega$).

A *deduction-system* consists of rules of the form

$$\frac{\Gamma}{\sigma}$$

where Γ is a (possibly empty) set of sequents and σ is a sequent. On should read this as "from the sequents in Γ we may derive σ". In the case that Γ is empty we say that σ is an *axiom* and we omit the line above it. We use the notation

$$\frac{\sigma_1 \qquad \dots \qquad \sigma_n}{\tau}$$

when $\Gamma = \{\sigma_1, \dots, \sigma_n\}$.

A *derivation* in such a deduction-system is a tree that is built using the rules in the deduction-system. The sequents with no line above them (the *leaves* of the tree) should be axioms. The tree can be infinite in width (for example, when using infinitary connectives $\bigvee$ or $\bigwedge$), but should have no infinite ascending chain. So trees have the following form:

$$\frac{\dfrac{\sigma_1 \qquad \sigma_2}{\sigma_3} \qquad \dfrac{\sigma_4 \qquad \sigma_5}{\sigma_6} \qquad \sigma_7}{\sigma_8}$$

We call the root of the tree, σ_8 in the example above, the *conclusion* of the derivation.

The logical axioms, which are listed below, will always be axioms of any deduction-system. Additionally, the sequents in a theory T are considered to be axioms as well. A *derivation relative to T* is a derivation whose axioms are either logical axioms or are in T. If a sequent σ is the conclusion of a derivation relative to T, then we say that σ is *T-provable* or σ is *derivable from T*. We call two formulas φ and ψ *T-provably equivalent* (in the context x) if $\varphi \vdash_x \psi$ and $\psi \vdash_x \varphi$ are both T-provable.

As promised we will now introduce the deduction-systems for κ-geometric logic and full κ-infinitary first-order logic.

Definition 3.5.1. The deduction-system for κ-geometric logic has the following rules. Here, φ, ψ and χ can be any κ-geometric formula. Furthermore, Ψ denotes any set of κ-geometric formulas of size $< \kappa$. Of course, the context of the sequent the formulas appear in has to be suitable.

(identity) The axiom:

$$\varphi \vdash_x \varphi.$$

(substitution) Let t be any term with its free variables contained in the string of variables y, then we have

$$\frac{\varphi \vdash_x \psi}{\varphi[t/x] \vdash_y \psi[t/x]}$$

(cut) The rule:

$$\frac{\varphi \vdash_x \psi \qquad \psi \vdash_x \chi}{\varphi \vdash_x \chi}$$

(equality1) The axiom:

$$\top \vdash_x x = x.$$

(equality2) The axiom:

$$(x = y) \wedge \varphi \vdash_z \varphi[y/x],$$

where z is any context containing x, y and the free variables from φ.

($\wedge$-introduction) The rule:

$$\frac{\varphi \vdash_x \psi \qquad \varphi \vdash_x \chi}{\varphi \vdash_x \psi \wedge \chi}$$

($\wedge$-elimination) The axioms:

$$\varphi \vdash_x \top, \qquad \varphi \wedge \psi \vdash_x \varphi, \qquad \varphi \wedge \psi \vdash_x \psi.$$

($\bigvee$-introduction) For every $\psi \in \Psi$ we have the axioms:

$$\bot \vdash_x \varphi, \qquad \psi \vdash_x \bigvee \Psi.$$

In particular, we have the axioms:

$$\bot \vdash_x \varphi, \qquad \varphi \vdash_x \varphi \vee \psi, \qquad \psi \vdash_x \varphi \vee \psi.$$

($\bigvee$-elimination) Let $\Gamma = \{\psi \vdash_x \varphi : \psi \in \Psi\}$, then we have the rule:

$$\frac{\Gamma}{\bigvee \Psi \vdash_x \varphi}$$

In particular, we have:

$$\frac{\varphi \vdash_x \chi \qquad \psi \vdash_x \chi}{\varphi \vee \psi \vdash_x \chi}$$

($\exists$-introduction) In the following rule, y is a single variable (which is not free in ψ):

$$\frac{\varphi \vdash_{x,y} \psi}{\exists y \varphi \vdash_x \psi}$$

($\exists$-elimination) In the following rule, y is a single variable (which is not free in ψ):

$$\frac{\exists y \varphi \vdash_x \psi}{\varphi \vdash_{x,y} \psi}$$

Note that this is exactly the converse of the ($\exists$-introduction) rule.

(distributivity) This axiom can be derived in the deduction-system for full κ-infinitary first-order logic (see Proposition 3.5.4), but it is not in general derivable in κ-geometric logic. So we have to list it explicitly here. Denote by $\varphi \wedge \Psi$ the set $\{\varphi \wedge \psi : \psi \in \Psi\}$, then we have the axiom:

$$\varphi \wedge \bigvee \Psi \vdash_x \bigvee \varphi \wedge \Psi.$$

In particular, we have:

$$\varphi \wedge (\psi \vee \chi) \vdash_x (\varphi \wedge \psi) \vee (\varphi \wedge \chi).$$

(Frobenius) Just as with the (distributivity) axiom, which links $\wedge$ and $\vee$, we need to explicitly link $\wedge$ and $\exists$. Again, this axiom can be derived in full κ-infinitary first-order logic (see Proposition 3.5.4). So we have the following axiom:

$$\varphi \wedge \exists y \psi \vdash_x \exists y(\varphi \wedge \psi).$$

Here y is a single variable that is not contained in the context x (and hence not free in φ).

Note that the converses of (distributivity) and (Frobenius) also hold, but these are derivable as is shown in the following proposition.

Proposition 3.5.2. *The converse axioms of (distributivity) and (Frobenius) are derivable in the deduction-system for κ-geometric logic. That is, the following two statements are true for this deduction-system.*

(i) Let Ψ be a set of κ-geometrical formulas with $|\Psi| < \kappa$, we denote by $\varphi \wedge \Psi$ the set $\{\varphi \wedge \psi : \psi \in \Psi\}$. We can derive:

$$\bigvee \varphi \wedge \Psi \vdash_x \varphi \wedge \bigvee \Psi.$$

(ii) Let y be a single variable, not contained in x and not free in φ, then we can derive:

$$\exists y(\varphi \wedge \psi) \vdash_x \varphi \wedge \exists y \psi.$$

Proof. We will prove this proposition by simply providing the derivations, which then directly serve as examples of how to use the deduction-system.

The derivation of (i) is as follows, where the "..." should be taken to mean "insert here a derivation like the one on the left for every $\psi \in \Psi$":

$$\dfrac{\dfrac{\dfrac{\varphi \wedge \psi \vdash_x \psi \qquad \psi \vdash_x \bigvee \Psi}{\varphi \wedge \psi \vdash_x \bigvee \Psi}\,\text{(cut)} \qquad \varphi \wedge \psi \vdash_x \varphi}{\varphi \wedge \psi \vdash_x \varphi \wedge \bigvee \Psi}\,(\wedge\text{-introduction}) \qquad \dots}{\bigvee \varphi \wedge \Psi \vdash_x \varphi \wedge \bigvee \Psi}\,(\bigvee\text{-elimination})$$

For (ii) we have:

$$\dfrac{\dfrac{\varphi \wedge \psi \vdash_{x,y} \varphi}{\exists y(\varphi \wedge \psi) \vdash_x \varphi}\,(\exists\text{-introduction}) \qquad \dfrac{\dfrac{\varphi \wedge \psi \vdash_{x,y} \psi \qquad \dfrac{\exists y \psi \vdash_x \exists y \psi}{\psi \vdash_{x,y} \exists y \psi}\,(\exists\text{-elimination})}{\varphi \wedge \psi \vdash_{x,y} \exists y \psi}\,\text{(cut)}}{\exists y(\varphi \wedge \psi) \vdash_x \exists y \psi}\,(\exists\text{-introduction})}{\exists y(\varphi \wedge \psi) \vdash \varphi \wedge \exists y \psi}\,(\wedge\text{-introduction})$$

□

Except for the structural rules and axioms, (identity) to (equality2), the rules consist of introduction and elimination rules for each of the connectives that are allowed in κ-geometric logic. It should be no surprise that the deduction-system for full κ-infinitary first-order logic is just an extension of this. We make this precise in the following definition.

Definition 3.5.3. The deduction-system for full κ-infinitary first-order logic is an extension of the deduction-system for κ-geometric logic (see Definition 3.5.1), where we now allow all $\mathscr{L}_\kappa$-formulas to appear in the rules and axioms. In particular, that means that Ψ can now be any set of $\mathscr{L}_\kappa$-formulas of size $< \kappa$. Furthermore, we add the following rules and axioms.

($\bigwedge$-introduction) Let $\Gamma = \{\varphi \vdash_x \psi : \psi \in \Psi\}$, then we have the rule:

$$\dfrac{\Gamma}{\varphi \vdash_x \bigwedge \Psi}$$

($\bigwedge$-elimination) For each $\psi \in \Psi$ we have the axioms:

$$\varphi \vdash_x \top, \qquad \bigwedge \Psi \vdash_x \psi.$$

($\to$-introduction) The rule:

$$\frac{\varphi \wedge \psi \vdash_x \chi}{\psi \vdash_x \varphi \to \chi}$$

(→-elimination) The rule:

$$\frac{\psi \vdash_x \varphi \to \chi}{\varphi \wedge \psi \vdash_x \chi}$$

Note that this is exactly the converse of the (→-introduction) rule.

(∀-introduction) In the following rule, y is a single variable (which is not free in φ):

$$\frac{\varphi \vdash_{x,y} \psi}{\varphi \vdash_x \forall y \psi}$$

(∀-elimination) In the following rule, y is a single variable (which is not free in φ):

$$\frac{\varphi \vdash_x \forall y \psi}{\varphi \vdash_{x,y} \psi}$$

Note that this is exactly the converse of the (∀-introduction) rule.

We interpret the formula $\neg\varphi$ to be an abbreviation for $\varphi \to \bot$. So there is no need for separate rules for negation.

In Definition 3.5.1 we claimed that the axioms (distributivity) and (Frobenius) can be derived from the other rules of full κ-infinitary first-order logic. This claim is proved in the proposition below.

Proposition 3.5.4. *The axioms of (distributivity) and (Frobenius) can be derived from the other rules in the deduction-system for full κ-infinitary first-order logic.*

Proof. We will once more simply provide the derivations. In the derivations we may sometimes swap the order of conjunctions, for readability. Technically we have not proved yet that this can be done, but this is rather easy to see and is not probably not worth our attention.

The derivation of (distributivity) is as follows, where the "…" should be taken to mean "insert here a derivation like the one on the left for every $\psi \in \Psi$":

$$\cfrac{\cfrac{\cfrac{\varphi \wedge \psi \vdash_x \bigvee \varphi \wedge \Psi}{\psi \vdash_x \varphi \to \bigvee \varphi \wedge \Psi}\,(\to\text{-introduction}) \qquad \cdots}{\bigvee \Psi \vdash_x \varphi \to \bigvee \varphi \wedge \Psi}\,(\vee\text{-elimination})}{\varphi \wedge \bigvee \Psi \vdash_x \bigvee \varphi \wedge \Psi}\,(\to\text{-elimination})$$

For (Frobenius) we have:

$$\cfrac{\cfrac{\cfrac{\cfrac{\exists y(\varphi \wedge \psi) \vdash_x \exists y(\varphi \wedge \psi)}{\varphi \wedge \psi \vdash_{x,y} \exists y(\varphi \wedge \psi)}\,(\exists\text{-elimination})}{\psi \vdash_{x,y} \varphi \to \exists y(\varphi \wedge \psi)}\,(\to\text{-introduction})}{\exists y \psi \vdash_x \varphi \to \exists y(\varphi \wedge \psi)}\,(\exists\text{-introduction})}{\varphi \wedge \exists y \psi \vdash_x \exists y(\varphi \wedge \psi)}\,(\to\text{-elimination})$$

□

As mentioned before, the deduction-systems we discussed are intuitionistic. However, occasionally we may want to talk about classical logic. This system can be easily obtained by adding one more axiom.

Definition 3.5.5. The deduction-system for *classical logic* is obtained by adding to the deduction-system from Definition 3.5.3 the following axiom ($\neg\neg$-elimination):

$$(\varphi \to \bot) \to \bot \vdash_x \varphi.$$

Or, if we use the abbreviation we discussed above:

$$\neg\neg\varphi \vdash_x \varphi.$$

When we talk about a sequent being provable or derivable, we will always mean relative to the intuitionistic deduction-system (whether we mean the geometric or first-order variant should be clear from the context). If we are interested in the deduction-system for classical logic, this will always be explicitly stated. For example we would say something is "classically provable" or "derivable in the deduction-system for classical logic".

For both the deduction-systems for κ-geometric logic and for full κ-infinitary first-order logic we have soundness and completeness theorems, with respect to the interpretation of logic as discussed in section 3.2.

Theorem 3.5.6 (Soundness). *Let T be a κ-geometric (κ-infinitary first-order) theory. Suppose σ is a κ-geometric (κ-infinitary first-order) sequent that is T-provable in the deduction-system for κ-geometric (κ-infinitary first-order) logic, then for any model M of T in any κ-geometric (κ-Heyting) category we have that σ is valid in M.*

Proof. The proof is fairly straightforward by induction on the construction of a derivation in the relevant deduction-system. For more information we refer to [Joh02b, D1.3, prop 1.3.2]. □

Theorem 3.5.6 also explains why we cannot in general use classical logic, because for arbitrary A in a subobject lattice in a κ-Heyting category, $\neg\neg A \leq A$ does not in general hold. However, this inequality holds if and only if such a lattice is a Boolean algebra. So the deduction-system for classical logic is sound for κ-Heyting categories where each subobject lattice is a Boolean algebra.

Theorem 3.5.7 (Completeness). *Let T be a κ-geometric (κ-infinitary first-order) theory. If for any model M of T in any κ-geometric (κ-Heyting) category we have that σ is valid in M, then σ is T-provable in the deduction-system for κ-geometric (κ-infinitary first-order) logic.*

The proof of the completeness theorem uses tools that we develop later on, so we refer to chapter 8 for the proof and similar (stronger) results. There we will also obtain a completeness result for the deduction-system for infinitary classical logic.

Chapter 4

Classifying topoi

We now have two different notions: that of geometric functors (i.e. ∞-geometric functors) and geometric morphisms. These should not be confused with each other. A geometric functor is a functor between geometric categories, while a geometric morphism is a morphism of topoi. However, the similarity in the names is no coincidence as will be made clear in the following proposition.

Proposition 4.0.8. *Let $f : \mathcal{E} \to \mathcal{F}$ be a geometric morphism, then its inverse image part $f^* : \mathcal{F} \to \mathcal{E}$ is a geometric functor.*

Proof. The inverse image part of a geometric morphism is by definition left exact, and since it is a left adjoint it preserves all colimits, so in particular coequalizers (i.e. regular epimorphisms) and all joins. $\square$

In particular we have that the inverse image part of a geometric morphism preserves geometric logic. As we mentioned before in section 2.2, the open geometric morphisms should be those geometric morphisms that preserve first-order logic. This is the case thanks to the following proposition.

Proposition 4.0.9. *Let $f : \mathcal{E} \to \mathcal{F}$ be a geometric morphism, then f is open if and only if f^* preserves the universal quantifier. We mean this in the following sense: let $h : X \to Y$ be an arrow in $\mathcal{F}$ and let A be a subobject of X, then $f^*(\forall_h(A)) = \forall_{f^*(h)}(f^*(A))$. In particular, f is open if and only if f^* is a Heyting functor.*

Proof. The first statement of the proposition are exactly cases (i) and (v) in [Joh80, Theorem 3.2]. If f^* is a Heyting functor, it preserves universal quantification and is thus open. If f is open, then f^* preserves universal quantification so we are just left to show that f^* preserves infinitary meets in any subobject lattice. The following proof for this fact is based on the proof of [MLM92, Theorem IX.6.3].

Let $\{A_i : i \in I\}$ be a family of subobjects of some object X in $\mathcal{F}$. Then $\coprod_i A_i$ is a subobject of $\coprod_i X$. Let $\alpha : \coprod_i X \to X$ be the arrow determined by the identity arrows on X, then we claim that

$$\forall_\alpha(\coprod_{i\in I} A_i) = \bigwedge_{i\in I} A_i.$$

Before proving this claim, we show how our desired result follows from it. Since f^* is a left adjoint, it preserves colimits and by assumption it preserves universal quantification. We thus find

$$f^*(\bigwedge_{i\in I} A_i) = f^*(\forall_\alpha(\coprod_{i\in I} A_i)) = \forall_{f^*(\alpha)}(\coprod_{i\in I} f^*(A_i)) = \bigwedge_{i\in I} f^*(A_i).$$

To prove the claim we will show that for any subobject B of X we have:

$$B \leq \forall_\alpha(\coprod_{i\in I} A_i) \iff \alpha^*(B) \leq \coprod_{i\in I} A_i \iff B \leq A_i \text{ (for all } i \in I).$$

The first "if and only if" is simply the adjunction $\alpha^* \dashv \forall_\alpha$. Before we prove the second one, we first note that $\alpha^*(B) = \coprod_{i\in I} B$ since pulling back in a topos preserves coproducts:

$$\begin{tikzcd}
\coprod_{i\in I} B = \coprod_{i\in I} b^*(X) = b^*(\coprod_{i\in I} X) = \alpha^*(B) \arrow[r] \arrow[d, tail] & B \arrow[d, tail, "b"] \\
\coprod_{i\in I} X \arrow[r, "\alpha"'] & X
\end{tikzcd}$$

where b denotes a monomorphism representing the subobject B. So the statement from the right to the left follows directly, because then $\coprod_{i\in I} A_i$ is a cocone for $\{B\}_{i\in I}$ as well. To prove the statement from the left to the right, we note that for every $j \in I$ we have the following commuting diagram (without the dashed arrow):

$$\begin{tikzcd}
A_j \arrow[r, hook] \arrow[d, tail, "a_j"'] & \coprod_{i\in I} A_i \arrow[d, tail] \\
X \arrow[r, hook] & \coprod_{i\in I} X \\
B \arrow[r, hook] \arrow[u, tail, "b"] \arrow[uu, dashed, bend left, "g"] & \coprod_{i\in I} B \arrow[u, tail] \arrow[uu, bend right, "f"']
\end{tikzcd}$$

The upper square is a pullback, as we will see below, so we obtain the dashed arrow and conclude that indeed $B \leq A_j$. To see that the upper square is indeed a pullback, we denote by χ_j the j-th inclusion of X into $\coprod_{i\in I} X$. Again using that coproducts are stable under pullbacks, we have

$$\chi_j^*(\coprod_{i\in I} A_i) = \coprod_{i\in I} \chi_j^*(A_i) = A_j.$$

Here the last equality follows because coproducts are disjoint, so

$$\chi_j^*(A_i) = \begin{cases} 0 & i \neq j \\ A_j & i = j \end{cases}$$

□

In particular we have established that a model M in $\mathcal{F}$ of a geometric theory T is sent to a model $f^*(M)$ in $\mathcal{E}$ for every geometric morphism $f : \mathcal{E} \to \mathcal{F}$. With this in mind we can define what we mean by a classifying topos.

Definition 4.0.10. Let T be a geometric theory, then a *classifying topos* of T, denoted by $\mathbf{Set}[T]$, is a topos such that there is an equivalence of categories

$$\mathbf{Topos}(\mathcal{E}, \mathbf{Set}[T]) \simeq T\text{–}\mathbf{Mod}(\mathcal{E})$$

that is natural in $\mathcal{E}$.

From the definition it follows that $\mathbf{Set}[T]$ is unique up to equivalence. Furthermore, as we will see in Theorem 6.2.1, the classifying topos of a geometric theory always exists. So from now on we will talk about 'the' classifying topos.

Under this correspondence, the identity on $\mathbf{Set}[T]$ should correspond to a model of T in $\mathbf{Set}[T]$. Denote this model by G_T, then by naturality we have for any $f : \mathcal{E} \to \mathbf{Set}[T]$ the following commuting diagram

$$\begin{tikzcd}
Id \in \mathbf{Topos}(\mathbf{Set}[T], \mathbf{Set}[T]) \arrow[r, dash, "\sim"] \arrow[d, "-\circ f"'] & T\text{–}\mathbf{Mod}(\mathbf{Set}[T]) \ni G_T \arrow[d, "f^*(-)"] \\
f \in \mathbf{Topos}(\mathcal{E}, \mathbf{Set}[T]) \arrow[r, dash, "\sim"] & T\text{–}\mathbf{Mod}(\mathcal{E}) \ni f^*(G_T)
\end{tikzcd}$$

So we see that the model in $\mathcal{E}$ corresponding to f is given by $f^*(G_T)$.

Definition 4.0.11. The model G_T in $\mathbf{Set}[T]$ is called the *generic model*.

Following the idea of the classifying topos for a geometric theory T we can also define something similar for arbitrary infinitary first-order theories. However, in this case we can no longer look at the geometric morphisms, because they may not preserve models of such a theory. If we want to preserve such models we have to preserve first-order logic, which includes universal quantification. By Proposition 4.0.9 the open geometric morphisms are precisely the geometric morphisms that do this (and they are the only candidate).

Definition 4.0.12. Let T be an infinitary first-order theory, then the *first-order classifying topos* $\mathbf{Set}^{\mathrm{fo}}[T]$ of T, if it exists, is the topos such that there is an equivalence of categories

$$\mathbf{Open}(\mathcal{E}, \mathbf{Set}^{\mathrm{fo}}[T]) \simeq T\text{–}\mathbf{Mod}(\mathcal{E})_\infty$$

that is natural in $\mathcal{E}$. As before, for such a first-order classifying topos there has to be a generic model G_T in $\mathbf{Set}^{\mathrm{fo}}[T]$ such that the model corresponding to some open $f : \mathcal{E} \to \mathbf{Set}^{\mathrm{fo}}[T]$ is given by $f^*(G_T)$.

We use different terminology for the classifying topos (the geometric case) and the first-order classifying topos, because they can be different objects. A geometric theory can of course also be considered to be a first-order theory, but while its classifying topos then always exists, it does not necessarily have a first-order classifying topos. An example of such a theory can be found in section 4.1, where we point out what kind of problems can arise that prevent the existence of a first-order classifying topos. Later, in chapter 9, we characterize which infinitary first-order theories have a first-order classifying topos.

We should also point out that in Definition 4.0.12 we had to cut down to the category $T\text{–}\mathbf{Mod}(\mathcal{E})_\infty$ instead of $T\text{–}\mathbf{Mod}(\mathcal{E})$. That is, we had to cut down to elementary morphisms instead of homomorphisms. This is because the inverse image parts of open geometric morphisms preserve full infinitary first-order logic.

Proposition 4.0.13. *The natural transformations between open geometric morphisms into a first-order classifying topos correspond to ∞-elementary embeddings. That is, suppose that for a certain theory T we have*

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

natural in $\mathcal{F}$. Then $T\text{–}\mathbf{Mod}(\mathcal{F}) = T\text{–}\mathbf{Mod}(\mathcal{F})_\infty$.

Proof. Let the situation be as in the statement of the proposition. Let furthermore M and N be models of T in $\mathcal{F}$ and let $h : M \to N$ be a homormorphism of models. Denote by G_T the model in $\mathcal{E}$ corresponding to the identity morphism on $\mathcal{E}$, then there are open geometric morphisms $m, n : \mathcal{F} \to \mathcal{E}$ such that $M \cong m^*(G_T)$ and $N \cong n^*(G_T)$. The goal will now be to show that given an arbitrary $\mathscr{L}_\infty$-formula in the language of T we have

$$\{x : \varphi(x)\}^{m^*(G_T)} \leq h_X^*(\{x : \varphi(x)\}^{n^*(G_T)}).$$

We let $\tau : m^* \to n^*$ be the natural transformation corresponding to h. That is, $\tau_{X^{G_T}} = h_X$. By naturality of τ we have that the outer square in the diagram below commutes.

$$\begin{tikzcd}
\{x:\varphi(x)\}^{m^*(G_T)} \arrow[dr, dashed] \arrow[drr, bend left, "\tau_{\{x:\varphi(x)\}^{G_T}}"] \arrow[ddr, tail, bend right] & & \\
& h_X^*(\{x:\varphi(x)\}^{n^*(G_T)}) \arrow[r] \arrow[d, tail] & \{x:\varphi(x)\}^{n^*(G_T)} \arrow[d, tail] \\
& X^{m^*(G_T)} \arrow[r, "\tau_{X^{G_T}} = h_X"'] & X^{n^*(G_T)}
\end{tikzcd}$$

So by the universal property of the pullback, the dashed arrow exists and we do indeed have the inequality that we needed. □

4.1 The problem with a first-order classifying topos

In this section we will provide an example of a first-order theory T that cannot have a first-order classifying topos in the sense of Definition 4.0.12. This argument comes directly from the introduction in [BJ98], but it is an important example so we will treat it here as well.

For this example, let us fix a language $\mathscr{L}$ that consists of just two propositional variables P and Q (i.e. 0-ary relation symbols) and nothing else. Our theory T will be the empty theory. We will show that there can be no first-order classifying topos $\mathbf{Set}^{\mathrm{fo}}[T]$. So let us suppose there is such a topos, and we will aim for a contradiction.

In [Jon80] it is established that there is a proper class of $\mathscr{L}_\infty$-formulas, none of which are provably equivalent. That means that for every cardinal κ there is a complete Heyting algebra H_κ of size at least κ generated by two elements. By "generated by two elements" we mean that there are $A, B \in H_\kappa$ such that any element of H_κ can be written as $\varphi(A, B)$ for some propositional $\mathscr{L}_\infty$-formula φ.

Considering H_κ as a locale, we obtain the localic topos $\mathbf{Sh}(H_\kappa)$. The subobject lattice Sub(1) in $\mathbf{Sh}(H_\kappa)$ is just (isomorphic to) H_κ. That means that $\mathbf{Sh}(H_\kappa)$ contains H_κ as a model for T, with A as the interpretation for P and B as the interpretation for Q.

Let $f : \mathbf{Sh}(H_\kappa) \to \mathbf{Set}^{\mathrm{fo}}[T]$ be the open geometric morphism corresponding to the model H_κ in $\mathbf{Sh}(H_\kappa)$. We claim that f^* is surjective on the subobjects of 1 in $\mathbf{Sh}(H_\kappa)$. To see this, let any such subobject be given. This subobject is then of the form $\varphi(A, B)$ for some propositional $\mathscr{L}_\infty$-formula φ. Since open geometric morphisms preserve full infinitary first-order logic (Proposition 4.0.9), we find

$$f^*(\{\emptyset : \varphi(P, Q)\}^{G_T}) = \{\emptyset : \varphi(P, Q)\}^{H_\kappa} = \varphi(A, B).$$

This means that for any cardinaly κ there is a surjection from the subobjects of 1 in $\mathbf{Set}^{\mathrm{fo}}[T]$ to the subobjects of 1 in $\mathbf{Sh}(H_\kappa)$. So Sub(1) in $\mathbf{Set}^{\mathrm{fo}}[T]$ would be a proper class, which cannot happen. So we conclude that $\mathbf{Set}^{\mathrm{fo}}[T]$ cannot exist.

4.2 Morita-equivalence

It is well-known that a Boolean ring is 'the same thing' as a Boolean algebra, even though they are defined in different languages. Nevertheless, we would like to consider the theory of Boolean rings and the theory of Boolean algebras as the same. Every Boolean ring determines, up to isomorphism, a Boolean algebra and vice versa. The same holds for homomorphisms of Boolean rings and Boolean algebras. In other words, we have an equivalence (even an isomorphism) of the category of Boolean rings and the category of Boolean algebras. This turns out to be a good definition of when two theories are considered to be the same.

Definition 4.2.1. We call two geometric theories T and T' *Morita-equivalent* if for every topos $\mathcal{E}$ we have an equivalence

$$T\text{–}\mathbf{Mod}(\mathcal{E}) \simeq T'\text{–}\mathbf{Mod}(\mathcal{E})$$

that is natural in $\mathcal{E}$ with respect to geometric morphisms.

From the definition of the classifying topos of a theory it is clear that a classifying topos for T is also a classifying topos for T', if T and T' are Morita-equivalent. In other words, we have that T and T' are Morita-equivlant if and only if they have the same classifying topos (up to equivalence).

To extend this definition to infinitary first-order theories, we can no longer require the equivalence to be natural in $\mathcal{E}$ for all geometric morphisms, because they may not preserve models. We therefore make the following definition.

Definition 4.2.2. We call two infinitary first-order theories T and T' *Morita-equivalent* if for every topos $\mathcal{E}$ we have an equivalence

$$T\text{–}\mathbf{Mod}(\mathcal{E})_\infty \simeq T'\text{–}\mathbf{Mod}(\mathcal{E})_\infty$$

that is natural in $\mathcal{E}$ with respect to open geometric morphisms.

Note that as always when we restrict ourselves to open geometric morphisms, we have to cut down to elementary morphisms between the models. Like before, we can link Morita-equivalence to equivalence of the first-order classifying topoi of infinitary first-order theories T and T'. However, in this case we need to take a little care because T and T' may not have first-order classifying topoi. If they have first-order classifying topoi and these are equivalent, then clearly T and T' are Morita-equivalent. In the other direction we have that if one of the theories, say T, has a first-order classifying topos, and the theories are Morita-equivalent, then T' has a first-order classifying topos and we have $\mathbf{Set}^{\mathrm{fo}}[T] \simeq \mathbf{Set}^{\mathrm{fo}}[T']$.

Chapter 5

The syntactic category for a geometric theory

For this entire chapter, we fix a κ-geometric theory T in some language $\mathscr{L}$. In this chapter we will show how to construct the syntactic category from T, this category is then later used to construct the classifying topos for T.

5.1 Construction of the syntactic category

In this section we will describe how to construct the *syntactic category* $\mathbf{Syn}_\kappa^{\mathrm{g}}(T)$. The objects of $\mathbf{Syn}_\kappa^{\mathrm{g}}(T)$ are equivalence classes $[x.\varphi]$ of κ-geometric formulas in context, where two formulas $\varphi(x)$ and $\psi(y)$ are equivalent if x and y have the same type and $\varphi = \psi[x/y]$. Given two objects $[x.\varphi]$ and $[y.\psi]$ we can thus assume x and y to have no variables in common. This is convenient for defining the arrows: an arrow $[x.\varphi] \to [y.\psi]$ is a T-provable equivalence class of κ-geometric formulas $[x, y.\theta]$ that are *T-provably functional from φ to ψ*. That is, the following κ-geometric axioms are derivable from T:

1. $\theta(x, y) \vdash_{x,y} \varphi(x) \wedge \psi(y)$,
2. $\varphi(x) \vdash_x \exists y \theta(x, y)$,
3. $\theta(x, y) \wedge \theta(x, y') \vdash_{x,y,y'} y = y'$.

Here it should be noted that the second and third axiom are automatically fulfilled when y is the empty context.

When proving that two formulas represent the same arrow, one has to show that they are T-provably equivalent. The following proposition tells us that we only need to show one direction of this equivalence and is thus a useful tool.

Proposition 5.1.1. *Let $\theta(x, y)$ and $\sigma(x, y)$ be two formulas that are T-provably functional from $[x.\varphi]$ to $[y.\psi]$. If $\theta(x, y) \vdash_{x,y} \sigma(x, y)$, then the converse $\sigma(x, y) \vdash_{x,y} \theta(x, y)$ must also hold and hence we have that $[x, y.\theta] = [x, y.\sigma]$ as arrows in $\mathbf{Syn}_\kappa^{\mathrm{g}}(T)$.*

Proof. We reason in the deduction-system for T. Suppose that x and y are such that $\sigma(x, y)$. Then since θ is T-provably functional, we find y' such that $\theta(x, y')$. Then because $\theta(x, y) \vdash_{x,y} \sigma(x, y)$ we have $\sigma(x, y')$. Since both $\sigma(x, y)$ and $\sigma(x, y')$ hold we see that $y = y'$ because σ is T-provably functional. We can now conclude that indeed $\theta(x, y)$ must hold. $\square$

Given two arrows $[x.\varphi] \xrightarrow{[x,y.\theta]} [y.\psi] \xrightarrow{[y,z.\sigma]} [z.\chi]$, their composition is given by $[x, z.\exists y(\theta(x, y) \wedge \sigma(y, z))]$. Since all representatives of the arrows are T-provably equivalent by definition, the definition of their composition does not depend on the choice of representatives. Of course, we need to verify that this really gives a composition operation and that there is an identity arrow. This is done in the following proposition.

Proposition 5.1.2. *The operation defined above is indeed a composition operation, with identity arrow* $[x, y.\varphi(x) \wedge x = y]$ *for every object* $[x.\varphi]$*.*

Proof. Given arrows $[x.\varphi] \xrightarrow{[x,y.\theta]} [y.\psi] \xrightarrow{[y,z.\sigma]} [z.\chi]$, it is straightforward to check that $\exists y(\theta(x, y) \wedge \sigma(y, z))$ is T-provably functional. Let now $[z, w.\tau] : [z.\chi] \to [w.\gamma]$, then it is also easy to check that $\exists z(\exists y(\theta(x, y) \wedge \sigma(y, z)) \wedge \tau(z, w))$ is equivalent to $\exists y(\theta(x, y) \wedge \exists z(\sigma(y, z) \wedge \tau(z, w)))$. This shows that the defined operation is associative and is therefore a composition operation.

Next we check that the composition of

$$[x.\varphi] \xrightarrow{[x,y.\varphi(x) \wedge x=y]} [y.\varphi] \xrightarrow{[y,z.\theta]} [z.\psi]$$

is just $[x, z.\theta]$. The composition is by definition represented by $\exists y(\varphi(x) \wedge x = y \wedge \theta(y, z))$, which is equivalent to $\varphi(x) \wedge \theta(x, z)$. Because $\theta(x, y)$ is T-provably functional from φ to ψ we see that $\varphi(x) \wedge \theta(x, z)$ is again equivalent to $\theta(x, z)$. The result for composition with the identity arrow on the right side then follows analogously. □

Now that we have established that $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$ is indeed a category, we will prove a quick useful fact about it.

Proposition 5.1.3. *Let* $[x.\varphi]$ *and* $[y.\psi]$ *be objects in* $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$*, with* x *and* y *of the same type. If* $\varphi(x)$ *and* $\psi(x)$ *are* T*-provably equivalent, then* $\theta(x, y) := \varphi(x) \wedge x = y \wedge \psi(y)$ *represents an arrow* $[x.\varphi] \to [y.\psi]$ *and this arrow is an isomorphism.*

Proof. We will first show that $[x, y.\theta]$ defines an arrow $[x.\varphi] \to [y.\psi]$. So we need to check that $\theta(x, y)$ is T-provably functional. From the definition of $\theta(x, y)$ the first and third axiom are clearly derivable from T. To derive the second axiom from T we let x be arbitrary and suppose that $\varphi(x)$ holds. Then since we have $\varphi \vdash_x \psi$ by assumption, we find $\psi(x)$. Taking $y = x$ we conclude that $\exists y \theta(x, y)$, as required.

We define $\sigma(y, z)$ to be $\psi(y) \wedge y = z \wedge \varphi(z)$, then analogous to the discussion above we see that $[y, z.\sigma]$ defines an arrow $[y.\psi] \to [z.\varphi(z)] = [x.\varphi]$. We claim that $[x, y.\theta]$ and $[y, z.\sigma]$ are inverse to each other. Their composition is represented by

$$\exists y(\theta(x, y) \wedge \sigma(y, z)),$$

which is just

$$\exists y(\varphi(x) \wedge x = y \wedge \psi(y) \wedge \psi(y) \wedge y = z \wedge \varphi(z)).$$

This clearly implies $\varphi(x) \wedge x = z$, which represents the identity arrow on $[x.\varphi]$ (as we saw in Proposition 5.1.2). By Proposition 5.1.1 we thus find that $[y, z.\sigma][x, y.\theta]$ is the identity arrow on $[x.\varphi]$. Likewise we see that $[x, y.\theta][y, z.\sigma]$ is the identity on $[y.\psi]$, and we conclude that $[x, y.\theta]$ is indeed an isomorphism. □

Note that $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$ is a small category as long as $\kappa < \infty$. However, there is no real reason to pay too much attention to the size here. We can just as well look at the (large) category $\mathbf{Syn}^{\mathrm{g}}(T)$ for any geometric theory T (even if T is κ-geometric). This is also why we do not have the problems described in section 4.1, when looking at the classifying topos for geometric theories, because there are no such issues with size. The reason for this is the following theorem and the corollary right after it.

Theorem 5.1.4. *Every geometric formula* $\varphi(x)$ *is provably equivalent to* $\bigvee_{i \in I} \varphi_i(x)$ *where each* $\varphi_i(x)$ *is a regular formula. In particular, there is a* κ *such that every geometric formula is provably equivalent to a* κ*-geometric formula.*

Recall that a *regular formula* is a formula constructed using only finite conjunction and existential quantification.

Proof. As is described in [Joh02b, D1.3, Lemma 1.3.8(ii)], one can use the (distributivity) rule, its provable converse and the the fact that $\exists x \bigvee \Psi$ is provably equivalent to $\bigvee_{\psi \in \Psi} \exists x \psi$ to move all the disjunctions to the outer most level. What is left is a formula of the form $\bigvee_{i \in I} \varphi_i(x)$, where each $\varphi_i(x)$ is a regular formula, that is equivalent to the formula we started with.

The last claim follows from the fact that there is only a set S of different regular formulas. We can assume the $\varphi_i(x)$ in $\bigvee_{i \in I} \varphi_i(x)$ to be different, so taking $\kappa > |S|$ we find that every geometric formula is provably equivalent to a disjunction of size $< \kappa$ of regular formulas. □

Corollary 5.1.5. *For large enough κ the inclusion $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T) \to \mathbf{Syn}^{\mathrm{g}}(T)$ is an equivalence of categories.*

Proof. By Proposition 5.1.3 we have that T-provable equivalent formulas give isomorphic objects in the syntactic category. So if we take κ as described in Theorem 5.1.4, then every object in $\mathbf{Syn}^{\mathrm{g}}(T)$, which is represented by a geometric formula, is isomorphic to some object from $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$, which is represented by a κ-geometric formula. Likewise we find that the inclusion is full and faithful, because arrows are also represented by (κ-)geometric formulas. $\square$

5.2 The syntactic category is geometric

The goal of this section will be to show that $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$ is κ-geometric. Some of the explicit constructions here will turn out to be useful later on.

In this section, when we say "formula" this means "κ-geometric formula". One reason for this is that the contents of this section will later also be applied to the more general case where we will mean "κ-infinitary first-order formula" instead of "formula".

Lemma 5.2.1. *The syntactic category $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$ has all finite products.*

Proof. The terminal object is given by $[\emptyset.\top]$, where $\emptyset$ denotes the empty context here. Let $[x.\varphi]$ be any other object, then $[x, \emptyset.\varphi(x)]$ is an arrow into $[\emptyset.\top]$. Any other arrow $[x, \emptyset.\theta(x)]$ going from $[x.\varphi]$ to $[\emptyset.\top]$ satisfies $\theta(x) \vdash_x \varphi(x)$ and so by Proposition 5.1.1 they are the same arrow.

Next we will show what binary products are in $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$. Let $[x.\varphi]$ and $[y.\psi]$ be two objects, we claim that their product is given by $[x, y.\varphi(x) \wedge \psi(y)]$ with projections

$$[x, y, x'.\varphi(x) \wedge \psi(y) \wedge x = x'] : [x, y.\varphi(x) \wedge \psi(y)] \to [x.\varphi],$$
$$[x, y, y'.\varphi(x) \wedge \psi(y) \wedge y = y'] : [x, y.\varphi(x) \wedge \psi(y)] \to [y.\psi].$$

Let now $[w.\chi]$ be another object with arrows

$$[w, x.\sigma_1] : [w.\chi] \to [x.\varphi],$$
$$[w, y.\sigma_2] : [w.\chi] \to [y.\psi].$$

Then $[w, x, y.\sigma_1(w, x) \wedge \sigma_2(w, y)]$ defines an arrow $[w.\chi] \to [x, y.\varphi(x) \wedge \psi(y)]$ such that its composition with the projections is $[w, x.\sigma_1]$ or $[w, y.\sigma_2]$ respectively. To see that this arrow is unique, we let $[w, x, y.\tau]$ be another arrow $[w.\chi] \to [x, y.\varphi(x) \wedge \psi(y)]$ such that its composition with the projections is $[w, x.\sigma_1]$ or $[w, y.\sigma_2]$ respectively. The composition of $[w, x, y.\tau]$ with the first projection is given by $\exists xy(\tau(w, x, y) \wedge \varphi(x) \wedge \psi(y) \wedge x = x')$. Thus we have $\exists xy(\tau(w, x, y) \wedge \varphi(x) \wedge \psi(y) \wedge x = x') \vdash_{w,x'} \sigma_1(w, x')$. Now, under the assumption that $\tau(w, x', y')$ we find that certainly $\exists xy(\tau(w, x, y) \wedge \varphi(x) \wedge \psi(y) \wedge x = x')$ and hence $\sigma_1(w, x')$. Likewise, we find $\tau(w, x', y') \vdash_{w,x',y'} \sigma_2(w, y')$. So we conclude that $\tau(w, x', y') \vdash_{w,x',y'} \sigma_1(w, x') \wedge \sigma_2(w, y')$ and by Proposition 5.1.1 we thus have that $[w, x, y.\tau] = [w, x, y.\sigma_1(w, x) \wedge \sigma_2(w, y)]$. $\square$

Lemma 5.2.2. *The syntactic category $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$ has equalizers.*

Proof. Let $[x, y.\theta], [x, y.\sigma] : [x.\varphi] \to [y.\psi]$ be two parallel arrows in $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$. Then

$$\exists y(\theta(z, y) \wedge \sigma(z, y)) \wedge z = x$$

represents an arrow e from $[z.\exists y(\theta(z, y) \wedge \sigma(z, y))]$ to $[x.\varphi]$ (here z is to be taken of the same type as x). To see that $[x, y.\theta]e = [x, y.\sigma]e$ we consider the composition $[x, y.\theta]e$, which is represented by

$$\exists x(\exists y'(\theta(z, y') \wedge \sigma(z, y')) \wedge z = x \wedge \theta(x, y)).$$

From this it follows that

$$\exists y'(\theta(z, y') \wedge \sigma(z, y')) \wedge \theta(z, y),$$

from which we can deduce that $\theta(z,y) \wedge \sigma(z,y)$ since we have both $\theta(z,y)$ and $\theta(z,y')$ and θ is T-provably functional. Then we can deduce further that

$$\exists y'(\theta(z,y') \wedge \sigma(z,y')) \wedge \sigma(z,y),$$

and hence

$$\exists x(\exists y'(\theta(z,y') \wedge \sigma(z,y')) \wedge z = x \wedge \sigma(x,y)),$$

which represents exactly the composition $[x,y.\sigma]e$. So by Proposition 5.1.1 we conclude that indeed $[x,y.\theta]e = [x,y.\sigma]e$.

Let now $[w,x.\tau] : [w.\chi] \to [x.\varphi]$ be any arrow such that $[x,y.\theta][w,x.\tau] = [x,y.\sigma][w,x.\tau]$. It is straightforward to check that $[w,z.\tau(w,z)]$ defines an arrow with codomain $[z.\exists y(\theta(z,y) \wedge \sigma(z,y))]$, such that its composition with e is exactly $[w,x.\tau]$. Furthermore, any arrow $f : [w.\chi] \to [z.\exists y(\theta(z,y) \wedge \sigma(z,y))]$ such that $ef = [w,x.\tau]$ must be represented by some formula that is derivable from $\tau(w,x)$ relative T, so we must have that $f = [w,z.\tau]$. $\square$

Corollary 5.2.3. *The syntactic category* $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$ *has all finite limits.*

Example 5.2.4. We can now use the explicit descriptions of the product (Lemma 5.2.1) and the equalizer (Lemma 5.2.2) to explicitly describe what the pullback of any diagram $[x.\varphi] \xrightarrow{[x,z.\theta]} [z.\chi] \xleftarrow{[y,z.\sigma]} [y.\psi]$ looks like. First we compute the product of $[x.\varphi]$ and $[y.\psi]$, which is given by $[x,y.\varphi(x) \wedge \psi(y)]$ with projections $\pi_x = [x,y,x'.\varphi(x) \wedge \psi(y) \wedge x = x']$ and $\pi_y = [x,y,y'.\varphi(x) \wedge \psi(y) \wedge y = y']$. The composition $[x,z.\theta]\pi_x$ is then $[x,y,z.\exists x'(\varphi(x) \wedge \psi(y) \wedge x = x' \wedge \theta(x',z)]$ which is just $[x,y,z.\psi(y) \wedge \theta(x,z)]$. We will call this arrow f_x. Likewise, we find that $[y,z.\sigma]\pi_y$ is $f_y = [x,y,z.\varphi(x) \wedge \sigma(y,z)]$.

To find the pullback we now only need to determine the equalizer of f_x and f_y. The object of the equalizer is given by $[x',y'.\exists z(\psi(y') \wedge \theta(x',z) \wedge \varphi(x') \wedge \sigma(y',z))]$, but by Proposition 5.1.3 this object is isomorphic to

$$[x',y'.\exists z(\theta(x',z) \wedge \sigma(y',z))].$$

The corresponding arrow to the product $[x,y.\varphi(x) \wedge \psi(y)]$ is then $[x',y',x,y.\exists z(\theta(x',z) \wedge \sigma(y',z)) \wedge x = x' \wedge y = y']$, which we will denote by e.

So the pullback of $[y,z.\sigma]$ along $[x,z.\theta]$ is the given by the composition $\pi_x e$, which is explicitly given by

$$\begin{aligned}&[x',y',x''.\exists xy(\exists z(\theta(x',z) \wedge \sigma(y',z)) \wedge x = x' \wedge y = y' \wedge \varphi(x) \wedge \psi(y) \wedge x = x'')] \qquad = \\ &[x',y',x.\exists z(\theta(x',z) \wedge \sigma(y',z)) \wedge x = x'].\end{aligned}$$

Likewise, the pullback of $[x,z.\theta]$ along $[y,z.\sigma]$ is $\pi_y e = [x',y',y.\exists z(\theta(x',z) \wedge \sigma(y',z)) \wedge y = y']$.

Proposition 5.2.5. *The syntactic category* $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$ *is a regular category.*

Proof. By Corollary 5.2.3 we only need to prove now that in $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$ the coequalizer of the kernel pair of any arrow exists, and that regular epimorphisms are stable under pullback. Here we will only give a quick proof of the fact that the coequalizer of the kernel pair of any arrow exists, because its explicit construction will be useful later on. The rest of the proof can be found in [Joh02b, D1.4, Lemma 1.4.10].

Let $p_0, p_1 : [z.\chi] \to [x.\varphi]$ be the kernel pair of some arrow $[z,x.\theta] : [x.\varphi] \to [y.\psi]$. We define an arrow $c : [x.\varphi] \to [y.\exists x\theta(x,y)]$, which is represented by θ (it is an easy check that θ is also T-provably functional when its codomain becomes $[y.\exists x\theta(x,y)]$). We claim that c is the coequalizer of p_0 and p_1. The composition cp_0 is represented by the same formula as the composition $[z,x.\theta]p_0$, which is $[z,x.\theta]p_1$, and this in turn is represented by the same formula as cp_1. We thus have that $cp_0 = cp_1$. From here on it will be useful to have an explicit description for p_0 and p_1. Using Example 5.2.4 we may assume that

$$\begin{aligned}p_0 &= [x,x',x''.\exists y(\theta(x,y) \wedge \theta(x',y)) \wedge x = x''],\\ p_1 &= [x,x',x''.\exists y(\theta(x,y) \wedge \theta(x',y)) \wedge x' = x''].\end{aligned}$$

Let now $[x,w.\sigma] : [x.\varphi] \to [w.\gamma]$ be any other arrow such that $[x,w.\sigma]p_0 = [x,w.\sigma]p_1$. Then we claim that $[y,w.\exists x(\theta(x,y) \wedge \sigma(y,w))] : [y.\exists x\theta(x,y)] \to [w.\gamma]$ is the unique arrow such that its composition with c is exactly $[x, w.\sigma]$. The hardest thing to check is the third axiom of $\exists x(\theta(x,y) \wedge \sigma(y,w))$ being T-provably functional. Checking the other things (the other two axioms for T-provable functionality, uniqueness and that its composition with c is $[x, w.\sigma]$) are straightforward but tedious to check, so for brevity's sake we will not do so here. We have to show that

$$\exists x(\theta(x,y) \wedge \sigma(x,w)) \wedge \exists x'(\theta(x',y) \wedge \sigma(x',w')) \vdash_{y,w,w'} w = w'$$

is T-provable. For this we reason in the deduction-system for T. We let x and x' be such that

$$\theta(x,y) \wedge \sigma(x,w) \wedge \theta(x',y) \wedge \sigma(x',w'),$$

then we can derive the following two formulas:

$$\exists x''(\exists y(\theta(x,y) \wedge \theta(x',y)) \wedge x = x'' \wedge \sigma(x'',w)),$$
$$\exists x''(\exists y(\theta(x,y) \wedge \theta(x',y)) \wedge x' = x'' \wedge \sigma(x'',w')).$$

These formulas represent the compositions $[x, w.\sigma]p_0$ and $[x, w.\sigma]p_1$ respectively, but those were by assumption the same arrow which means they are T-provably equivalent. So from the second formula we derive that

$$\exists x''(\exists y(\theta(x,y) \wedge \theta(x',y)) \wedge x = x'' \wedge \sigma(x'',w')).$$

Because this formula is T-provably functional we conclude that indeed $w = w'$. □

The regular epi-mono factorization of $f : X \to Y$ is constructed by pulling back f along itself to find its kernel pair. Then the coequalizer e of this kernel pair is the regular epimorphism in the factorization. Since e is the coequalizer of the kernel pair of f, there is a unique arrow from the coequalizer to Y. This is the monomorphism m in the factorization (see also [Oos16, Proposition 4.2]).

In the proof of Proposition 5.2.5 we have thus found an explicit description of the epi-mono factorization of an arrow $[x, y.\theta] : [x.\varphi] \to [y.\psi]$:

$$[x.\varphi] \xrightarrow{[x,y.\theta]} [y.\exists x\theta(x,y)] \xrightarrow{[y,y'.\exists x\theta(x,y)\wedge y=y']} [y.\psi].$$

For the monomorphism we used here that $[y, y'.\exists x(\theta(x,y) \wedge \theta(x,y'))] = [y, y'.\exists x\theta(x,y) \wedge y = y']$. This explicit description will be useful in the following lemma.

Lemma 5.2.6. *Let $[x'.\psi]$ and $[x.\varphi]$ be objects in $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$, with x and x' of the same type. If $\psi(x) \vdash_x \varphi(x)$ is derivable from T, then there is a monomorphism $[x'.\psi] \to [x.\varphi]$ represented by $\theta(x',x) := \psi(x') \wedge x' = x$. Conversely, if $[y, x.\sigma] : [y.\chi] \to [x.\varphi]$ represents a subobject of $[x.\varphi]$ then there is a formula $\psi(x)$ such that $\psi(x) \vdash_x \varphi(x)$ is T-provable and the corresponding monomorphism represents the same subobject as $[y, x.\sigma]$.*

Proof. We will first prove the first part of the lemma. The formula $\theta(x',x)$ defined there is T-provably functional, here we used the fact that $\psi(x) \vdash_x \varphi(x)$ to derive the first and second axiom.

To see that $[x', x.\theta]$ is a monomorphism we let $\alpha(y,x')$ and $\beta(y,x')$ represent two parallel arrows into $[x'.\psi]$, such that $[x', x.\theta][y, x'.\alpha] = [x', x.\theta][y, x'.\beta]$. We will show that $\alpha(y,x') \vdash_{y,x'} \beta(y,x')$, because then we can apply Proposition 5.1.1 to conclude that $[y, x'.\alpha] = [y, x'.\beta]$. Because $\alpha(y,x')$ is T-provably functional we have $\alpha(y,x') \vdash_{y,x'} \psi(x')$. So we have $\alpha(y,x) \vdash_{y,x} \exists x'(\psi(x') \wedge x = x')$ from which we find $\alpha(y,x) \vdash_{y,x} \exists x'(\alpha(y,x') \wedge \theta(x',x))$. The last formula on the right side represents the composition $[x', x.\theta][y, x'.\alpha]$ and so we have $\alpha(y,x) \vdash_{y,x} \exists x'(\beta(y,x') \wedge \theta(x',x))$, from which our desired result follows.

For the converse we consider the image of $[y, x.\sigma]$, which by the discussion above is

$$[x'.\exists y\sigma(y,x')] \xrightarrow{[x',x.\exists y\sigma(y,x')\wedge x=x']} [x.\varphi].$$

Since the regular epi-mono factorization is unique up to isomorphism and $[y, x.\sigma]$ was already a monomorphism, we have that it represents the same subobject as its image. Furthermore, because $[x', x.\exists y\sigma(y,x') \wedge x = x']$ is functional we have

$$\exists y\sigma(y,x') \wedge x = x' \vdash_{x,x'} \exists y\sigma(y,x') \wedge \varphi(x),$$

from which it follows that indeed $\exists y\sigma(y,x) \vdash_x \varphi(x)$. □

Corollary 5.2.7. *If $\chi(x) \vdash_x \psi(x)$ and $\psi(x) \vdash_x \varphi(x)$ are T-provable, then $[x.\chi] \leq [x.\psi]$ as subobjects of $[x.\varphi]$.*

Corollary 5.2.8. *Let $\varphi(x)$ and $\psi(x)$ be two formulas in the same context, then $\varphi(x)$ and $\psi(x)$ are T-provably equivalent if and only if $[x.\varphi]$ and $[x.\psi]$ are isomorphic in $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$.*

Proof. The 'only if' part is Proposition 5.1.3. For the 'if' part we use that $[x.\varphi] \leq [x.\psi]$ and $[x.\psi] \leq [x.\varphi]$ as subobjects (of, for example, $[x.\varphi]$). $\square$

By the previous lemma we may assume that any subobject of an object $[x.\varphi]$ is given by the arrow $[x, x'.\varphi'(x) \wedge x = x']$ with domain some $[x.\varphi']$ in the same context as $[x.\varphi]$. This is useful for simplifying statements about subobjects, like the next lemma.

Lemma 5.2.9. *Let $\{[x.\varphi_i]\}_{i\in I}$ be a family of size $< \kappa$ of subobjects of some object $[x.\varphi]$ in $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$, then their join (in $\mathrm{Sub}([x.\varphi])$) is given by*

$$[x', x.x = x' \wedge \bigvee_{i\in I} \varphi_i(x')] : [x'. \bigvee_{i\in I} \varphi_i(x')] \to [x.\varphi].$$

Proof. By Lemma 5.2.6 we have that $\varphi_i(x) \vdash_x \varphi(x)$ for all $i \in I$. So $\bigvee_{i\in I} \varphi_i(x) \vdash_x \varphi(x)$ and by using Lemma 5.2.6 again we see that the proposed subobject is indeed a subobject.

Since we have $\varphi_i(x) \vdash_x \bigvee_{i\in I} \varphi_i(x)$ for all $i \in I$, we have by Corollary 5.2.7 that indeed $[x.\varphi_i] \leq [x. \bigvee_{i\in I} \varphi_i(x)]$ for all $i \in I$.

Finally, let $[x', x.\psi(x') \wedge x = x'] : [x'.\psi] \to [x.\varphi]$ represent a subobject that is larger than $[x.\varphi_i]$ for all $i \in I$ (again, we used Lemma 5.2.6 to assume the representative to be of this form). Then $\varphi_i(x) \vdash_x \psi(x)$ for all $i \in I$, so $\bigvee_{i\in I} \varphi_i(x) \vdash_x \psi(x)$ which means that $[x. \bigvee_{i\in I} \varphi_i(x)] \leq [x.\psi]$ by Corollary 5.2.7.

Note that in the case I is empty we get the bottom element of $\mathrm{Sub}([x.\varphi])$, namely the one represented by $[x', x.\bot] : [x'.\bot] \to [x.\varphi]$. $\square$

Proposition 5.2.10. *The syntactic category $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$ is a κ-geometric category.*

Proof. We already have established that $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$ is regular (Proposition 5.2.5) and that the subobject posets have joins of size $< \kappa$ (Lemma 5.2.9). So we only need to check that these joins are stable under pullback. For this we will use their explicit description and the explicit description of pullbacks in Example 5.2.4. Let

$$[x, x'.x = x' \wedge \bigvee_{i\in I} \varphi_i(x)] : [x. \bigvee_{i\in I} \varphi_i(x)] \to [x.\varphi]$$

represent a join subobjects of $[x.\varphi]$. Its pullback along $[z, x.\theta(z, x)] : [z.\psi] \to [x.\varphi]$ is then given by

$$[x', z', z.\exists x(\bigvee_{i\in I} \varphi_i(x) \wedge x = x' \wedge \theta(z', x)) \wedge z = z'] : [x', z'.\exists x(\bigvee_{i\in I} \varphi_i(x) \wedge x = x' \wedge \theta(z', x))] \to [z.\psi].$$

By equivalence of formulas we find that the following is isomorphic (and hence also gives the pullback):

$$[x', z', z. \bigvee_{i\in I} (\varphi_i(x') \wedge \theta(z', x')) \wedge z = z'] : [x', z'. \bigvee_{i\in I} (\varphi_i(x') \wedge \theta(z', x))] \to [z.\psi].$$

Then calculating the pullback of $[x', x.\varphi_i(x') \wedge x = x'] : [x'.\varphi_i(x')] \to [x.\varphi]$ along $[z, x.\theta(z, x)]$ gives us

$$[x', z', z.\exists x(\varphi_i(x) \wedge x = x' \wedge \theta(z', x)) \wedge z = z'] : [x', z'.\exists x(\varphi_i(x) \wedge x = x' \wedge \theta(z', x))] \to [z.\psi],$$

which is isomorphic to

$$[x', z', z.\varphi_i(x') \wedge \theta(z', x)) \wedge z = z'] : [x', z'.\varphi_i(x) \wedge \theta(z', x)] \to [z.\psi].$$

So we see that the pullback of the join is isomorphic to the join of the pullbacks, which means that indeed joins are stable under pullback. $\square$

5.3 The universal syntactic model

There is a canonical $\mathscr{L}$-structure U_T in $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$ defined as follows. Let S be some sort and let s be a variable of sort S, then $S^{U_T} = [s.\top]$. In particular, by Lemma 5.2.1, this means that the interpretation of a type X is $[x.\top]$, where x is a list of distinct variables of type X. For a relation symbol R of type X we note that $R(x) \vdash_x \top$ is always T-provable, so by Lemma 5.2.6 we have that there is a monomorphism $[x.R(x)] \rightarrowtail [x.\top]$. We take R^{U_T} to be the subobject represented by this monomorphism. Finally, for a function symbol $f : X \to S$ from type X to sort S one can easily check that the formula $f(x) = s$ is T-provably functional from $X^{U_T} = [x.\top]$ to $S^{U_T} = [s.\top]$, so we let f^{U_T} be $[x, s.f(x) = s]$.

Proposition 5.3.1. *The interpretation $\{x : \varphi(x)\}^{U_T}$ of a formula $\varphi(x)$ in U_T is $[x.\varphi]$. In particular, the $\mathscr{L}$-structure U_T is actually a model of T.*

Proof. The first claim is just by a straightforward induction on the construction of the formula. To see that then U_T is actually a model of T, we let $\varphi \vdash_x \psi$ be a sequent in T. Then by Corollary 5.2.7 we have that $[x.\varphi] \leq [x.\psi]$ as subobjects of $[x.\top]$, and thus $\{x : \varphi(x)\}^{U_T} \leq \{x : \psi(x)\}^{U_T}$ as subobjects of X^{U_T}. □

Proposition 5.3.1 now justifies the following definition.

Definition 5.3.2. The model U_T defined above is called the *universal syntactic model*.

Corollary 5.3.3. *The universal syntactic model U_T has the property that the κ-geometric axioms valid in U_T are precisely those that are derivable from T.*

Proof. Clearly every κ-geometric axiom that is derivable from T is valid in U_T, since it is a model for T and $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$ is sound with respect to the deduction-system for κ-geometric logic. The converse follows from Lemma 5.2.6. □

Let us fix some κ-geometric category $\mathcal{C}$. By Corollary 3.3.6 a κ-geometric functor $F : \mathbf{Syn}^{\mathrm{g}}_\kappa(T) \to \mathcal{C}$ sends U_T to a model $F(U_T)$ of T in $\mathcal{C}$. Conversely we have that any model M in $\mathcal{C}$ corresponds to a κ-geometric functor that sends $[x.\varphi]$ to $\{x : \varphi\}^M$. Similarly, for every arrow in $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$ there is also only one possible arrow to be sent to in $\mathcal{C}$, because they are precisely the arrows that are definable from T. So M gives rise to a functor F_M, such that $F_M(U_T) \cong M$. We thus see that, up to isomorphism, models of T in $\mathcal{C}$ are the same thing as κ-geometric functors $\mathbf{Syn}^{\mathrm{g}}_\kappa(T) \to \mathcal{C}$.

We also note that a homomorphism of models M and N is exactly the same thing as a natural transformation of their corresponding κ-geometric functors. We recall that $\mathbf{Geom}_\kappa$ denotes the category of κ-geometric categories, so by $\mathbf{Geom}_\kappa(\mathcal{C}, \mathcal{D})$ we denote the category of κ-geometric functors $\mathcal{C} \to \mathcal{D}$ and natural transformations between them. We have now essentially established the following proposition.

Proposition 5.3.4. *For κ-geometric $\mathcal{C}$ there is an equivalence of categories*

$$\mathbf{Geom}_\kappa(\mathbf{Syn}^{\mathrm{g}}_\kappa(T), \mathcal{C}) \simeq T\text{–}\mathbf{Mod}(\mathcal{C})$$

that is natural in $\mathcal{C}$. This equivalence is given by sending a κ-geometric functor F to the T-model $F(U_T)$.

Chapter 6

The classifying topos for a geometric theory

In chapter 4 we have defined what a classifying topos for a geometric theory T is. We now have the necessary tools to construct such a classifying topos $\mathbf{Set}[T]$. The last step in doing so is defining the right Grothendieck topology on $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$, and then we can piece everything together.

6.1 The covering topology

Definition 6.1.1. We define J_κ, the *κ-covering (Grothendieck) topology*, on a κ-geometric category $\mathcal{C}$ to be as follows. A sieve R on an object A is covering if R contains a family $\{\alpha_i : B_i \to A\}_{i\in I}$ with $|I| < \kappa$ such that

$$\bigvee_{i\in I} \mathrm{Im}(\alpha_i) = A.$$

We call such a family a *covering family*.

Of course, we need to check that J_κ does indeed define a Grothendieck topology.

Proposition 6.1.2. *In the setting of Definition* 6.1.1*, J_κ does indeed define a Grothendieck topology on $\mathcal{C}$.*

Proof. We shortly recall the three axioms of a Grothendieck topology from Definition 2.1.5.

1. For any object A, the maximal sieve $\max(A)$ is covering.
2. For any R covering A, and any $f : A' \to A$, $f^*(R) = \{g : B \to A' : fg \in R\}$ is covering.
3. Let R be any sieve on A and let S be covering for A such that for every $f : A' \to A$ in S, $f^*(R)$ is covering for A', then R is covering for A.

The first axiom is seen easily to hold, because $Id_A \in \max(A)$ and $\mathrm{Im}(Id_A) = A$.

For the second axiom we let $\{\alpha_i : B_i \to A : i \in I\}$ be a covering family in R. Define $\alpha'_i : B'_i \to A'$ to be the pullback of α_i along f. Note that then $f\alpha'_i = \alpha_i\beta_i$ for some $\beta_i : B'_i \to B_i$, so $\alpha'_i \in f^*(R)$. Since images are stable under pullback, we have that $\mathrm{Im}(\alpha'_i)$ is the pullback of $\mathrm{Im}(\alpha_i)$. Because $\mathcal{C}$ is κ-geometric (and thus joins are stable under pullback) and $|I| < \kappa$ we have that the pullback of $\bigvee_{i\in I} \mathrm{Im}(\alpha_i)$ along f is $\bigvee_{i\in I} \mathrm{Im}(\alpha'_i)$, but $\bigvee_{i\in I} \mathrm{Im}(\alpha_i)$ was just A by assumption and the pullback of A along f is just A'. So we have that $\bigvee_{i\in I} \mathrm{Im}(\alpha'_i) = A'$ and hence that $\{\alpha'_i : B'_i \to A' : i \in I\}$ is a covering family in $f^*(R)$. We therefore conclude that $f^*(R)$ is covering.

For the third axiom we let $\{\alpha_i : B_i \to A : i \in I\}$ be a covering family in S. Then for each $i \in I$ we find a covering family $\{\beta_{ik} : C_{ik} \to B_i : K_i\}$ in $\alpha_i^*(R)$. In particular this means that $\alpha_i\beta_{ik} \in R$ for all $i \in I$ and $k \in K_i$. We claim that $\{\alpha_i\beta_{ik} : i \in I, k \in K_i\}$ is a covering family in R. Clearly the index set is still strictly smaller than κ since I and each K_i are strictly smaller than κ. We thus have to prove that

$$\bigvee_{i\in I}\bigvee_{k\in K_i} \mathrm{Im}(\alpha_i\beta_{ik}) = A.$$

Since $\{\alpha_i : i \in I\}$ was already a covering family, it suffices to prove that $\mathrm{Im}(\alpha_i) \le \bigvee_{k\in K_i} \mathrm{Im}(\alpha_i\beta_{ik})$ for all $i \in I$. Fix i and let $k \in K_i$ we have that β_{ik} factors as

$$\begin{tikzcd} C_{ik} \arrow[r, two heads, "e"] & \mathrm{Im}(\beta_{ik}) \arrow[r, tail, "\iota_k"] & B_i \end{tikzcd}$$

Then $\alpha_i\iota_k$ factors as

$$\begin{tikzcd} \mathrm{Im}(\beta_{ik}) \arrow[r, two heads, "h"] & \mathrm{Im}(\alpha_i\iota_k) \arrow[r, tail, "g"] & A \end{tikzcd}$$

So $\alpha_i\beta_{ik} = \alpha_i\iota_k e = ghe$, which means that g together with he is a factorization of $\alpha_i\beta_{ik}$ into a regular epimorphism and a monomorphism. We thus have that $\mathrm{Im}(\alpha_i\beta_{ik}) = \mathrm{Im}(\alpha_i\iota_k)$, and we have reduced the problem to showing that $\mathrm{Im}(\alpha_i) \le \bigvee_{k\in K_i} \mathrm{Im}(\alpha_i\iota_k)$.

Let P_k be the pullback of α_i along g to obtain a pullback diagram as shown below. Note that the arrow from P_k to B_i is a monomorphism, since g is a monomorphism. So P_k represents a subobject of B_i and we have that $\mathrm{Im}(\beta_{ik}) \le P_k$ as shown below.

$$\begin{tikzcd} \mathrm{Im}(\beta_{ik}) \arrow[drr, tail, bend left, "\iota"] \arrow[dr, dashed] \arrow[ddr, two heads, bend right, "h"'] & & \\ & P_k \arrow[r, tail] \arrow[d] & B_i \arrow[d, "\alpha_i"] \\ & \mathrm{Im}(\alpha_i\iota_k) \arrow[r, tail, "g"'] & A \end{tikzcd}$$

The pullback of $\bigvee_k \mathrm{Im}(\alpha_i\iota_k)$ along α_i is thus $\bigvee_k P_k \ge \bigvee_k \mathrm{Im}(\beta_{ik}) = B_i$. So since joins are stable under pullback we have the following pullback diagram

$$\begin{tikzcd} \bigvee_k P_k \arrow[r, equal] \arrow[d] & B_i \arrow[d, "\alpha_i"] \\ \bigvee_k \mathrm{Im}(\alpha_i\iota_k) \arrow[r, tail] & A \end{tikzcd}$$

In other words, α_i factors through $\bigvee_{k\in K_i} \mathrm{Im}(\alpha_i\iota_k)$ which means that $\mathrm{Im}(\alpha_i) \le \bigvee_{k\in K_i} \mathrm{Im}(\alpha_i\iota_k)$ which is what we had to show. $\square$

Definition 6.1.3. If for a Grothendieck topology J on a category $\mathcal{C}$ all representable presheaves on $\mathcal{C}$ are sheaves, we say that J is *subcanonical*. If J is the largest subcanonical Grothendieck topology on $\mathcal{C}$, we say that J is *canonical*.

Proposition 6.1.4. *The κ-covering Grothendieck topology J_κ is subcanonical.*

Proof. This is essentially proved in [MLM92, Lemma X.5.4], albeit in a somewhat different setting. So we will provide a (not so detailed) proof here that applies to our setting. Checking the details that are left out can be somewhat tedious but should be straightforward.

Let a covering sieve S on some object Y be given. We will show that every compatible family in yZ has an amalgamation. A compatible family for S in yZ consists of arrows $f_s : X_s \to Z$ for each $s : X_s \to Y$ in S, such that for any $g : W \to X_s$ we have $f_s g = f_{sg}$. Let such a compatible family be given. Since S is covering, there is a covering family $\{s_i : i \in I\}$ in S (denote X_i for X_{s_i} and f_i for f_{s_i}). For each $i \in I$, we have a subobject $\mathrm{Im}(\langle s_i, f_i\rangle)$ of $Y \times Z$. These subobjects give rise to a subobject $\bigvee_i \mathrm{Im}(\langle s_i, f_i\rangle)$. We claim that this last subobject is the graph of some arrow $f : Y \to Z$. This arrow f will then be the amalgamation we were looking for.

To see that $\bigvee_i \mathrm{Im}(\langle s_i, f_i\rangle)$ is the graph of some arrow, we reason in the internal language of our category $\mathcal{C}$. We have to show that $\bigvee_i \mathrm{Im}(\langle s_i, f_i\rangle)$ is

(i) total, that is $\vdash_y \exists z \bigvee_i \exists x_i(s_i(x_i) = y \wedge f_i(x_i) = z)$, and

(ii) functional, that is $\bigvee_i \exists x_i(s_i(x_i) = y \wedge f_i(x_i) = z) \wedge \bigvee_i \exists x_i(s_i(x_i) = y \wedge f_i(x_i) = z') \vdash_{y,z,z'} z = z'$.

Let us first prove (i). Let y be arbitrary. Note that since the s_i form a covering family, we have that $\vdash_y \bigvee_i \exists x_i(s_i(x_i) = y)$ (which is the internal variant of saying that $\bigvee_i \mathrm{Im}(s_i) = Y$). So we find x_i be such that $s_i(x_i) = y$. Now take z to be $f_i(x_i)$ and we are done.

For (ii) we let y, z and z' be such that the antecedent is satisfied. We note that the antecedent is equivalent to

$$\bigvee_{i,j} \exists x_i x_j(s_i(x_i) = y \wedge f_i(x_i) = z \wedge s_j(x_j) = y \wedge f_j(x_j) = z').$$

We will show that for all i, j we can deduce $z = z'$ from

$$\exists x_i x_j(s_i(x_i) = y \wedge f_i(x_i) = z \wedge s_j(x_j) = y \wedge f_j(x_j) = z').$$

So let x_i and x_j be such, and consider the pullback

$$\begin{tikzcd} P \arrow[r, "p_i"] \arrow[d, "p_j"'] & X_i \arrow[d, "s_i"] \\ X_j \arrow[r, "s_j"'] & Y \end{tikzcd}$$

Then there must be some p of sort P such that $p_i(p) = x_i$ and $p_j(p) = x_j$. We now conclude that

$$z = f_i(x_i) = f_i p_i(p) = f_j p_j(p) = f_j(x_j) = z',$$

where the middle equality follows from the fact that the f_i are part of a compatible family. $\square$

In particular this means that we can talk about the Yoneda embedding into $\mathbf{Sh}(\mathcal{C}, J_\kappa)$. That is, technically $y : \mathcal{C} \to \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$ does not have codomain $\mathbf{Sh}(\mathcal{C}, J_\kappa)$. If we let $a : \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}} \to \mathbf{Sh}(\mathcal{C}, J_\kappa)$ be the sheafification functor then ay is (isomorphic to) y. So we will omit the a and just talk about the Yoneda embedding.

Proposition 6.1.5. *For κ-geometric $\mathcal{C}$, a functor $F : \mathcal{C} \to \mathcal{E}$ into any topos $\mathcal{E}$ is flat and J_κ-continuous precisely when it is a κ-geometric functor. In symbols, we have:*

$$\mathbf{FlatCon}((\mathcal{C}, J_\kappa), \mathcal{E}) = \mathbf{Geom}_\kappa(\mathcal{C}, \mathcal{E}).$$

Proof. Since $\mathcal{C}$ has all finite limits, a functor $F : \mathcal{C} \to \mathcal{E}$ is left exact precisely when it is flat by Lemma 2.3.10. So, to complete the proof we will show that such F is J_κ-continuous if and only if it preserves regular epimorphisms and joins of size $< \kappa$.

Suppose that F is J_κ-continuous, then we have two characterizations of this. By definition, F takes J_κ-covering sieves to epimorphic families, and the other characterization is that it takes such sieves to colimits. First, let e be a regular epimorphism in $\mathcal{C}$ and let S be the sieve generated by e (i.e. it contains all arrows that factor through e). Clearly S is J_κ-covering, since $\mathrm{Im}(e)$ is (isomorphic to) the codomain of e. So S is sent to an epimorphic family in $\mathcal{E}$. In particular this means that $F(e)$ is epimorphic, but every epimorphism in a topos is a regular epimorphism (see for example [MLM92, Theorem IV.7.8]), so $F(e)$ is a regular epimorphism. Now, let $\{A_i : i \in I\}$, $|I| < \kappa$, be a family of subobjects of some object X in $\mathcal{C}$. Then as subobjects of $F(X)$ we certainly have $F(A_i) \leq F(\bigvee_i A_i)$ for all $i \in I$, and hence $\bigvee_i F(A_i) \leq F(\bigvee_i A_i)$. The sieve generated by the monomorphisms representing the A_i is J_κ-covering for $\bigvee_i A_i$. So it is sent to a colimit in $\mathcal{E}$, and $\bigvee_i F(A_i)$ is a cocone for this same diagram. We conclude that we also have $F(\bigvee_i A_i) \leq \bigvee_i F(A_i)$ and hence $F(\bigvee_i A_i) = \bigvee_i F(A_i)$. We have now shown that F is κ-geometric if it is J_κ-continuous.

For the other direction we let F be κ-geometric and let S be a covering sieve for some object X of $\mathcal{C}$. Then there is a covering family of size $< \kappa$ in S. Since F preserves joins of size $< \kappa$ we see that S is indeed sent to an epimorphic family. $\square$

Proposition 6.1.6. *The Yoneda embedding $y : \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J_\kappa)$ is κ-geometric.*

Proof. In the restricted version of Diaconescu's theorem (Corollary 2.3.9), we can fill in $\mathbf{Sh}(\mathcal{C}, J_\kappa)$ in the role of $\mathcal{E}$ and chase the identity functor on $\mathbf{Sh}(\mathcal{C}, J_\kappa)$:

$$\begin{array}{ccc} i \in \mathbf{Topos}(\mathbf{Sh}(\mathcal{C}, J_\kappa), \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}) & \xrightarrow{\rho} & \mathbf{Flat}(\mathcal{C}, \mathbf{Sh}(\mathcal{C}, J_\kappa)) \ni \rho(i) = ay \\ \uparrow & & \uparrow \\ Id \in \mathbf{Topos}(\mathbf{Sh}(\mathcal{C}, J_\kappa), \mathbf{Sh}(\mathcal{C}, J_\kappa)) & \xrightarrow{\sim} & \mathbf{FlatCon}(\mathcal{C}, \mathbf{Sh}(\mathcal{C}, J_\kappa)) \ni \rho(i) = ay \end{array}$$

where

$$i : \mathbf{Sh}(\mathcal{C}, J_\kappa) \rightleftarrows \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}} : a$$

denotes the inclusion and sheafification functor. So ay (which is just $y : \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J_\kappa)$) is flat and continuous, but by Proposition 6.1.5 that is equivalent to saying that it is κ-geometric. $\square$

6.2 Piecing everything together

Note that a geometric theory is a set of geometric axioms. So for every geometric theory there is a large enough κ such that the theory is κ-geometric.

Theorem 6.2.1. *Let T be a geometric theory, and let κ be such that T is κ-geometric, then the classifying topos $\mathbf{Set}[T]$ is given by $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{g}}_\kappa(T), J_\kappa)$. The generic model G_T is given by $y(U_T)$ where y is the Yoneda embedding.*

Proof. By Proposition 6.1.5 we have that the flat J_κ-continuous functors $\mathbf{Syn}^{\mathrm{g}}_\kappa(T) \to \mathcal{E}$ are precisely the κ-geometric functors. So we obtain (all natural in $\mathcal{E}$)

$$\mathbf{Topos}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}^{\mathrm{g}}_\kappa(T), J_\kappa)) \simeq \mathbf{FlatCon}((\mathbf{Syn}^{\mathrm{g}}_\kappa(T), J_\kappa), \mathcal{E}) = \mathbf{Geom}_\kappa(\mathbf{Syn}^{\mathrm{g}}_\kappa(T), \mathcal{E}) \simeq T\text{–}\mathbf{Mod}(\mathcal{E}),$$

where the first equivalence is given by the continuous version of Diaconescu's theorem (Corollary 2.3.9) and the last equivalence is from Proposition 5.3.4.

The generic model corresponds to the identity on $\mathbf{Set}[T]$, which corresponds under Diaconescu's theorem to the J_κ-continuous flat functor $Id^* y = y$, which corresponds to the model $y(U_T)$ under the last equivalence. $\square$

The explicit description of the generic model allows us to prove some useful facts about it, in the form of the following two propositions.

Proposition 6.2.2. *The generic model has the property that the geometric axioms that are valid in G_T are precisely those that are derivable from T.*

Proof. By soundness every geometric axiom derivable from T holds in G_T. For the other direction we suppose $\varphi \vdash_x \psi$ is some geometric axiom that is is valid in G_T, and let λ be such that it is a λ-geometric axiom. For now, let us denote the universal syntactic model in $\mathbf{Syn}^{\mathrm{g}}_\lambda(T)$ by U^λ_T to make explicit which syntactic category we are looking at. Then consider the model $y(U^\lambda_T)$ of T in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{g}}_\lambda(T), J_\lambda)$, which is indeed a model because the Yoneda embedding is λ-geometric (Proposition 6.1.6). Since $y(U^\lambda_T)$ is a model of T it must be classified by some geometric morphism $f : \mathbf{Sh}(\mathbf{Syn}^{\mathrm{g}}_\lambda(T), J_\lambda) \to \mathbf{Set}[T]$. So $y(U^\lambda_T) \cong f^*(G_T)$, and thus we see that $\varphi \vdash_x \psi$ must also be valid in $y(U^\lambda_T)$. That means that we have

$$y[x.\varphi] = \{x : \varphi(x)\}^{y(U^\lambda_T)} \leq \{x : \psi(x)\}^{y(U^\lambda_T)} = y[x.\psi],$$

so because y is full and faithful we have that $[x.\varphi] \leq [x.\psi]$ (as subobjects of $[x.\top]$). We conclude that $\varphi \vdash_x \psi$ is also valid in U^λ_T and thus by Corollary 5.3.3 it must be derivable from T. $\square$

Proposition 6.2.3. *Let X be a type in the language of T, then every subobject of X^{G_T} is the interpretation of some geometric formula in the language of T. In particular, this formula is always a disjunction of κ-geometric formulas.*

Before we prove this proposition we first note the following. If C is any object in the site $(\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T), J_{\kappa})$, then by Proposition 6.1.4 we have that yC is a sheaf. So using Corollary 2.1.17, we have that the subobjects of yC in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T), J_{\kappa})$ are precisely the closed sieves on C.

Given an arrow $m : D \to C$ in $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$, we get an arrow $ym : yD \to yC$ in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T), J_{\kappa})$ by sending $f : X \to D$ to $mf : X \to C$. This yields the following explicit description for any closed sieve S on D:

$$\exists_{ym}(S) = \{mf : f \in S\}.$$

In particular we have:

$$\exists_{ym}(\max(D)) = \{mf : \mathrm{cod}(f) = D\}.$$

Proof. For a type X, the interpretation X^{G_T} is given by $y[x.\top]$, where x is a string of variables of type X. The subobjects of X^{G_T} in $\mathbf{Set}[T]$ are by the discussion above then exactly the closed sieves on $[x.\top]$. Let S be a closed sieve on $[x.\top]$, we will show that S can be written as the join of representable subobjects in the following sense: we will show that there is a collection $\{D_i\}_{i \in I}$ of subobjects of $[x.\top]$ such that

$$S = \bigvee_{i \in I} yD_i$$

as subobjects of $y[x.\top]$. This would prove the proposition because each object D_i in $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$ is the interpretation in U_T of some κ-geometric formula. So yD_i is then the interpretation of that formula in the generic model G_T (because y is κ-geometric by Proposition 6.1.6). We conclude then that S is the join of the interpretations of κ-geometric formulas, and is thus the interpretation of a geometric formula itself.

To construct such a collection of subobjects of $[x.\top]$, we will first prove the following claim. Let $f : D \to [x.\top]$ be some arrow in S, and write

$$D \xrightarrow{e} \mathrm{Im}(f) \xrightarrow{m} C$$

for its regular epi-mono factorization. We claim that $m \in S$. By definition we have $me = f \in S$, so $e \in m^*(S)$. Since $\mathrm{Im}(e) = \mathrm{Im}(f)$ (as subobjects of $\mathrm{Im}(f)$) we see that $m^*(S)$ is covering, so because S is closed we conclude that indeed $m \in S$.

Using the claim we now see that

$$S = \bigcup_{\substack{m : D \to C \in S \\ m \text{ mono}}} \exists_{ym}(\max(D)),$$

as follows. Clearly $\exists_{ym}(\max(D)) \subseteq S$ for any $m : D \to C$ in S, since S is a sieve. For the other inclusion we let $f \in S$, and let again $D \xrightarrow{e} \mathrm{Im}(f) \xrightarrow{m} C$ be its regular epi-mono factorization. By the claim then $m \in S$, and since $e \in \max(\mathrm{Im}(f))$ we have $f = me \in \exists_{ym}(\max(\mathrm{Im}(f)))$.

As we had seen in the discussion before this proof, $\exists_{ym}(\max(D))$ is the closed sieve representing the subobject yD. So we are left to show that this union actually gives the join of subobjects. Let a denote the sheafification functor from the presheaf category to our sheaf category. Then since a is the inverse image part of a geometric morphism we find (here the equalities stand for equality of subobjects):

$$\begin{aligned}
S &= a(S) && (S \text{ is already a sheaf}) \\
&= a\Big(\bigcup_m \exists_{ym}(\max(D))\Big) && (\text{the equality we proved earlier}) \\
&= a\Big(\bigvee_m \exists_{ym}(\max(D))\Big) && (\text{joins in the presheaf category are given by unions}) \\
&= \bigvee_m a(\exists_{ym}(\max(D))) && (a \text{ preserves joins}) \\
&= \bigvee_m \exists_{ym}(\max(D)) && (\exists_{ym}(\max(D)) \text{ is already a sheaf})
\end{aligned}$$

Note that we have shortened the notation below the union and join symbols for better readability, but this is supposed to be read as "$m : D \to C$ in S, with m a monomorphism". □

Chapter 7

The syntactic category for a first-order theory

The construction of a first-order classifying topos for some first-order theory T will be similar to the construction of a classifying topos for a geometric theory. In particular, we will need a first-order version of the syntactic category.

As in chapter 5, let us fix for the rest of this chapter a theory T in some language $\mathscr{L}$. Only now we allow T to be κ-infinitary first-order. We define $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ exactly as we did with $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$ but now we allow all $\mathscr{L}_\kappa$-formulas and we use the deduction-system for full κ-infinitary first-order logic.

The proofs in chapter 5 up to section 5.3 apply word for word to $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ if we interpret "formula" as "$\mathscr{L}_\kappa$-formula" in the statements. Just to be clear, Theorem 5.1.4 does still only apply to geometric formulas, as is made explicit in its statement. There is in general no such theorem for $\mathscr{L}_\infty$-formulas.

In particular, $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ is a κ-geometric category. It will be no surprise that it is in fact a κ-Heyting category as follows from the following two propositions. Here we can again use Lemma 5.2.6 to assume that a subobject of $[x.\varphi]$ is given by an arrow $[x, x'.\varphi'(x) \wedge x = x']$ with domain $[x.\varphi']$ in the same context as $[x.\varphi]$.

Proposition 7.0.4. *Let $\{[x.\varphi_i]\}_{i\in I}$ be a family of size $< \kappa$ of subobjects of some object $[x.\varphi]$ in $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$, then their meet (in $\mathrm{Sub}([x.\varphi])$) is given by*

$$[x', x.x = x' \wedge \bigwedge_{i\in I} \varphi_i(x')] : [x'. \bigwedge_{i\in I} \varphi_i(x')] \to [x.\varphi].$$

Proof. The proof is analogous to Lemma 5.2.9. □

Proposition 7.0.5. *Let $[x, y.\theta] : [x.\varphi] \to [y.\psi]$ be some arrow in $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$, then $[x, y.\theta]^* : \mathrm{Sub}([y.\psi]) \to \mathrm{Sub}([x.\varphi])$ has a right adjoint $\forall_{[x,y.\theta]} : \mathrm{Sub}([x.\varphi]) \to \mathrm{Sub}([y.\psi])$ given by*

$$[x.\varphi'] \mapsto [y.\forall x(\theta(x, y) \to \varphi'(x))].$$

Proof. Throughout this proof, we denote by $[y.\psi']$ an arbitrary subobject of $[y.\psi]$. Using Example 5.2.4 we can explicitly calculate $[x, y.\theta]^*[y.\psi']$ to be

$$[x.\exists y(\theta(x, y) \wedge \psi'(y))].$$

We have to show that

$$[x.\exists y(\theta(x, y) \wedge \psi'(y))] \leq [x.\varphi'] \iff [y.\psi'] \leq [y.\forall x(\theta(x, y) \to \varphi'(x))].$$

By Lemma 5.2.6 this comes down to showing that

$$\exists y(\theta(x, y) \wedge \psi'(y)) \vdash_x \varphi'(x)$$

is provable in T if and only if

$$\psi'(y) \vdash_y \forall x(\theta(x,y) \to \varphi'(x))$$

is provable in T. So let us show this by reasoning in the deduction-system for T.

($\Longrightarrow$) Assume $\psi'(y)$ and let x be such that $\theta(x,y)$. Then we have $\exists y(\theta(x,y) \wedge \psi'(y))$, so we find $\varphi'(x)$. From this we conclude $\forall x(\theta(x,y) \to \varphi'(x))$, which proves the first direction.

($\Longleftarrow$) Assume now $\exists y(\theta(x,y) \wedge \psi'(y))$. Let y be such that $\theta(x,y) \wedge \psi'(y)$. Then we have $\forall x(\theta(x,y) \to \varphi'(x))$, but because $\theta(x,y)$ we can conclude $\varphi'(x)$, which concludes our proof. $\square$

Corollary 7.0.6. *The category* $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ *is* κ*-Heyting.*

Now that we have established that $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ is κ-Heyting we can again define a canonical $\mathscr{L}$-structure U_T in $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ like we did in section 5.3. We would like to conclude again that U_T is a model of T. To do so, we want to apply the proof of Proposition 5.3.1 in this case and for that we only need that the interpretation $\{x : \varphi(x)\}^{U_T}$ of a $\mathscr{L}_\kappa$-formula is represented by $[x, y.\varphi(x) \wedge x = y] : [x.\varphi] \to [x.\top]$. This was already shown for κ-geometric formulas by induction, so we provide the induction steps for implication, infinite meets and universal quantification.

Lemma 7.0.7. *We have the following facts about the subobjects of* $[x.\top]$ *for some context* x*:*

(i) $\bigwedge_{i\in I}[x.\varphi_i] = [x.\bigwedge_{i\in I}\varphi_i]$*, with* $|I| < \kappa$*;*

(ii) $[x.\forall y\varphi(x,y)] = \forall_\pi[x, y.\varphi(x,y)]$*, where* $\pi : [x, y.\top] \to [x.\top]$ *is the projection;*

(iii) $[x.\varphi] \to [x.\psi] = [x.\varphi \to \psi]$*.*

Proof. Fact (i) is just Proposition 7.0.4. For (ii) we note that by the explicit description in the proof of Lemma 5.2.1 the projection π is given by $[x, y, x'.x = x']$. So by Proposition 7.0.5 we have

$$\forall_\pi[x, y.\varphi(x,y)] = [x'.\forall xy(x = x' \to \varphi(x,y))] = [x.\forall y\varphi(x,y)].$$

Finally, for (iii) we will show that $[x.\chi] \wedge [x.\varphi] \leq [x.\psi]$ if and only if $[x.\chi] \leq [x.\varphi \to \psi]$. Note that $[x.\chi] \wedge [x.\varphi] = [x.\chi \wedge \varphi]$. By Lemma 5.2.6 we have that $[x.\chi \wedge \varphi] \leq [x.\psi]$ if and only if $\chi \wedge \varphi \vdash_x \psi$ is T-provable. The latter is equivalent to $\chi \vdash_x \varphi \to \psi$ being T-provable, which again by Lemma 5.2.6 happens if and only if $[x.\chi] \leq [x.\varphi \to \psi]$. $\square$

Corollary 7.0.8. *The* $\mathscr{L}$*-structure* U_T *in* $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ *is actually a model of* T*, and the* $\mathscr{L}_\kappa$*-formulas that are valid in* U_T *are precisely those that are derivable from* T*.*

Proof. As mentioned before, Lemma 7.0.7 provides us with the missing induction steps to prove that the interpretation $\{x : \varphi(x)\}^{U_T}$ of an $\mathscr{L}_\kappa$-formula is represented by $[x, y.\varphi(x) \wedge x = y] : [x.\varphi] \to [x.\top]$. The second statement is analogous to Corollary 5.3.3. $\square$

Proposition 7.0.9. *For* κ*-Heyting categories* $\mathcal{C}$ *we have an equivalence*

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

natural in $\mathcal{C}$*. This equivalence is given by sending a* κ*-Heyting functor* F *to the model* $F(U_T)$*.*

Proof. This appears as [BJ98, Proposition 2.4], and can be verified in the same way as Proposition 5.3.4. $\square$

7.1 Sheaves on the syntactic category for a first-order theory

In a way similar to chapter 6 we will want to look at the category of sheaves $\mathbf{Sh}(\mathbf{Syn}^{\text{fo}}_\kappa(T), J_\kappa)$. Even though this may not always be a first-order classifying topos for T (as was demonstrated in section 4.1), there are still some results analogous to those in chapter 6 that will be useful. First of all, we can strengthen Proposition 6.1.6 as follows (its proof is based on [BJ98, Lemma 3.1]).

Proposition 7.1.1. *Let $\mathcal{C}$ be a κ-Heyting category, then the Yoneda embedding $y : \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J_\kappa)$ is κ-Heyting.*

Proof. We had already established in Proposition 6.1.6 that the Yoneda embedding is κ-geometric. The Yoneda embedding also preserves all meets of subobjects that exist in $\mathcal{C}$. So all that is left to check is that universal quantification is preserved.

Let $f : A \to B$ be an arrow in $\mathcal{C}$ and let A' be a subobject of A. We define $B' = \forall_f(A')$. We first claim that for any subobject R of yB with $(yf)^*(R) \leq yA'$, we have $R \leq yB'$. To prove this claim, we consider R as a closed sieve on B and prove that every arrow in R factors through B'. Let $g : C \to B$ be an arrow in R, and consider the pullback

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

Then $h \in f^*(R)$, so we must have that h factors through A' because $(yf)^*(R) \leq yA'$. That means that $\mathrm{Im}(h) \leq A'$, and since images are stable under pullback we find

$$f^*(\mathrm{Im}(g)) = \mathrm{Im}(h) \leq A',$$

hence

$$\mathrm{Im}(g) \leq \forall_f(A') = B'.$$

This proves our claim.

Let us now apply the claim with $\forall_{yf}(yA')$ in the role of R, then we have that

$$(yf)^*(\forall_{yf}(yA')) \leq yA' \implies \forall_{yf}(yA') \leq yB'.$$

The antecedent of this implication is just the counit of the adjunction $(yf)^* \dashv \forall_{yf}$, so we conclude that $\forall_{yf}(yA') \leq yB'$. From the counit of the adjunction $f^* \dashv \forall_f$ we have that $f^*(B') = f^*\forall_f(A') \leq A'$, so because y preserves pullbacks:

$$(yf)^*(yB') = y(f^*(B')) \leq yA',$$

and hence $yB' \leq \forall_{yf}(yA')$. We conclude that indeed $y(\forall_f(A')) = y(B') = \forall_{yf}(yA')$. □

This allows us to give a result similar to Proposition 6.2.2. However, we cannot speak about the generic model here because that term is reserved for the model in a (first-order) classifying topos corresponding to the identity morphism. So instead, we will just talk about the embedding of the universal syntactic model.

Proposition 7.1.2. *The $\mathscr{L}$-structure $y(U_T)$ in $\mathbf{Sh}(\mathbf{Syn}^{\text{fo}}_\kappa(T), J_\kappa)$ is a model of T and the $\mathscr{L}_\kappa$-formulas that are valid in $y(U_T)$ are precisely those that are derivable from T.*

Proof. By Proposition 7.1.1, the Yoneda embedding is a κ-Heyting functor. So it follows directly that $y(U_T)$ is a model of T. By soundness this model needs to at least satisfy those formulas that are derivable from T. For the other direction we let $\varphi(x)$ be a valid $\mathscr{L}_\kappa$-formula in $y(U_T)$. Then

$$y[x.\varphi] = \{x : \varphi(x)\}^{y(U_T)} = X^{y(U_T)} = y[x.\top],$$

so because y is full and faithful it follows that $[x.\varphi] = [x.\top]$ (as subobjects of $[x.\top]$, as objects they are just isomorphic). Thus $[x.\varphi]$ is valid in U_T and using Corollary 7.0.8 we conclude that $\varphi(x)$ must be derivable from T. □

If we attempt to apply the same technique as we did in the proof of Theorem 6.2.1, we end up with the following proposition (which also appears as [BJ98, (5) on page 45]).

Proposition 7.1.3. *There is a full and faithfull functor*

$$T\text{–}\mathbf{Mod}(\mathcal{E})_\kappa \to \mathbf{Topos}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), J_\kappa)).$$

This functor is defined as follows: for a model M let F be the κ-Heyting functor $\mathbf{Syn}_\kappa^{\mathrm{fo}}(T) \to \mathcal{E}$ corresponding to it. Then M is sent to the geometric morphism corresponding to F. Thus, for any such model M there is geometric morphism f such that $f^*(y(U_T)) \cong M$.

Proof. Recall from Proposition 7.0.9 that we have an equivalence

$$T\text{–}\mathbf{Mod}(\mathcal{E})_\kappa \simeq \mathbf{Heyt}_\kappa(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), \mathcal{E}).$$

Diaconescu's theorem (Corollary 2.3.9) gives us an equivalence

$$\mathbf{Geom}_\kappa(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), \mathcal{E}) \simeq \mathbf{Topos}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), J_\kappa)),$$

like we have seen in the proof of Theorem 6.2.1. The final step linking these two equivalences is noting that every κ-Heyting functor is also κ-geometric. $\square$

Since $y(U_T)$ is itself a model of T (Proposition 7.1.2) we see that any open geometric morphism $f : \mathcal{E} \to \mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), J_\kappa)$ corresponds to some model $f^*(y(U_T))$ in $\mathcal{E}$. Thus every open geometric morphism is in the image of the functor described in Proposition 7.1.3. However, it is in general not the case that this image only contains open geometric morphisms, nor that an arbitrary geometric morphism f into $\mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), J_\kappa)$ is determined up to isomorphism by the $\mathscr{L}$-structure $f^*(y(U_T))$.

Finally, the proof of Proposition 6.2.3 applies also to the case of $y(U_T)$ in $\mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), J_\kappa)$. For clarity, we will state the proposition again is these terms, but for the proof we refer to Proposition 6.2.3.

Proposition 7.1.4. *Let X be a type in $\mathscr{L}$, then every subobject of $X^{y(U_T)}$ in $\mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), J_\kappa)$ is the interpretation of some $\mathscr{L}_\infty$-formula. In particular, this formula is always a disjunction of $\mathscr{L}_\kappa$-formulas.*

Chapter 8

Completeness theorems

In this chapter we will provide the proof for the completeness theorem that we mentioned in section 3.5. After that, we will also provide a few strengthenings of this theorem that allow us to restrict ourselves to certain topoi instead of all κ-geometric (or κ-Heyting) categories. We recall the statement of Theorem 3.5.7.

Theorem 3.5.7**, repeated.** *Let T be a κ-geometric (κ-infinitary first-order) theory. If for any model M of T in any κ-geometric (κ-Heyting) category we have that σ is valid in M, then σ is T-provable in the deduction-system for κ-geometric (κ-infinitary first-order) logic.*

Proof of Theorem 3.5.7**.** If σ is valid in every model M of T in any κ-geometric category, it is in particular valid in the model U_T in $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$. The valid sequents in this model are exactly those that are T-provable by Corollary 5.3.3, so we conclude that σ must be T-provable. The case for κ-infinitary first-order logic is similar, using $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ instead. $\square$

Theorem 3.5.7 is about completeness concerning all κ-geometric categories (or κ-Heyting categories). We are mostly interested in logic inside topoi, and it turns out that we can actually restrict ourselves to only topoi. We have actually already seen completeness results in previous chapters, where we only had to look at one particular topos. That is, Proposition 6.2.2 tells us that for geometric logic we only need to look at the generic model G_T in the classifying topos $\mathbf{Set}[T]$. For first-order logic we had a somewhat weaker result, namely Proposition 7.1.2 where we had to restrict ourselves to $\mathscr{L}_\kappa$-formulas. However, this allows us to derive the following completeness theorem.

Theorem 8.0.5 (Completeness for topoi)**.** *Let T be an infinitary first-order theory in some language $\mathscr{L}$. Let φ be an $\mathscr{L}_\infty$-formula that is valid in every model of T in every topos, then φ is derivable from T.*

Proof. The formula φ is in particular an $\mathscr{L}_\kappa$-formula for some large enough κ. In this, we may also assume κ to be large enough to express the axioms in T as $\mathscr{L}_\kappa$-formulas. By assumption, φ is valid in every model of T in every topos. In particular it is valid in $y(U_T)$ in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(T), J_\kappa)$. By Proposition 7.1.2 we then conclude that φ is derivable from T. $\square$

We can now already derive a stronger version of completeness, like the one we have for geometric logic in Proposition 6.2.2. This comes at the cost of having to restrict ourselves to those theories that admit a first-order classifying topos.

Theorem 8.0.6. *Let T be an infinitary first-order theory in some language $\mathscr{L}$, such that T has a first-order classifying topos $\mathbf{Set}^{\mathrm{fo}}[T]$. Then the $\mathscr{L}_\infty$-formulas that are valid in G_T in $\mathbf{Set}^{\mathrm{fo}}[T]$ are precisely those that are derivable from T.*

Proof. Since G_T is a model of T, we have by soundness that any $\mathscr{L}_\infty$-formula that is derivable from T must be valid in G_T. For the converse we let $\varphi(x)$ be any $\mathscr{L}_\infty$-formula that is valid in G_T. Let now M be some model of T in some topos $\mathcal{E}$, then since $\mathbf{Set}^{\mathrm{fo}}[T]$ is the first-order classifying topos for T there must be an open geometric morphism $f : \mathcal{E} \to \mathbf{Set}^{\mathrm{fo}}[T]$ corresponding to M. In particular this means that

$$\{x : \varphi(x)\}^M = f^*(\{x : \varphi(x)\}^{G_T}) = f^*(X^{G_T}) = X^M,$$

so we see that $\varphi(x)$ is valid in M as well. Since M and $\mathcal{E}$ were arbitrary, we can apply Theorem 8.0.5 to conclude that $\varphi(x)$ is derivable from T. □

8.1 Completeness for classical logic

For classical logic one would expect that restricting ourselves to Boolean topoi would give a completeness result. In this section we will provide such results. In particular, we obtain similar results for Boolean topoi and classical logic, as we have just seen for general topoi and intuitionistic logic. I have not been able to find the results from this section anywhere else in literature.

Let us first introduce a new kind of logic, that lives between geometric logic and full first-order infinitary logic.

Definition 8.1.1. We call a formula *κ-infinitary sub-first-order* if it is built from κ-geometric formulas and implication.

Note that in κ-infinitary sub-first-order logic negation is allowed, because $\neg\varphi$ is just $\varphi \to \bot$.

The reason for the name is explained by the following proposition.

Proposition 8.1.2. *Let $f : \mathcal{F} \to \mathcal{E}$ be a sub-open geometric morphism, then its inverse image part f^* preserves infinitary sub-first-order logic.*

Proof. In Proposition 4.0.8 we had already seen that f^* is a geometric functor and thus preserves all connectives and existential quantification in any infinitary sub-first-order formula, except for possibly implication. By [Joh80, Lemma 3.1] we have that f is sub-open if and only if f^* preserves implication, and so we conclude that f^* preserves infinitary sub-first-order logic. □

In a classical setting infinitary sub-first-order logic actually has all expressive power that full infinitary first-order logic has. This is made precise in the following lemma.

Lemma 8.1.3. *For every $\mathscr{L}_\kappa$-formula $\varphi(x)$ there is a κ-infinitary sub-first-order formula $\varphi'(x)$ such that $\varphi(x)$ and $\varphi'(x)$ are classically equivalent.*

Proof. Given an $\mathscr{L}_\kappa$-formula $\varphi(x)$ we can replace every occurrence of $\forall y \psi(x, y)$ by $\neg\exists y \neg\psi(x, y)$, where $\psi(x, y)$ is a subformula of $\varphi(x)$. For infinite conjunctions we can replace $\bigwedge_i \psi_i(x)$ by $\neg \bigvee_i \neg\psi_i(x)$. These replacements do not change the formula, up to classical equivalence. □

We can now prove that when taking the double negation sheaves, a model of a sub-first-order theory T will remain a model of that theory. To be more precise, we have the following proposition.

Proposition 8.1.4. *Let $\mathcal{E}$ be a topos, let M be a model of some sub-first-order theory T in $\mathcal{E}$ and let*

$$i : \mathbf{Sh}_{\neg\neg}(\mathcal{E}) \rightleftarrows \mathcal{E} : a$$

be the inclusion of double negation sheaves, then $a(M)$ is a model of T.

Proof. By Proposition 2.2.8 we have that i is sub-open, so by Proposition 8.1.2 a preserves infinitary sub-first-order logic. So any sub-first-order formula that is valid in M, must be valid in $a(M)$. In particular this holds for all sub-first-order formulas in T. □

For the rest of this section a will denote the sheafification for double negation sheaves (in which topos will be clear from the context), and i will denote the corresponding inclusion.

The assumption that T is a sub-first-order theory is crucial in Proposition 8.1.4, as is shown in the following example.

Example 8.1.5. Let T be a theory in a language with one sort and single unary relation symbol A. We have just one axiom in T, namely

$$\neg\forall x(A(x) \vee \neg A(x)).$$

This theory is consistent in intuitionistic logic, so $\mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), J_\kappa)$ is not the trivial topos. Hence $\mathbf{Sh}_{\neg\neg}(\mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), J_\kappa))$ is not the trivial topos. Now suppose that the double negation sheafification $ay(U_T)$ of the model $y(U_T)$ would remain a model of T. Then

$$0 = \{\emptyset : \neg\forall x(A(x) \vee \neg A(x))\}^{ay(U_T)} = 1,$$

where the first equality holds because $\neg\forall x(A(x) \vee \neg A(x))$ is classically false and the second equality holds because we assumed $ay(U_T)$ to be a model of T. Since $\mathbf{Sh}_{\neg\neg}(\mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T), J_\kappa))$ is not the trivial topos, this is a contradiction and we see that $ay(U_T)$ cannot be a model of T.

Nevertheless, we can formulate an analogue of Proposition 7.1.2 for classical logic.

Theorem 8.1.6. *Let T be a first-order theory expressible in $\mathscr{L}_\kappa$. Then there is a Boolean topos $\mathcal{E}$ with a model M_T of T, such that the $\mathscr{L}_\kappa$-formulas valid in M_T are precisely those that are classically derivable from T.*

Proof. Using Lemma 8.1.3 we can replace every first-order formula in T by a classically equivalent sub-first-order formula. Let us denote the sub-first-order theory we obtain in this way by T'. Classically, T and T' are equivalent (every axiom in the one is derivable from the other). So in every Boolean topos, the models of T and T' coincide.

Let us consider $\mathcal{E} = \mathbf{Sh}_{\neg\neg}(\mathbf{Sh}(\mathbf{Syn}_\kappa^{\mathrm{fo}}(T'), J_\kappa))$, then by Proposition 8.1.4 we have that $ay(U_{T'})$ is a model of T' in $\mathcal{E}$. As we have just seen, $ay(U_{T'})$ is also a model of T, and we claim that this is the M_T as described in the statement of this theorem.

By soundness, $ay(U_{T'})$ at least satisfies all formulas that are classically derivable from T. For the converse we let $\varphi(x)$ be some $\mathscr{L}_\kappa$-formula that is valid in $ay(U_{T'})$. Let $\varphi'(x)$ be its sub-first-order equivalent (from Lemma 8.1.3). Then

$$\{x : \neg\neg\varphi'(x)\}^{y(U_{T'})} = \neg\neg\{x : \varphi'(x)\}^{y(U_{T'})}$$

is closed in $\mathrm{Sub}(X^{y(U_{T'})})$ for the $\neg\neg$-topology. We also have

$$\begin{array}{lcr}
a(\{x : \neg\neg\varphi'(x)\}^{y(U_{T'})}) & = & \text{(since } \neg\neg\varphi'(x) \text{ is sub-first-order and } a \text{ is sub-open)} \\
\{x : \neg\neg\varphi'(x)\}^{ay(U_{T'})} & = & \text{(since } \neg\neg\varphi'(x) \leftrightarrow \varphi'(x) \text{ and } \varphi'(x) \leftrightarrow \varphi(x) \text{ hold classically)} \\
\{x : \varphi(x)\}^{ay(U_{T'})} & = & \text{(since } \varphi(x) \text{ is valid in } ay(U_{T'}) \text{ by assumption)} \\
X^{ay(U_{T'})} & = & \\
a(X^{y(U_{T'})}). & &
\end{array}$$

So we can apply Proposition 2.1.15, to obtain $\{x : \neg\neg\varphi'(x)\}^{y(U_{T'})} = X^{y(U_{T'})}$. This means that $\neg\neg\varphi'(x)$ is valid in $y(U_{T'})$. By Proposition 7.1.2 that means that there is an intuitionistic deduction of $\neg\neg\varphi'(x)$ from T'. This deduction is also a classical deduction, and classically we also have $\neg\neg\varphi'(x) \leftrightarrow \varphi'(x)$ and $\varphi'(x) \leftrightarrow \varphi(x)$, so we obtain a classical deduction of $\varphi(x)$ from T', and hence from T. $\square$

We can now also provide an analogue of Theorem 8.0.5.

Corollary 8.1.7 (Completeness for Boolean topoi)**.** *Let T be an infinitary first-order theory in some language $\mathscr{L}$. Let φ be an $\mathscr{L}_\infty$-formula that is valid in every model of T in every Boolean topos, then φ is classically derivable from T.*

Proof. Essentially the same as that of Theorem 8.0.5, only we now consider the model M_T in the topos $\mathcal{E}$, as provided by Theorem 8.1.6. $\square$

In chapter 10 we will also find an analogue of Theorem 8.0.6. That is, we show that given the existence of a first-order classifying topos, we only have to look at a single model in a single topos. The result we obtain is even slightly stronger: we only need the existence of a so-called Boolean classifying topos. What this all means precisely will be discussed in chapter 10.

8.2 Barr's theorem

To round off this chapter, we will mention one more useful consequence of Proposition 6.2.2. It basically states that the deduction-system for classical logic is conservative over κ-geometric logic.

Theorem 8.2.1 (Barr)**.** *Let $\mathcal{E}$ be a topos, then there is a surjective geometric morphism $f : \mathcal{F} \to \mathcal{E}$ where $\mathcal{F}$ satisfies the axiom of choice.*

Proof. For the proof we refer to [Joh77, Theorem 7.57]. $\square$

Corollary 8.2.2. *Let T be a κ-geometric theory, and let σ be a κ-geometric sequent that is T-provable in the deduction-system for classical logic, then σ was already T-provable in the deduction-system for κ-geometric logic.*

Proof. Let $\mathbf{Set}[T]$ be the classifying topos of T, and let G_T be its generic model. By Barr's theorem (Theorem 8.2.1) there is a surjective geometric morphism $f : \mathcal{F} \to \mathbf{Set}[T]$ such that $\mathcal{F}$ satisfies the axiom of choice. Then $f^*(G_T)$ is a model of T in $\mathcal{F}$.

Let $\varphi \vdash_x \psi$ be a κ-geometric sequent that is T-provable in the deduction-system for classical logic. Since $\mathcal{F}$ satisfies the axiom of choice, it is in particular Boolean, so this derivation is sound for $f^*(G_T)$ and hence $\varphi \vdash_x \psi$ is valid in $f^*(G_T)$. That is, we have

$$f^*(\{x : \varphi(x)\}^{G_T}) = \{x : \varphi(x)\}^{f^*(G_T)} \leq \{x : \psi(x)\}^{f^*(G_T)} = f^*(\{x : \psi(x)\}^{G_T}).$$

By [MLM92, Lemma VII.4.3] we have that a geometric morphism is a geometric surjection if and only if it reflects the order of subobjects. So since f is surjective, we have for any two subobjects A, B of some object C in $\mathbf{Set}[T]$ that

$$A \leq B \text{ in } \mathrm{Sub}(C) \iff f^*(A) \leq f^*(B) \text{ in } \mathrm{Sub}(f^*(C)).$$

In particular we have that

$$\{x : \varphi(x)\}^{G_T} \leq \{x : \psi(x)\}^{G_T}.$$

So $\varphi \vdash_x \psi$ is also valid in G_T, which is equivalent to being derivable using the deduction-system for κ-geometric logic (by Proposition 6.2.2). $\square$

Chapter 9

The first-order classifying topos

As we have seen in section 4.1, not every infinitary first-order theory can have a first-order classifying topos. In this chapter we will characterize those theories that have a first-order classifying topos. This result is due to Carsten Butz and Peter Johnstone (see [BJ98]). The problem was that there could be too many inequivalent formulas for a theory. This turns out to be the only problem, so let us introduce some terminology to characterize theories that are 'small enough'.

Definition 9.0.3. A theory T is called *locally small* in a context x if there is a set S_x of formulas in context x, such that every $\mathscr{L}_\infty$-formula in context x is T-provably equivalent to a formula in S_x. We call T *locally small* if it is locally small for every context over its language.

After one more simple definition, we can already state the theorem that characterizes theories with a first-order classifying topos (Theorem 9.0.7). The rest of this chapter is devoted to the proof of this theorem.

Definition 9.0.4. A theory T is called *geometrically saturated* if every $\mathscr{L}_\infty$-formula is T-provably equivalent to some geometric formula in the same context.

Example 9.0.5. An example of a geometrically saturated theory is DLO, the theory of dense linear orders without endpoints. That is, we define DLO in a single-sorted language $\mathscr{L}$ with just one binary relation symbol $<$. The axioms in DLO then express that $<$ is a dense linear order, together with the axioms

$$\forall x \exists y (x < y),$$
$$\forall x \exists y (y < x),$$

which express that $<$ has no endpoints.

In Example 10.5.3, we will see that $\mathbf{Set}[\text{DLO}] = \mathbf{Set}^{\text{fo}}[\text{DLO}]$, and that their generic models are the same. Let $\varphi(x)$ be any $\mathscr{L}_\infty$-formula, then by Proposition 6.2.3 we have that $\{x : \varphi(x)\}^{G_T} = \{x : \psi(x)\}^{G_T}$ for some geometric formula $\psi(x)$. So $\varphi(x)$ and $\psi(x)$ are equivalent in the generic model G_T in $\mathbf{Set}^{\text{fo}}[T]$, and so by Theorem 8.0.6 we conclude that $\varphi(x)$ and $\psi(x)$ are T-provably equivalent.

Proposition 9.0.6. *If T is geometrically saturated, then every homomorphism of models of T is an elementary morphism. So we have*

$$T\text{–}\mathbf{Mod}(\mathcal{E}) = T\text{–}\mathbf{Mod}(\mathcal{E})_\infty$$

for all topoi $\mathcal{E}$.

Proof. Let $\varphi(x)$ be some infinitary first-order formula, and let $h : M \to N$ be a homomorphism of models of T in some topos $\mathcal{E}$. We recall from Definition 3.3.1 that for a type X we denote the arrow $X^M \to X^N$ that belongs to h by h_X. Since T is geometrically saturated, we can find a geometric $\psi(x)$ that is T-provably equivalent to $\varphi(x)$. So we have

$$\{x : \varphi(x)\}^M = \{x : \psi(x)\}^M \leq h_X^*(\{x : \psi(x)\}^N) = h_X^*(\{x : \varphi(x)\}^N),$$

and we conclude that h is indeed an elementary morphism. $\square$

Theorem 9.0.7. *Let T be an infinitary first-order theory, then the following are equivalent:*

(i) T is locally small,

(ii) T is Morita-equivalent to a geometrically saturated theory,

(iii) T has a first-order classifying topos $\mathbf{Set}^{\mathrm{fo}}[T]$ (in the sense of Definition 4.0.12).

Proof outline. The proof of (i) $\implies$ (ii) can be found in section 9.1, while the proof of (ii) $\implies$ (iii) can be found in section 9.4.

Finally, (iii) $\implies$ (i) can be shown here already. By Theorem 8.0.6 we have that the generic model G_T satisfies exactly those $\mathscr{L}_\infty$-formulas that are T-provable. So for a context x of type X we have that every $\mathscr{L}_\infty$-formula in context x is represented by some subobject of X^{G_T}, where two such formulas are T-provably equivalent if and only if they are represented by the same subobject. Since $\mathrm{Sub}(X^{G_T})$ is a set, we have a set of different $\mathscr{L}_\infty$-formulas in context x, up to T-provable equivalence. $\square$

9.1 Morleyization

For a locally small theory T we can define its *Morleyization*, $\mathrm{Mor}(T)$, as follows. Denote by S_x the set of different formulas (up to T-provable equivalence) in context x. Then we extend the language of T by adding, for each context x and each $\varphi \in S_x$, a relation symbol R_φ which has the same type as x. Since there is only a set of different contexts, and each context only has a set of formulas S_x, we only add a set of relation symbols, so our language is still a set. We define $\mathrm{Mor}(T)$ to be T with an axiom $\forall x(R_\varphi(x) \leftrightarrow \varphi(x))$ added for every relation symbol we added to the language (again $\mathrm{Mor}(T)$ is still a set).

Proposition 9.1.1. *Let T be a locally small theory, then $\mathrm{Mor}(T)$ is geometrically saturated. Moreover, given a model of $\mathrm{Mor}(T)$ we can forget its extra structure to obtain a model of T. This operation gives a Morita-equivalence between $\mathrm{Mor}(T)$ and T.*

Proof. This is essentially [BJ98, Lemma 5.2]. Given a formula in the language of $\mathrm{Mor}(T)$ we can replace every relation symbol that we had added to the language with a formula in the language of T. So every formula in the language of $\mathrm{Mor}(T)$ is equivalent to one in the language of T, which in turn is equivalent to some relation symbol. Thus for every formula we can even find an atomic formula equivalent to it, relative to $\mathrm{Mor}(T)$.

To see that the described operation is a Morita-equivalence we note there is a unique way of interpreting the additional relation symbols in any model of T. Similarly, the ∞-elementary morphisms are the same for models of T and $\mathrm{Mor}(T)$. $\square$

Note that we have now proved (i) $\implies$ (ii) of Theorem 9.0.7. There is one more useful link between T and $\mathrm{Mor}(T)$ that we will present as the following proposition.

Proposition 9.1.2. *Let T be a locally small theory, then $\mathrm{Mor}(T)$ is conservative over T. That is, if we denote by $\mathscr{L}$ the language of T, then any $\mathscr{L}_\infty$-formula derivable from $\mathrm{Mor}(T)$ is already derivable from T.*

Proof. Let $\varphi(x)$ be an $\mathscr{L}_\infty$-formula derivable from $\mathrm{Mor}(T)$, then by soundness $\varphi(x)$ is valid in every model of $\mathrm{Mor}(T)$ in every topos. By Proposition 9.1.1 we have that $\mathrm{Mor}(T)$ and T are Morita-equivalent and that this equivalence is given by forgetting the extra structure on models of $\mathrm{Mor}(T)$. That means that $\varphi(x)$ is valid in every model of T in every topos, so by completeness we conclude that $\varphi(x)$ is derivable from T. $\square$

9.2 The theory of open geometric morphisms

For this section, let us fix some site $(\mathcal{C}, J)$ where $\mathcal{C}$ has all finite limits. In Example 3.4.4 we have seen how we can define a theory T in the internal language $\mathscr{L}(\mathcal{C})$, such that a model of T in some topos[1] $\mathcal{E}$ is the same thing as a left exact functor $\mathcal{C} \to \mathcal{E}$. Since we have assumed $\mathcal{C}$ to have all finite limits, the left exact functors and flat functors with domain $\mathcal{C}$ coincide by Lemma 2.3.10.

We can extend the theory T so that its models are actually flat continuous functors. For this we add an axiom

$$\top \vdash_x \bigvee_{f \in S} \exists y_f (f(y_f) = x),$$

for every covering sieve S on X, where we denote the domain of $f \in S$ by Y_f (and y_f is of type Y_f). This ensures that covering sieves are sent to epimorphic families, and so we have now established a geometric theory T' such that we have an equivalence

$$\mathbf{FlatCon}((\mathcal{C}, J), \mathcal{E}) \simeq T'\text{–}\mathbf{Mod}(\mathcal{E})$$

natural in $\mathcal{E}$.

By the continuous version of Diaconescu's theorem (Corollary 2.3.9), we thus see that $\mathbf{Sh}(\mathcal{C}, J)$ is the classifying topos for T':

$$\mathbf{Topos}(\mathcal{E}, \mathbf{Sh}(\mathcal{C}, J)) \simeq \mathbf{FlatCon}((\mathcal{C}, J), \mathcal{E}) \simeq T'\text{–}\mathbf{Mod}(\mathcal{E}).$$

In fact, this yields the following well-known theorem that is interesting in its own right.

Theorem 9.2.1. *Every topos $\mathcal{E}$ is the classifying topos of some geometric theory T.*

Proof. Let $(\mathcal{C}, J)$ be some site such that $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, J)$. By Proposition 2.1.12 we may assume $\mathcal{C}$ to have all finite limits, so we can apply the above construction to obtain the theory T' of flat and continuous functors on $(\mathcal{C}, J)$. As we saw, this gives us that $\mathbf{Sh}(\mathcal{C}, J) \simeq \mathcal{E}$ is the classifying topos for T'. $\square$

The goal of this section will be to extend T' to T'' in such a way that we actually obtain a theory of open geometric morphisms, in the sense that

$$\mathbf{Open}(\mathcal{E}, \mathbf{Sh}(\mathcal{C}, J)) \simeq T''\text{–}\mathbf{Mod}(\mathcal{E})_\infty.$$

Note that in doing so, we will need to allow T'' to become a first-order theory. To do this, we want to characterize which flat continuous functors correspond to open geometric morphisms, under Diaconescu's theorem. We have already seen how to do this, namely in Lemma 2.3.11. The following lemma (which is [BJ98, Corollary 1.2]) is basically a restatement of this result.

Lemma 9.2.2. *Let $F : (\mathcal{C}, J) \to \mathcal{E}$ be a flat continuous functor, then the geometric morphism f corresponding to it under Diaconescu's theorem is open if and only if the following two conditions hold:*

(i) for each object X in $\mathcal{C}$ and each family $\{S_i : i \in I\}$ of closed sieves on X, we have

$$\bigwedge_{i \in I} \bigvee \{\exists_{F(\beta)}(F(Y)) \mid \beta : Y \to X \in S_i\} = \bigvee \{\exists_{F(\beta)}(F(Y)) \mid \beta : Y \to X \in \bigcap_{i \in I} S_i\}$$

as subobjects of $F(X)$ in $\mathcal{E}$;

(ii) for each $\alpha : X \to Y$ in $\mathcal{C}$ and each closed sieve S on X we have

$$\forall_{F(\alpha)} \bigvee \{\exists_{F(\beta)}(F(Z)) \mid \beta : Z \to X \in S\} = \bigvee \{\exists_{F(\gamma)}(F(W)) \mid \gamma : W \to Y \in \forall_\alpha(S)\}$$

as subobjects of $F(Y)$.

Here $\forall_\alpha(S)$ denotes the closed sieve $\{\gamma : W \to Y \mid \forall \beta : W \to X(\alpha\beta = \gamma \implies \beta \in S)\}$.

[1] Actually, that example required the codomain of the functor to be just a geometric category, but we will now be interested in the case where it is a topos.

Proof. The key ingredient is that for each closed sieve S on X we have

$$\bar{\tau}_X(S) = \bigvee \{\exists_{F(\beta)}(F(Y)) \mid \beta : Y \to X \in S\},$$

which we will prove now. Here $\bar{\tau}_X$ is as in Lemma 2.3.11.

We recall that the action of $\bar{\tau}$ on subobjects is given by f^*, the inverse image part of the geometric morphism f, and also that under the correspondence of Diaconescu's theorem (Corollary 2.3.9) $F = f^*ay$, where a denotes the sheafification functor. So since f^* preserves the left adjoint of the pullback functor, we have the following commuting diagram (for $\beta : Y \to X$):

$$\begin{tikzcd}
\mathrm{Sub}(ay(Y)) \arrow[r, "\bar{\tau}_Y"] \arrow[d, "\exists_{ay(\beta)}"'] & \mathrm{Sub}(F(Y)) \arrow[d, "\exists_{F(\beta)}"] \\
\mathrm{Sub}(ay(X)) \arrow[r, "\bar{\tau}_X"'] & \mathrm{Sub}(F(X))
\end{tikzcd}$$

If we let $\beta \in S$, then $\exists_{ay(\beta)}(\max(Y))$ is the closure of the sieve on X containing of all arrows into X that factor through β (here $\max(Y)$ denotes the maximal sieve on Y, regarded as a subobject of $ay(Y)$). So we have

$$S = \bigvee \{\exists_{ay(\beta)}(\max(Y)) \mid \beta : Y \to X \in S\}.$$

Applying $\bar{\tau}_X$ to both sides, and again using that this action is given by f^* which preserves arbitrary joins, we indeed find

$$\begin{aligned}
\bar{\tau}_X(S) &= \bigvee \{\bar{\tau}_X(\exists_{ay(\beta)}(\max(Y))) \mid \beta : Y \to X \in S\} \\
&= \bigvee \{\exists_{F(\beta)}(\bar{\tau}_Y(\max(Y))) \mid \beta : Y \to X \in S\} \\
&= \bigvee \{\exists_{F(\beta)}(F(Y)) \mid \beta : Y \to X \in S\}.
\end{aligned}$$

□

We can translate conditions (i) and (ii) of Lemma 9.2.2 into first-order sentences as follows. For each family $\{S_i : i \in I\}$ of closed sieves on X, condition (i) becomes:

$$\forall x \left(\bigwedge_{i \in I} \bigvee_{j \in J_i} \exists y_j (\beta_j(y_j) = x) \leftrightarrow \bigvee_{k \in K} \exists y_k (\beta_k(y_k) = x) \right), \tag{9.1}$$

where $\{\beta_k : Y_j \to X\}_{j \in J_i}$ is the set of arrows in S_i and $\{\beta_k : Y_k \to X\}_{k \in K}$ is the set of arrows in in $\bigcap_{i \in I} S_i$. For every $\alpha : X \to Y$ and sieve S on X, condition (ii) becomes:

$$\forall y \left(\forall x \left(\alpha(x) = y \to \bigvee_{i \in I} \exists z_i (\beta_i(z_i) = x) \right) \leftrightarrow \bigvee_{j \in J} \exists w_j (\gamma_j(w_j) = y) \right), \tag{9.2}$$

where $\{\beta_i : Z_i \to X\}_{i \in I}$ is the set of arrows in S and $\{\gamma_j : W_j \to Y\}_{j \in J}$ is the set of arrows in $\forall_\alpha(S)$.

By adding the axioms (9.1) and (9.2) to T' to obtain T'' we thus obtain an infinitary first-order theory (in $\mathscr{L}(\mathcal{C})$) such that the models of T'' in some topos $\mathcal{E}$ are essentially open geometric morphisms $\mathbf{Sh}(\mathcal{C}, J) \to \mathcal{E}$. To conclude this section, we wrap up this result in the following proposition.

Proposition 9.2.3. *Let $(\mathcal{C}, J)$ be some site where $\mathcal{C}$ has all finite limits. Then there is an infinitary first-order theory T in the language $\mathscr{L}(\mathcal{C})$, such that*

$$\mathbf{Open}(\mathcal{E}, \mathbf{Sh}(\mathcal{C}, J)) \simeq T\text{–}\mathbf{Mod}(\mathcal{E})_\infty,$$

natural in $\mathcal{E}$.

Proof. The only thing left to justify is the switch to ∞-elementary morphisms on the right side, but this is because natural transformation of open geometric morphisms correspond to ∞-elementary embeddings (Proposition 4.0.13). □

Again, we obtain a result that is of interest in its own right (this also appears as [BJ98, Proposition 4.1]), like Theorem 9.2.1.

Theorem 9.2.4. *Every topos $\mathcal{E}$ is the first-order classifying topos of some infinitary first-order theory T.*

Proof. Just like in Theorem 9.2.1 we take some site $(\mathcal{C}, J)$ for $\mathcal{E}$ such that $\mathcal{C}$ has all finite limits. This time we consider the theory of open geometric morphisms, as we constructed in this section. $\square$

9.3 The first-order classifying topos for an extended theory

The main result of this section will be what is essentially [BJ98, Proposition 4.4]. That is, we will be proving the following proposition.

Proposition 9.3.1. *Let T be an infinitary first-order theory in $\mathscr{L}_\kappa$. Then there exists a cardinal $\lambda > \kappa$ and a theory $\overline{T}$ in $\mathscr{L}_\lambda$ such that we have an equivalence*

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

natural in $\mathcal{E}$.

Moreover, $\overline{T}$ is an $\mathscr{L}_\kappa$-conservative extension of T, and every $\mathscr{L}_\infty$-formula is $\overline{T}$-provably equivalent to a disjunction of $\mathscr{L}_\kappa$-formulas.

Throughout the remainder of this section, fix a T as in the statement of Proposition 9.3.1. We will first describe the construction of $\overline{T}$ and then prove some properties about it that will result in the proof of Proposition 9.3.1.

The idea is to pick λ large enough such that we can let $\overline{T}$ be the theory of open geometric morphisms described in section 9.2. There is a slight issue here: the theory described in that section would be in the internal language $\mathscr{L}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(T))$ of $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$, but we want $\overline{T}$ to be in the language $\mathscr{L}$ (the same language as T). Since every object and arrow in $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ is represented by some $\mathscr{L}_\kappa$-formula, we can rewrite every formula in the language $\mathscr{L}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(T))$ to one in the language of $\mathscr{L}$. For example, if $f : X \to Y$ is an arrow $[x, y.\theta] : [x.\varphi] \to [y.\psi]$ in $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$, then we can rewrite the formula

$$\exists x(f(x) = y)$$

to

$$\exists x\theta(x, y).$$

The syntactic category $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ is small, so there is a set of all its arrows. Let γ denote the cardinality of this set. Since each sieve is just a subset of this set of arrows, there are at most 2^γ sieves on any object in $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$. The indexing set I in (9.1) in section 9.2 has thus at most cardinality 2^γ. The other indexing sets in both (9.1) and (9.2) all have cardinality at most γ. Pick λ strictly bigger than 2^γ and κ, and let $\overline{T}$ be the set of all $\mathscr{L}_\lambda$-sentences that are valid in $y(U_T)$ (in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(T), J_\kappa)$). Then $\overline{T}$ contains every possible instance of (9.1) and (9.2), that is valid in $y(U_T)$.

Lemma 9.3.2. *Let M be a model of T in $\mathcal{E}$, and let f be the geometric morphism corresponding to it (as in Proposition* 7.1.3*). Then M is a model of $\overline{T}$ if and only if f is open.*

Proof. If f is open then since $M \cong f^*(y(U_T))$ and f^* preserves infinitary first-order logic (Proposition 4.0.9), we have that every $\mathscr{L}_\lambda$-formula valid in $y(U_T)$ is also valid in M. So in particular M is a model of $\overline{T}$.

For the converse we note that $y(U_T)$ is a model of the theory of open geometric morphisms as described in section 9.2. So all instances of (9.1) and (9.2) described in that section are valid in $y(U_T)$, which means that they are in $\overline{T}$. Therefore, f is open if M is a model of $\overline{T}$. $\square$

Corollary 9.3.3. *There is a natural equivalence*

$$\mathbf{Open}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}^{\text{fo}}_{\kappa}(T), J_\kappa)) \simeq \overline{T}\text{–}\mathbf{Mod}(\mathcal{E})_\kappa.$$

Proof. By definition T is a subset of $\overline{T}$, so any model of $\overline{T}$ in $\mathcal{E}$ is also a model of T. We thus have that the full and faithful functor

$$T\text{–}\mathbf{Mod}(\mathcal{E})_\kappa \to \mathbf{Topos}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}^{\text{fo}}_{\kappa}(T), J_\kappa))$$

from Proposition 7.1.3 restricts to an equivalence

$$\overline{T}\text{–}\mathbf{Mod}(\mathcal{E})_\kappa \simeq \mathbf{Open}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}^{\text{fo}}_{\kappa}(T), J_\kappa)),$$

where we could restrict the right side of the equivalence to the open geometric morphisms because of Lemma 9.3.2. □

Lemma 9.3.4. *The theory $\overline{T}$ is essentially the full first-order theory of $y(U_T)$, in the sense that any $\mathscr{L}_\infty$-formula that is valid in $y(U_T)$ is derivable from $\overline{T}$.*

Proof. Let φ be an $\mathscr{L}_\infty$-formula valid in $y(U_T)$. Let M be a model of $\overline{T}$ in $\mathcal{E}$, then there is an open geometric morphism $f : \mathcal{E} \to \mathbf{Sh}(\mathbf{Syn}^{\text{fo}}_{\kappa}(T, J_\kappa))$ corresponding to M under the equivalence from Corollary 9.3.3. Since $M \cong f^*(y(U_T))$, we must have that φ is valid in M as well. So φ is valid in every model of $\overline{T}$ is every topos, and thus by completeness (Theorem 8.0.5) φ is derivable from $\overline{T}$. □

Lemma 9.3.5. *The theory $\overline{T}$ is an $\mathscr{L}_\kappa$-conservative extension of T.*

Proof. By Proposition 7.1.2 we have that the $\mathscr{L}_\kappa$-formulas valid in $y(U_T)$ are precisely those that are derivable from T. Since $y(U_T)$ is also a model for $\overline{T}$, we have that any $\mathscr{L}_\kappa$-formula derivable from $\overline{T}$ is valid in $y(U_T)$ and must thus be derivable from T. □

Lemma 9.3.6. *Every $\mathscr{L}_\infty$-formula is $\overline{T}$-provably equivalent to a disjunction of $\mathscr{L}_\kappa$-formulas.*

Proof. Let $\varphi(x)$ be an $\mathscr{L}_\infty$-formula. Then its interpretation $\{x : \varphi(x)\}^{y(U_T)}$ is a subobject of $X^{y(U_T)}$, and is thus the interpretation of some disjunction of $\mathscr{L}_\kappa$-formulas by Proposition 7.1.4. Let us call this disjunction $\psi(x)$. Then by Lemma 9.3.4 we have that the equivalence of $\varphi(x)$ and $\psi(x)$ is derivable from $\overline{T}$. □

We can now complete the main proof of this section.

Proof of Proposition 9.3.1**.** We have essentially seen everything from the statement of the proposition in this section already. The conservativity of $\overline{T}$ is Lemma 9.3.5 and the last claim of the proposition is just Lemma 9.3.6.

All that remains is to check the equivalence. From Corollary 9.3.3 we have

$$\mathbf{Open}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}^{\text{fo}}_{\kappa}(T), J_\kappa)) \simeq \overline{T}\text{–}\mathbf{Mod}(\mathcal{E})_\kappa,$$

which is almost what we want. The only difference is that we have $\overline{T}\text{–}\mathbf{Mod}(\mathcal{E})_\kappa$ on the right side of the equivalence. However, every $\mathscr{L}_\infty$-formula is $\overline{T}$-provably equivalent to a disjunction of $\mathscr{L}_\kappa$-formulas, so we directly have that every κ-elementary morphism is also ∞-elementary. We thus find $\overline{T}\text{–}\mathbf{Mod}(\mathcal{E})_\kappa = \overline{T}\text{–}\mathbf{Mod}(\mathcal{E})_\infty$, and we are done. □

Corollary 9.3.7. *Let T be an infinitary first-order theory. Then if T is equivalent to $\overline{T}$, in the sense that every formula of $\overline{T}$ can be derived from T (and vice versa), then T has a first-order classifying topos.*

Proof. This appears as [BJ98, Corollary 4.5]. If $\overline{T}$ and T are equivalent, then every model of T is also a model of $\overline{T}$ (the converse is already trivially true). So by Proposition 9.3.1 we have

$$\mathbf{Open}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}^{\text{fo}}_{\kappa}(T), J_\kappa)) \simeq \overline{T}\text{–}\mathbf{Mod}(\mathcal{E})_\infty = T\text{–}\mathbf{Mod}(\mathcal{E})_\infty,$$

and we may take $\mathbf{Sh}(\mathbf{Syn}^{\text{fo}}_{\kappa}(T), J_\kappa)$ to be the first-order classifying topos of T. □

9.4 Finishing the theorem

We are now ready to finish the proof of Theorem 9.0.7 by proving that (ii) implies (iii). That is, if T is Morita-equivalent to a geometrically saturated theory, then there is a first-order classifying topos of T. We will give a direct proof of this fact, unlike [BJ98], where a few more tools are developed. In particular, we fully skip [BJ98, Corollary 4.6] and the discussion around it by providing a more proof-theoretic argument.

Proof of Theorem 9.0.7**, (ii) $\implies$ (iii).** If two theories are Morita-equivalent, then one has a first-order classifying topos if and only if the other does, and this is the same topos. So we may as well assume T to be geometrically saturated.

By Theorem 5.1.4 we can find κ large enough such that every geometric formula in the language of T is provably equivalent to a κ-geometric formula. We also make sure that κ is large enough to express T in $\mathscr{L}_\kappa$. Let $\overline{T}$ and λ be as in Proposition 9.3.1. We will aim to show that $\overline{T}$ and T are equivalent, so that we can apply Corollary 9.3.7.

Recall that by definition $\overline{T}$ is the set of all $\mathscr{L}_\lambda$-sentences that are true in the model $y(U_T)$ in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\kappa(T), J_\kappa)$. Thus by definition, T is a subset of $\overline{T}$ and we only need to show that every formula in $\overline{T}$ is derivable from T. So let φ be a formula in $\overline{T}$. Then, because T is geometrically saturated, φ is T-provably equivalent to some geometric formula ψ, which we may assume to be κ-geometric. Then ψ is valid in $y(U_T)$, and by Proposition 7.1.2 we have that ψ is derivable from T. Therefore φ is derivable from T.

So we conclude that T and $\overline{T}$ are equivalent and thus by applying Corollary 9.3.7 we have that T has a first-order classifying topos. □

Chapter 10

The Boolean classifying topos

In this chapter we introduce a new concept, namely that of a Boolean classifying topos. We show that there is a necessary and sufficient condition on a first-order theory to have such a Boolean classifying topos, and we show how to construct such a topos given a theory satisfying this condition (Theorem 10.1.8). The results in this chapter do not appear anywhere else in literature, to the best of my knowledge.

10.1 Definition of a Boolean classifying topos

Let us for now consider a sub-first-order theory T that has a first-order classifying topos $\mathbf{Set}^{\mathrm{fo}}[T]$. Denote by

$$i : \mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{\mathrm{fo}}[T]) \rightleftarrows \mathbf{Set}^{\mathrm{fo}}[T] : a$$

the inclusion of the double negation sheaves and the corresponding sheafification functor. We have seen in Proposition 8.1.4 that the sheafification of the generic model $a(G_T)$ in $\mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{\mathrm{fo}}[T])$ is again a model for T. Let now $\mathcal{E}$ be any Boolean topos, then any model M of T in $\mathcal{E}$ must be classified by some open geometric morphism $f : \mathcal{E} \to \mathbf{Set}^{\mathrm{fo}}[T]$. We can thus apply Corollary 2.2.10 to obtain a factorization of f as

$$\mathcal{E} \xrightarrow{h} \mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{\mathrm{fo}}[T]) \xrightarrow{i} \mathbf{Set}^{\mathrm{fo}}[T]$$

where h is an open geometric morphism. So $M \cong f^*(G_T) = h^*(a(G_T))$. Conversely, a geometric morphism $h : \mathcal{E} \to \mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{\mathrm{fo}}[T])$ is automatically open, because $\mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{\mathrm{fo}}[T])$ is Boolean (Proposition 2.2.9). So $h^*(a(G_T))$ is a model of T in $\mathcal{E}$. It looks like $\mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{\mathrm{fo}}[T])$ is a Boolean classifying topos for T in the following sense.

Definition 10.1.1. Let T be an infinitary first-order theory, then the *Boolean classifying topos* $\mathbf{Set}^{\mathrm{b}}[T]$ of T, if it exists, is the Boolean topos such that there is an equivalence of categories

$$\mathbf{Topos}(\mathcal{E}, \mathbf{Set}^{\mathrm{b}}[T]) = \mathbf{Open}(\mathcal{E}, \mathbf{Set}^{\mathrm{b}}[T]) \simeq T\text{–}\mathbf{Mod}(\mathcal{E})_\infty,$$

where $\mathcal{E}$ is Boolean. Furthermore, this equivalence should be natural in $\mathcal{E}$. Like we have seen in chapter 4, this topos contains a generic model G_T such that the model corresponding to $f : \mathcal{E} \to \mathbf{Set}^{\mathrm{b}}[T]$ is given by $f^*(G_T)$.

There are a few subtle things in this definition that are worth noting. First of all, we talk about 'the' topos because if such a Boolean classifying topos exists, it must be unique up to equivalence. Another thing is that $\mathbf{Topos}(\mathcal{E}, \mathbf{Set}^{\mathrm{b}}[T]) = \mathbf{Open}(\mathcal{E}, \mathbf{Set}^{\mathrm{b}}[T])$ appears in the definition. This is because of Proposition 2.2.9, which says that geometric morphisms and open geometric morphisms are the same since $\mathbf{Set}^{\mathrm{b}}[T]$ is Boolean. This also explains why we look at ∞-elementary embeddings on the right hand side, instead of just homomorphisms. Finally, we have to include the phrase "if it exists" once more, because not every first-order theory may have a Boolean classifying topos in this sense. In section 10.2 we will see that there may be issues with size, just like in the case of first-order classifying topoi.

Even though it looks like $\mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{\mathrm{fo}}[T])$ is a good candidate for such a Boolean classifying topos, this is in general not the case. For the above construction to work we needed T to be sub-first-order. The problem here is that open geometric morphisms into $\mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{\mathrm{fo}}[T])$ correspond to sub-open geometric morphisms into $\mathbf{Set}^{\mathrm{fo}}[T]$. This problem can be solved by looking at the Boolean core of a topos. So we take the following definition from [Joh80, page 22].

Definition 10.1.2. We recall that an *open subtopos* of a topos $\mathcal{E}$ is one of the form $\mathcal{E}/U$, for U a subobject of the terminal object in $\mathcal{E}$. The *Boolean core* of a topos $\mathcal{E}$ is the open subtopos $\mathcal{E}/U$ for the largest U, such that $\mathcal{E}/U$ is included in $\mathbf{Sh}_{\neg\neg}(\mathcal{E})$. We denote the Boolean core of $\mathcal{E}$ by $\mathcal{B}(\mathcal{E})$.

By definition $\mathcal{B}(\mathcal{E})$ is a subtopos of $\mathbf{Sh}\neg\neg(\mathcal{E})$, so by Proposition 2.1.22 we have that $\mathcal{B}(\mathcal{E})$ is Boolean.

To see that the Boolean core of a topos is really different from the topos of double negation sheaves, we consider the following example.

Example 10.1.3. Let us fix some monoid M and view it as a category. That is, the category M has one object, an arrow for each element of M and composition is given by the monoid operation. We will consider the topos $\mathbf{Set}^{M^{\mathrm{op}}}$.

As explained in [MLM92, pages 35 and 274], the subobject classifier Ω of $\mathbf{Set}^{M^{\mathrm{op}}}$ is given by the set of right ideals on M (a right ideal on M is a subset R of M, such that for all $r \in R$ and $m \in M$ we have $r \cdot m \in R$). The action of $m \in M$ on Ω is then given by sending a right ideal R to $R \cdot m := \{h \in M : m \cdot h \in R\}$.

The subobjects of the terminal object 1 correspond to their classifying arrows $1 \to \Omega$. Such a classifying arrow is then just a right ideal R such that $R \cdot m = R$ for all $m \in M$. So if R is non-empty, it must contain some $m \in M$ and then $R = R \cdot m = M$. So there are exactly two such right ideals: $\emptyset$ and M itself. We thus see that 1 has exactly two subobjects (i.e. $\mathbf{Set}^{M^{\mathrm{op}}}$ is two-valued).

So the Boolean core $\mathcal{B}(\mathbf{Set}^{M^{\mathrm{op}}})$ of $\mathbf{Set}^{M^{\mathrm{op}}}$ is either $\mathbf{Set}^{M^{\mathrm{op}}}/1$ or $\mathbf{Set}^{M^{\mathrm{op}}}/0$. The former is just $\mathbf{Set}^{M^{\mathrm{op}}}$, and the latter is the trivial topos. The topos $\mathbf{Set}^{M^{\mathrm{op}}}$ is Boolean if and only if M is a group. So $\mathcal{B}(\mathbf{Set}^{M^{\mathrm{op}}}) = \mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{M^{\mathrm{op}}}) = \mathbf{Set}^{M^{\mathrm{op}}}$ if M is a group, but if M is not a group then we have that $\mathcal{B}(\mathbf{Set}^{M^{\mathrm{op}}})$ is the trivial topos, and thus that $\mathcal{B}(\mathbf{Set}^{M^{\mathrm{op}}}) \neq \mathbf{Sh}_{\neg\neg}(\mathbf{Set}^{M^{\mathrm{op}}})$.

We can give a slightly more explicit description of the Boolean core, in terms of the subobjects of Ω in $\mathcal{E}$.

Proposition 10.1.4. *Let $\mathcal{E}$ be a topos and let U be the subterminal object such that $\mathcal{E}/U$ is the Boolean core of $\mathcal{E}$. Then U is the largest subterminal object, such that*

$$U \leq \bigwedge_{P \in \mathrm{Sub}(\Omega)} \neg\neg P \to P.$$

Proof. By definition, U is the largest subterminal object such that $\mathcal{E}/U$ is included in $\mathbf{Sh}_{\neg\neg}(\mathcal{E})$. We recall from [Joh02a, Section A4.5] that $\mathcal{E}/U$ is equivalent to the subtopos of $\mathcal{E}$ corresponding to the topology J_U, which is given by $U \to (-)$ as Lawvere-Tierney topology. So U is the largest subterminal object such that $J_{\neg\neg} \leq J_U$, which is equivalent to saying that $\neg\neg P \leq U \to P$ for all $P \in \mathrm{Sub}(\Omega)$. This is precisely the case when $U \leq \neg\neg P \to P$ for all $P \in \mathrm{Sub}(\Omega)$. So we find indeed that U is the largest subterminal object such that

$$U \leq \bigwedge_{P \in \mathrm{Sub}(\Omega)} \neg\neg P \to P.$$

$\square$

There is an interesting link between the Boolean core of some topos $\mathcal{E}$ and open geometric morphisms into $\mathcal{E}$.

Proposition 10.1.5. *Let $\mathcal{F}$ be a Boolean topos, then a geometric morphism $f : \mathcal{F} \to \mathcal{E}$ is open if and only if f factors through the Boolean core of $\mathcal{E}$.*

Proof. This is precisely [Joh80, Proposition 3.6(ii)], but we will give a short proof here anyway. First we note that by definition, $\mathcal{B}(\mathcal{E})$ is an open subtopos of $\mathcal{E}$. By [Joh80, Lemma 1.4] this means that the inclusion $i : \mathcal{B}(\mathcal{E}) \to \mathcal{E}$ is open.

The image of an open geometric morphism $f : \mathcal{F} \to \mathcal{E}$ is an open subtopos of $\mathcal{E}$, by [Joh80, Lemma 1.5]. We also know that f has to factor through $\mathbf{Sh}_{\neg\neg}(\mathcal{E})$ by Proposition 2.2.8, so f has to factor through $\mathcal{B}(\mathcal{E})$.

For the converse, we suppose that a geometric morphism f factors through $\mathcal{B}(\mathcal{E})$ as

$$\mathcal{F} \xrightarrow{h} \mathcal{B}(\mathcal{E}) \xrightarrow{i} \mathcal{E}.$$

Then h is open by Proposition 2.2.9, because $\mathcal{B}(\mathcal{E})$ is Boolean, and i is open because $\mathcal{B}(\mathcal{E})$ is an open subtopos. So their composition, which is f, is open. □

Using the concept of the Boolean core, we can show that a theory T at least has a Boolean classifying topos if it has a first-order classifying topos.

Proposition 10.1.6. *Let T be an infinitary first-order theory that has a first-order classifying topos $\mathbf{Set}^{\mathrm{fo}}[T]$. Then $\mathcal{B}(\mathbf{Set}^{\mathrm{fo}}[T])$ is the Boolean classifying topos for T, and its generic model is given by $a(G_T)$. Here $a \dashv i$ denotes the sheafification functor and the inclusion for the Boolean core.*

Proof. Throughout this proof, we let $\mathcal{E}$ denote an arbitrary Boolean topos. The operation

$$\begin{gathered}\mathbf{Topos}(\mathcal{E}, \mathcal{B}(\mathbf{Set}^{\mathrm{fo}}[T])) \to \mathbf{Open}(\mathcal{E}, \mathbf{Set}^{\mathrm{fo}}[T]),\\ f \mapsto if,\end{gathered}$$

is well-defined and (essentially) surjective by Proposition 10.1.5. Furthermore, since i is full and faithful, natural transformations $f \to f'$ correspond one-to-one with natural transformations $if \to if'$. We thus have that

$$\mathbf{Topos}(\mathcal{E}, \mathcal{B}(\mathbf{Set}^{\mathrm{fo}}[T])) \simeq \mathbf{Open}(\mathcal{E}, \mathbf{Set}^{\mathrm{fo}}[T]) \simeq T\text{–}\mathbf{Mod}(\mathcal{E})_\infty,$$

and we conclude that $\mathcal{B}(\mathbf{Set}^{\mathrm{fo}}[T])$ is the Boolean classifying topos for T. The generic model in $\mathcal{B}(\mathbf{Set}^{\mathrm{fo}}[T])$ corresponds to the identity morphism, so under the above equivalence this evaluates to $a(G_T)$. □

We can now use Theorem 9.0.7 to see that all locally small theories have a Boolean classifying topos. So this condition is sufficient, but it will turn out to be not necessary. There is a similar condition that is necessary and sufficient. That is when we consider the property of being locally small relative to the classical deduction-system. To make this precise we have the following definition.

Definition 10.1.7. A theory T is called *classically locally small* in a context x if there is a set S_x of formulas in context x, such that every $\mathscr{L}_\infty$-formula in context x is classically T-provably equivalent to a formula in S_x. We call T *classically locally small* if T is classically locally small for every context over its language.

In section 10.2 we will see very concrete examples of why not every theory should satisfy this and why this is a weaker condition that just being locally small.

We can now look at the main result of this chapter, we will prove this result in section 10.4.

Theorem 10.1.8. *An infinitary first-order theory T has a Boolean classifying topos $\mathbf{Set}^{\mathrm{b}}[T]$ if and only if T is classically locally small.*

At the end of section 8.1 we promised to provide a result for classical logic, that is similar to Theorem 8.0.6. Now that we have the notion of a Boolean classifying topos, we can actually give this result.

Corollary 10.1.9. *Let T be an infinitary first-order theory in some language $\mathscr{L}$, such that T has a Boolean classifying topos $\mathbf{Set}^{\mathrm{b}}[T]$. Then the $\mathscr{L}_\infty$-formulas that are valid in its generic model G_T are precisely those that are classically derivable from T.*

Proof. Since G_T is a model of T in a Boolean topos, it satisfies at least all $\mathscr{L}_\infty$-formulas that are classically derivable from T. For the converse we will want to apply the completeness theorem for classical logic (Corollary 8.1.7), so let M be any model in some boolean topos $\mathcal{E}$ and let $\varphi(x)$ be some $\mathscr{L}_\infty$-formula that is valid in G_T. Then M is classified by an open geometric morphism $f : \mathcal{E} \to \mathbf{Set}^{\mathrm{b}}[T]$, and we have

$$\{x : \varphi(x)\}^M = f^*(\{x : \varphi(x)\}^{G_T}) = f^*(X^{G_T}) = X^M.$$

So $\varphi(x)$ must be valid in all models of T in every Boolean topos $\mathcal{E}$, and we conclude that $\varphi(x)$ must be derivable from T in the classical deduction-system. □

10.2 The problem with a Boolean classifying topos

In section 4.1 we had constructed a theory T that cannot have a first-order classifying topos. For this we used a language $\mathscr{L}$ with just two propositional variables P and Q and we let T be the empty theory. We then proceeded to show that T is not locally small. However, this T is classically locally small, since the free Boolean algebra on two generators is finite and hence complete. We thus see that being classically locally small really is a weaker condition than being locally small.

To give an example of a theory that is not classically locally small, we can essentially use the same argument as we did in section 4.1, only now we let our language have a countable infinity of propositional variables. By [Sol66, Theorem 1] we then have for every κ a complete Boolean algebra B_κ of size at least κ, on countably many generators. We note that to repeat the argument from section 4.1, one has to use the fact that $\mathbf{Sh}(B_\kappa)$ is Boolean, because B_κ is a complete Boolean algebra.

10.3 Syntactic category for classical logic

For this section we consider the category $\mathbf{Syn}^{\mathrm{c}}_\kappa(T)$, which is like $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$, but now using the deduction-system for classical logic. All the proofs and properties for $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$ in chapter 7 also apply to $\mathbf{Syn}^{\mathrm{c}}_\kappa(T)$, with the exception of Proposition 7.0.9 and its consequence Proposition 7.1.3. In particular, $\mathbf{Syn}^{\mathrm{c}}_\kappa(T)$ is a κ-Heyting category and contains a universal model U_T of T. In this case, the $\mathscr{L}_\kappa$-formulas that are valid in U_T are precisely those that are derivable from T in the deduction-system for classical logic.

The reason that Proposition 7.0.9 does not hold, is because we used the classical deduction-system. So we cannot have that every model in every κ-Heyting category $\mathcal{C}$ corresponds to some κ-Heyting functor $\mathbf{Syn}^{\mathrm{c}}_\kappa(T) \to \mathcal{C}$. When we restrict ourselves to categories where the classical deduction-system is sound, we do find a version of Proposition 7.0.9. We are primarily interested in topoi, so we will consider Boolean topoi. This gives us the following proposition.

Proposition 10.3.1. *For a Boolean topos $\mathcal{E}$ we have an equivalence*

$$\mathbf{Heyt}_\kappa(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), \mathcal{E}) \simeq T\text{–}\mathbf{Mod}(\mathcal{E})_\kappa,$$

natural in $\mathcal{E}$. This equivalence is given by sending a κ-Heyting functor F to the model $F(U_T)$.

For the rest of this section, fix a classically locally small infinitary first-order theory T in some language $\mathscr{L}$. Denote by S_x the set of different formulas (up to classical T-provable equivalence) in context x. Let κ be a regular cardinal strictly larger than the cardinalities of all S_x[1] and such that T is expressible in $\mathscr{L}_\kappa$. Also, for the rest of this section we will only use the classical deduction-system. So now "equivalent formulas", "T-provable", and so on, are meant relative the classical deduction-system.

Our next goal will be to show that $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$ is a Boolean topos. To do so, we will show that the κ-covering topology J_κ coincides with the double negation topology $J_{\neg\neg}$. However, this may give a little trouble since there may be empty covering sieves in J_κ. We note that by the definition of J_κ an object $[x.\varphi]$ of $\mathbf{Syn}^{\mathrm{c}}_\kappa(T)$ admits an empty covering sieve precisely when φ is inconsistent with T. Objects with an empty covering sieve can be left out and the resulting category of sheaves will be equivalent to the original one (as follows directly from a simple application of the Comparison Lemma, see [Joh02b, Theorem 2.2.3]). This inspires the following definition.

[1]In fact, we just need κ to be such that in every context there are $< \kappa$ different $\mathscr{L}_\kappa$-formulas in that context, up to classical T-provable equivalence.

Definition 10.3.2. We define $\mathbf{SynCons}^{\mathrm{c}}_{\kappa}(T)$ to be the full subcategory of $\mathbf{Syn}^{\mathrm{c}}_{\kappa}(T)$ consisting of those objects $[x.\varphi]$ such that φ is consistent with T (equivalently: $[x.\varphi]$ is not isomorphic to $[x.\bot]$). We also denote by J_κ the topology on $\mathbf{SynCons}^{\mathrm{c}}_{\kappa}(T)$ that is induced by the topology J_κ on $\mathbf{Syn}^{\mathrm{c}}_{\kappa}(T)$.

So now we are just left to show that J_κ coincides with $J_{\neg\neg}$ as topologies on $\mathbf{SynCons}^{\mathrm{c}}_{\kappa}(T)$. We will do this in the following two lemmas.

Lemma 10.3.3. *We have $J_\kappa \subseteq J_{\neg\neg}$ as topologies on $\mathbf{SynCons}^{\mathrm{c}}_{\kappa}(T)$.*

Proof. Let S be a covering sieve on $[y.\varphi]$ in J_κ and let $[x, y.\theta] : [x.\psi] \to [y.\varphi]$ be an arrow. By definition, there is a covering family $\{[x_i, y.\sigma_i] : [x_i.\chi_i] \to [y.\varphi]\}_{i\in I}$ in S. This family is non-empty, because $\varphi(x)$ is consistent with T.

We claim that there is an $i \in I$ such that the pullback of $[x, y.\theta]$ along $[x_i, y.\sigma_i]$ exists. Suppose that this is not the case, then for each $i \in I$ the formula

$$\exists y(\theta(x, y) \wedge \sigma_i(x_i, y)),$$

which would be the pullback according to Example 5.2.4, is inconsistent with T. Then

$$\bigvee_{i\in I} \theta(x, y) \wedge \sigma_i(x_i, y)$$

is inconsistent with T, and this formula is equivalent to

$$\theta(x, y) \wedge \bigvee_{i\in I} \sigma_i(x_i, y).$$

Then we must have that

$$\theta(x, y) \wedge \bigvee_{i\in I} \exists x_i \sigma_i(x_i, y)$$

is inconsistent with T, but this formula is equivalent to $\theta(x, y)$ since the σ_i form a covering family. Then $\psi(x)$ cannot be consistent with T, because θ is T-provably functional which means that $\psi(x) \vdash_x \exists y\theta(x, y)$, and so we have reached a contradiction.

We have thus found that there must be an $i \in I$ such that the pullback of $[x, y.\theta]$ and $[x_i, y.\sigma_i]$ exists. This gives us an arrow into $[x.\psi]$ such that its composition with $[x, y.\theta]$ factors through $[x_i, y.\sigma_i]$, and thus this composition is in S. We conclude that S is also a covering sieve in $J_{\neg\neg}$. $\square$

Lemma 10.3.4. *We have $J_{\neg\neg} \subseteq J_\kappa$ as topologies on $\mathbf{SynCons}^{\mathrm{c}}_{\kappa}(T)$.*

Proof. Let S be a covering sieve on $[y.\varphi]$ in $J_{\neg\neg}$ and denote its arrows by $[x_i, y.\theta_i]$. Then because T is classically locally small and by the choice of κ, there is $I_0 \subseteq I$ such that $|I_0| < \kappa$ and for all $i \in I$ there is $i_0 \in I_0$ such that $\exists x_i\theta_i(x_i, y)$ is T-provably equivalent to $\exists x_{i_0}\theta_{i_0}(x_{i_0}, y)$.

Assume for a contradiction that $\{[x_i, y.\theta_i]\}_{i\in I_0}$ is not a covering family. Then define $\eta(y)$ to be

$$\varphi(y) \wedge \neg \bigvee_{i\in I_0} \exists x_i\theta_i(x_i, y),$$

and note that $\eta(y)$ is consistent with T. Then by Lemma 5.2.6, $[y.\eta]$ is a subobject of $[y.\varphi]$. Let us denote the corresponding monomorphism by m. Since S is covering in $J_{\neg\neg}$, there must be an arrow f into $[y.\eta]$ such that $mf \in S$. That is, $mf = [x_i, y.\theta_i]$ for some $i \in I$. Considering the regular epi-mono factorization of $[x_i, y.\theta_i]$ we find the following commuting diagram:

$$\begin{tikzcd}
{[x_i.\chi]} \arrow[r] \arrow[d, "f"'] \arrow[dr, "{[x_i,y.\theta_i]}"] & {[y.\exists x_i\theta_i(x_i, y)]} \arrow[d] \\
{[y.\eta]} \arrow[r, "m"'] & {[y.\varphi]}
\end{tikzcd}$$

There must then be $i_0 \in I_0$ such that $\exists x_i \theta_i(x_i, y)$ is T-provably equivalent to $\exists x_{i_0} \theta_{i_0}(x_{i_0}, y)$, which means that $[y.\exists x_i \theta_i(x_i, y)]$ and $[y.\exists x_{i_0} \theta_{i_0}(x_{i_0}, y)]$ are isomorphic. We thus have an arrow $[x_i, y.\sigma] : [x_i.\chi] \to [y.\varphi]$ factoring through both $[y.\eta]$ and $[y.\exists x_{i_0} \theta_{i_0}(x_{i_0}, y)]$. Then from $\sigma(x_i, y)$ we would be able to derive both $\eta(y)$ and $\exists x_{i_0} \theta_{i_0}(x_{i_0}, y)$ relative T, which is a contradiction. So $\sigma(x_i, y)$ is inconsistent with T and can thus not exist. Thus we conclude that $\{[x_i, y.\theta_i]\}_{i \in I_0}$ is a covering family in S and hence that S is covering for J_κ. $\square$

Corollary 10.3.5. *If T is a classically locally small infinitary first-order theory, then for large enough κ, $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$ is Boolean.*

Once more, we have the following proposition, which we had already seen in the context of $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$ and $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$.

Proposition 10.3.6. *Let X be a type in $\mathscr{L}$, then every subobject of $X^{y(U_T)}$ in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$ is the interpretation of some $\mathscr{L}_\infty$-formula, this formula is always a disjunction of $\mathscr{L}_\kappa$-formulas.*

Proof. This is the same as Proposition 7.1.4, where we had this exact same statement for the case of $\mathbf{Syn}^{\mathrm{fo}}_\kappa(T)$. However, the proof was already given for the case of $\mathbf{Syn}^{\mathrm{g}}_\kappa(T)$ in Proposition 6.2.3. $\square$

10.4 Existence of Boolean classifying topoi

We now have enough tools to prove Theorem 10.1.8. Let us first recall its statement.

Theorem 10.1.8**, repeated.** *An infinitary first-order theory T has a Boolean classifying topos $\mathbf{Set}^{\mathrm{b}}[T]$ if and only if T is classically locally small.*

Proof. The proof from the left to the right uses that the $\mathscr{L}_\infty$-formulas valid in G_T are precisely those that are classically derivable from T (Corollary 10.1.9), and is essentially the same as the proof of (iii) $\implies$ (i) of Theorem 9.0.7. Fix some context x and its type X. Then every $\mathscr{L}_\infty$-formula in context x is represented by some subobject of X^{G_T}, and two such formulas are classically T-provably equivalent if and only if they are represented by the same subobject. Since $\mathrm{Sub}(X^{G_T})$ is a set, there is a set of different $\mathscr{L}_\infty$-formulas in context x, up to classical T-provable equivalence.

For the proof from the right to the left assume that T is classically locally small. By Corollary 10.3.5 we then find a large enough κ such that $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$ is Boolean. We claim that $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$ gives us the Boolean classifying topos for T. By Diaconescu's theorem (Corollary 2.3.9) and Proposition 6.1.5 we have that

$$\mathbf{Topos}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)) \simeq \mathbf{FlatCon}((\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa), \mathcal{E}) = \mathbf{Geom}_\kappa(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), \mathcal{E}).$$

So given any κ-geometric functor $F : \mathbf{Syn}^{\mathrm{c}}_\kappa(T) \to \mathcal{E}$, we find a corresponding geometric morphism $f : \mathcal{E} \to \mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$. Under the equivalence of Diaconescu's theorem, we have that F is isomorphic to $f^* y$. Here we use that J_κ is subcanonical for $\mathbf{Syn}^{\mathrm{c}}_\kappa(T)$ (Proposition 6.1.4), which is why we can leave out the sheafification functor. Because $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$ is Boolean, we have by Proposition 2.2.9 that f is open. So Proposition 4.0.9 tells us that f^* is a κ-Heyting functor, and by Proposition 7.1.1 we have that the Yoneda embedding into the category of sheaves is κ-Heyting. So F is isomorphic to a κ-Heyting functor, and we conclude that the inclusion

$$\mathbf{Heyt}_\kappa(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), \mathcal{E}) \hookrightarrow \mathbf{Geom}_\kappa(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), \mathcal{E})$$

is essentially surjective and thus gives us an equivalence of categories.

We can now apply Proposition 10.3.1 to conclude that

$$\mathbf{Topos}(\mathcal{E}, \mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)) \simeq \mathbf{Geom}_\kappa(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), \mathcal{E}) \simeq \mathbf{Heyt}_\kappa(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), \mathcal{E}) \simeq T\text{–}\mathbf{Mod}(\mathcal{E})_\kappa,$$

for Boolean $\mathcal{E}$. All that remains is to check that $T\text{–}\mathbf{Mod}(\mathcal{E})_\kappa = T\text{–}\mathbf{Mod}(\mathcal{E})_\infty$. This follows from Proposition 4.0.13, and the fact that $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$ is Boolean so every geometric morphism into it is open. Note that technically Proposition 4.0.13 is about $T\text{–}\mathbf{Mod}(\mathcal{E})$ and not $T\text{–}\mathbf{Mod}(\mathcal{E})_\kappa$, but its proof works equally well for $T\text{–}\mathbf{Mod}(\mathcal{E})_\kappa$. $\square$

10.5 For geometric theories

Suppose that T is a geometric theory, then by Theorem 6.2.1 its classifying topos $\mathbf{Set}[T]$ exists. In Proposition 10.1.6 we already established a connection between the first-order classifying topos of a theory, and its Boolean classifying topos. In this section we will look at the connection between the classifying topos, the first-order classifying topos and the Boolean classifying topos.

Proposition 10.5.1. *Let T be a geometric theory, then the following are equivalent:*

(i) $\mathbf{Set}[T]$ *is Boolean,*

(ii) $\mathbf{Set}[T] = \mathbf{Set}^{\mathrm{fo}}[T] = \mathbf{Set}^{\mathrm{b}}[T]$,

(iii) $\mathbf{Set}[T] = \mathbf{Set}^{\mathrm{b}}[T]$,

(iv) $\mathbf{Set}[T] = \mathbf{Set}^{\mathrm{fo}}[T]$.

If any of these equivalent conditions holds, the generic models are also the same.

Proof. The implications (ii) $\implies$ (iii), (ii) $\implies$ (iv) and (iii) $\implies$ (i) are trivial, so let us prove (i) $\implies$ (ii) and (iv) $\implies$ (i).

(i) $\implies$ (ii) If $\mathbf{Set}[T]$ is Boolean, then every geometric morphism into $\mathbf{Set}[T]$ is open. So for any topos $\mathcal{E}$ we have

$$\mathbf{Open}(\mathcal{E}, \mathbf{Set}[T]) = \mathbf{Topos}(\mathcal{E}, \mathbf{Set}[T]) \simeq T\text{–}\mathbf{Mod}(\mathcal{E}) = T\text{–}\mathbf{Mod}(\mathcal{E})_\infty.$$

The last equality follows because every homomorphism of models of T corresponds to a natural transformation of open geometric morphisms, which means that the homomorphism is actually ∞-elementary. So indeed $\mathbf{Set}[T] = \mathbf{Set}^{\mathrm{fo}}[T]$. Now it also follows directly that $\mathbf{Set}[T] = \mathbf{Set}^{\mathrm{b}}[T]$, either by restricting ourselves to Boolean $\mathcal{E}$ in the above equivalence, or by considering Proposition 10.1.6 and the fact that since $\mathbf{Set}^{\mathrm{fo}}[T]$ is Boolean we have $\mathcal{B}(\mathbf{Set}^{\mathrm{fo}}[T]) \simeq \mathbf{Set}^{\mathrm{fo}}[T]$.

(iv) $\implies$ (i) Let $f : \mathcal{E} \to \mathbf{Set}[T]$ be any geometric morphism. Then f classifies a model M of T in $\mathcal{E}$. Since $\mathbf{Set}[T] = \mathbf{Set}^{\mathrm{fo}}[T]$, this model is also classified by some open geometric morphism g into $\mathbf{Set}[T]$. Then f and g are two geometric morphisms (into $\mathbf{Set}[T]$) classifying the same model, so they have to be isomorphic. So f must be an open geometric morphism. Since f was arbitrary, we can conclude that every geometric morphism into $\mathbf{Set}[T]$ is open, and hence by Proposition 2.2.9 that $\mathbf{Set}[T]$ is Boolean.

The last claim follows because the generic model is the model corresponding to the identity on $\mathbf{Set}[T] = \mathbf{Set}^{\mathrm{fo}}[T] = \mathbf{Set}^{\mathrm{b}}[T]$. □

There is a more concrete characterization of when $\mathbf{Set}[T]$ is Boolean, for coherent T, by Andreas Blass and Andrej Ščedrov in [BŠ83]. This description will allow us to find some easy examples of theories for which their classifying topos is Boolean. Then by Proposition 10.5.1 such theories will also have a first-order classifying topos and a Boolean classifying topos.

Theorem 10.5.2. *Let T be a coherent theory in some language $\mathscr{L}$, then* $\mathbf{Set}[T]$ *is Boolean if and only if the following two conditions hold:*

(i) every $\mathscr{L}_\omega$-formula is classically T-provably equivalent to a coherent formula;

(ii) and for every context x, there are only finitely $\mathscr{L}_\omega$-formulas in context x, up to classical T-provable equivalence.

Proof. This is precisely [BŠ83, Theorem 1]. □

Example 10.5.3. In Example 9.0.5 we had considered the theory DLO of dense linear order without endpoints. This theory is actually an example of a theory where $\mathbf{Set}[\mathrm{DLO}] = \mathbf{Set}^{\mathrm{fo}}[\mathrm{DLO}] = \mathbf{Set}^{\mathrm{b}}[\mathrm{DLO}]$. To see this, we will verify the two properties listed in Theorem 10.5.2. For the rest of this example we will only use the classical deduction-system, so "equivalent" will mean "classically equivalent", etc.

It is well-known that every $\mathscr{L}_\omega$-formula is DLO-equivalent to a quantifier-free formula (DLO has quantifier elimination, see [Mar02, Theorem 3.1.3]). We can bring such a quantifier-free formula in disjunctive normal form. So every formula is DLO-equivalent to one of the form

$$\bigvee_i \bigwedge_{j_i} A_{j_i},$$

where the disjunction and conjunctions are finite, and each A_{j_i} is an atomic formula or a negation thereof. The only atomic formulas are $x = y$, $x < y$, $x > y$, $\top$ and $\bot$ (with the possibility that x and y are the same variable). Their negations are DLO-equivalent to $x < y \vee x > y$, $x > y \vee x = y$, $x < y \vee x = y$, $\bot$ and $\top$ respectively. So every $\mathscr{L}_\omega$-formula is DLO-equivalent to a coherent formula, which shows that DLO satisfies property (i) from Theorem 10.5.2.

For the second property we fix some context x, which is just a finite string of variables $x_1, \ldots, x_n$. There are only finitely many possible atomic formulas in at most these free variables. By the discussion above, every $\mathscr{L}_\omega$-formula in context x is equivalent to a disjunction of conjunctions of atomic formulas with free variables among x. There are, up to equivalence, only finitely many conjunctions of atomic formulas in context x, and hence there are only finitely many $\mathscr{L}_\omega$-formulas in context x. So DLO also satisfies property (ii) from Theorem 10.5.2, and so we have that $\mathbf{Set}[\mathrm{DLO}]$ is Boolean. Then by Proposition 10.5.1 we indeed have $\mathbf{Set}[\mathrm{DLO}] = \mathbf{Set}^{\mathrm{fo}}[\mathrm{DLO}] = \mathbf{Set}^{\mathrm{b}}[\mathrm{DLO}]$.

Using Proposition 10.5.1 we can also generalize Theorem 10.5.2 to infinitary logic.

Theorem 10.5.4. *Let T be a geometric theory. Then $\mathbf{Set}[T]$ is Boolean if and only if the following holds: there is a κ such that T is expressible in $\mathscr{L}_\kappa$ and both of the following conditions are satisfied:*

(i) every $\mathscr{L}_\kappa$-formula is classically T-provably equivalent to a κ-geometric formula;

(ii) and for every context x, there are $< \kappa$ different $\mathscr{L}_\kappa$-formulas, up to classical T-provable equivalence.

Note that one direction is slightly weaker than in the case of Theorem 10.5.2. That is, if we assume that $\mathbf{Set}[T]$ is Boolean (for coherent T), then we only get some κ for which conditions (i) and (ii) hold, while Theorem 10.5.2 also tells us that $\kappa = \omega$. The other direction is the same, because if (i) and (ii) hold for ω, then they certainly hold for some κ. The reason for the slight weakening of one direction is that the proof in [BŠ83] heavily relies on model-theoretic arguments, which do not directly generalize to infinitary logic.

Proof of Theorem 10.5.4**.** Suppose that $\mathbf{Set}[T]$ is Boolean. Then by Proposition 10.5.1 we have that $\mathbf{Set}^{\mathrm{b}}[T] = \mathbf{Set}[T]$. So since $\mathbf{Set}^{\mathrm{b}}[T]$ exists, we have that T is classically locally small by Theorem 10.1.8. Denote by S_x the set of different $\mathscr{L}_\infty$-formulas in context x, up to classical T-provable equivalence. Then we choose κ to be strictly larger than the cardinality S_x, for all contexts x. Then clearly (ii) is satisfied. For (i) we let $\varphi(x)$ be any $\mathscr{L}_\kappa$-formula, then by Proposition 6.2.3 we have that $\{x : \varphi(x)\}^{G_T}$ is also the interpretation of a disjunction of κ-geometric formulas. Let us call this disjunction $\psi(x)$. We had already seen that there are $< \kappa$ different κ-geometric formulas, up to classical T-provable equivalence. So $\psi(x)$ itself is κ-geometric. Since we have $\{x : \varphi(x)\}^{G_T} = \{x : \psi(x)\}^{G_T}$ in $\mathbf{Set}[T] = \mathbf{Set}^{\mathrm{b}}[T]$, we have by Corollary 10.1.9 that $\varphi(x)$ and $\psi(x)$ are classically T-provably equivalent, which establishes (i).

Now suppose that there is κ such that T is a κ-geometric theory satsifying conditions (i) and (ii). By (ii) we have that κ satisfies the necessary conditions that the κ described in section 10.3 does. In particular, we have that $\mathbf{Set}^{\mathrm{b}}[T] = \mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$. We claim that the inclusion $\mathbf{Syn}^{\mathrm{g}}_\kappa(T) \hookrightarrow \mathbf{Syn}^{\mathrm{c}}_\kappa(T)$ gives an equivalence of categories, because then we have that

$$\mathbf{Set}^{\mathrm{b}}[T] = \mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa) \simeq \mathbf{Sh}(\mathbf{Syn}^{\mathrm{g}}_\kappa(T), J_\kappa) = \mathbf{Set}[T].$$

To prove the claim we note that it follows directly from (i) that the inclusion is essentially surjective. This also shows that the inclusion is full: any arrow $[x, y.\theta]$ in $\mathbf{Syn}^{\mathrm{c}}_\kappa(T)$ must be represented by some κ-geometric formula. So we may assume $\theta(x, y)$ to be κ-geometric. There is a small detail here that needs some consideration: $\theta(x, y)$ is T-provably functional relative to the classical deduction-system. To see that $\theta(x, y)$ is also T-provably functional relative the deduction-system for geometric logic, we can apply Barr's theorem (or, technically, a consequence thereof: Corollary 8.2.2). Now for faithfulness, suppose that $\theta(x, y)$ and $\sigma(x, y)$ represent the same arrow in $\mathbf{Syn}^{\mathrm{c}}_\kappa(T)$. Then they must be classically T-provably equivalent. Using Barr's theorem again we then conclude that $\theta(x, y)$ and $\sigma(x, y)$ are already T-provably equivalent relative to the deduction-system for geometric logic. So they also represent the same arrow in $\mathbf{Syn}^{\mathrm{g}}_{\kappa}(T)$. This proves the claim, and thus concludes the proof of our theorem. $\square$

A final application of Proposition 10.5.1 is to obtain a result like Theorem 9.2.1 and Theorem 9.2.4, but now for Boolean classifying topoi. Of course, in this case we have to only look at Boolean topoi. That is, we have the following theorem.

Theorem 10.5.5. *Every Boolean topos $\mathcal{E}$ is the Boolean classifying topos of some geometric theory T.*

Proof. By Theorem 9.2.1 we have that $\mathcal{E} = \mathbf{Set}[T]$ for some geometric theory T. Since $\mathcal{E}$ is Boolean by assumption, we have by Proposition 10.5.1 that $\mathcal{E} = \mathbf{Set}[T] = \mathbf{Set}^{\mathrm{b}}[T]$. $\square$

Chapter 11

Connections with Model Theory

In this chapter we will provide an example and non-example of a connection between Model Theory and Boolean classifying topoi. We will be looking at Boolean classifying topoi, because the logic we study in Model Theory is classical. This does come with two issues. The first is that in Model Theory we study finitary logic (i.e. we only consider $\mathscr{L}_\omega$-formulas). The second issue is that not every theory may have a Boolean classifying topos. There are two solutions to these issues. We can either just consider suitable theories, or we can look at a suitable extension of a theory.

The first solution would be to consider theories in which there are only finitely many different $\mathscr{L}_\omega$-formulas, up to classical provable equivalence.

Definition 11.0.6. We call the formulas expressible in $\mathscr{L}_\omega$ *finitary*. A theory expressible in $\mathscr{L}_\omega$ will be called a *finitary theory* or *finitary first-order theory*.

Definition 11.0.7. A finitary theory T is called *classically locally finite* in a context x if there are only finitely many different $\mathscr{L}_\omega$-formulas in context x, up to classical T-provable equivalence. We say that T is *classically locally finite* if it is classically locally finite in every context over its language.

The following proposition explains why this solves both issues we mentioned before.

Proposition 11.0.8. *Let T be a classically locally finite theory. Then $\mathbf{Set}^{\mathrm{b}}[T]$ exists and is given by $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\omega(T), J_\omega)$. Furthermore, every $\mathscr{L}_\infty$-formula is classically T-provably equivalent to a disjunction of $\mathscr{L}_\omega$-formulas.*

Proof. In sections 10.3 and 10.4 we saw that for large enough κ we could construct the Boolean classifying topos of a theory as $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\kappa(T), J_\kappa)$. In section 10.3 we saw that such a κ must be such that up to classical T-provable equivalence there are $< \kappa$ different $\mathscr{L}_\kappa$-formulas in any context, and such that T is expressible in $\mathscr{L}_\kappa$. In this case we can thus take $\kappa = \omega$, so we do indeed have that $\mathbf{Set}^{\mathrm{b}}[T]$ exists and is given by $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{c}}_\omega(T), J_\omega)$.

We can now also apply Proposition 10.3.6 to see the interpretation of any $\mathscr{L}_\infty$-formula in G_T is the same as the interpretation of a disjunction of $\mathscr{L}_\omega$-formulas. So by Corollary 10.1.9 we conclude that every $\mathscr{L}_\infty$-formula is classically T-provably equivalent to a disjunction of $\mathscr{L}_\omega$-formulas. $\square$

The second solution was to consider a suitable extension. We had already seen such an extension in section 9.3. Let us give this extension a name.

Definition 11.0.9. The *classifying extension* $\overline{T}$ of a finitary first-order theory T is defined to be the full infinitary first-order theory of $y(U_T)$ in $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\omega(T), J_\omega)$.

Recall from the construction in section 9.3 that we can find λ big enough such that $\overline{T}$ is expressible in $\mathscr{L}_\lambda$ (Lemma 9.3.4). Furthermore, we recall the following three facts about $\overline{T}$ from Proposition 9.3.1.

1. The first-order classifying topos $\mathbf{Set}^{\mathrm{fo}}[\overline{T}]$ is given by $\mathbf{Sh}(\mathbf{Syn}^{\mathrm{fo}}_\omega(T), J_\omega)$.
2. The theory $\overline{T}$ is an $\mathscr{L}_\omega$-conservative extension of T.
3. Every $\mathscr{L}_\infty$-formula is $\overline{T}$-provably equivalent to a disjunction of $\mathscr{L}_\omega$-formulas.

In particular, from the first fact we find that $\mathbf{Set}^{\mathrm{b}}[\overline{T}]$ exists and is given by $\mathcal{B}(\mathbf{Set}^{\mathrm{fo}}[\overline{T}])$.

11.1 Complete theories

In Model Theory we are often interested in complete theories. There is a simple condition on the $\mathbf{Set}^{\mathrm{b}}[T]$ for a classically locally finite theory T that characterizes when such a theory is complete. Let us make precise what all of this means.

Definition 11.1.1. We call a finitary first-order theory *complete* if every finitary sentence is either true in all **Set**-models or false in all **Set**-models. Note that by soundness and Gödel's completeness theorem this is equivalent to saying that for every finitary sentence φ we either have that φ is T-provable or $\neg\varphi$ is T-provable in the deduction-system for classical logic.

We also recall the following definition from Lattice Theory.

Definition 11.1.2. We call a lattice *two-valued* if it is of the form $\{0, 1\}$ with $0 \leq 1$. In particular, such a lattice is a Boolean algebra. Accordingly, we call a topos $\mathcal{E}$ a *two-valued topos* if $\mathrm{Sub}(1)$ in $\mathcal{E}$ is two-valued.

We are now ready to state the first main result of this section. We note that we implicitly applied Proposition 11.0.8 already in this statement to assume that $\mathbf{Set}^{\mathrm{b}}[T]$ actually exists.

Theorem 11.1.3. *Let T be a classically locally finite theory. Then $\mathbf{Set}^{\mathrm{b}}[T]$ is two-valued if and only if T is complete.*

Part of the proof of this theorem may have value on its own, so we present it as a lemma first.

Lemma 11.1.4. *For any finitary first-order theory T we have that if $\mathbf{Set}^{\mathrm{b}}[T]$ exists and is two-valued, then T is complete.*

Proof. Let φ be any finitary sentence, then $\{\emptyset : \varphi\}^{G_T} = 0$ or $\{\emptyset : \varphi\}^{G_T} = 1$. So we must have that either φ or $\neg\varphi$ is valid in the generic model G_T. Then by Corollary 10.1.9 we conclude that either φ or $\neg\varphi$ can be derived classically from T. □

Proof of Theorem 11.1.3**.** The proof from the left to the right is just Lemma 11.1.4. For the converse we have by Proposition 10.3.6 that any subobject of 1 in $\mathbf{Set}^{\mathrm{b}}[T]$ is the interpretation of an $\mathscr{L}_\infty$-formula, which is classically T-provably equivalent to a disjunction of $\mathscr{L}_\omega$-formulas by Proposition 11.0.8. So let $\{\emptyset : \bigvee_i \varphi_i\}^{G_T}$ be any subobject of 1 in $\mathbf{Set}^{\mathrm{b}}[T]$. Then for every i we have that either φ_i or $\neg\varphi_i$ is classically derivable from T, to we must have that either $\bigvee_i \varphi_i$ or $\neg\bigvee_i \varphi_i$ is classically derivable from T. So $\{\emptyset : \bigvee_i \varphi_i\}^{G_T} = 0$ or $\{\emptyset : \bigvee_i \varphi_i\}^{G_T} = 1$, and we conclude that $\mathbf{Set}^{\mathrm{b}}[T]$ is two-valued. □

Example 11.1.5. In Example 9.0.5 and Example 10.5.3 we had already considered the theory DLO, and in the latter we saw that DLO is classically locally finite. It is well-known that DLO is complete (in fact, this follows directly from the fact that it has quantifier elimination and the only atomic sentences are $\top$ and $\bot$). So we can apply Theorem 11.1.3 directly to see that $\mathbf{Set}^{\mathrm{b}}[\mathrm{DLO}]$ is two-valued. Recall from Example 10.5.3 that $\mathbf{Set}^{\mathrm{b}}[\mathrm{DLO}] = \mathbf{Set}^{\mathrm{fo}}[\mathrm{DLO}] = \mathbf{Set}[\mathrm{DLO}]$, so we can actually conclude that $\mathbf{Set}^{\mathrm{fo}}[\mathrm{DLO}]$ and $\mathbf{Set}[\mathrm{DLO}]$ are two-valued as well.

Example 11.1.6. Let T be the theory of dense linear orders, possibly with endpoints. That is, T has the same axioms as DLO except for the two explicitly stated in Example 9.0.5, which were

$$\forall x \exists y (x < y),$$
$$\forall x \exists y (y < x).$$

Now that the endpoints are not specified by T, we have that T is not complete. There are five possible completions of T, by specifying which endpoints there are, and, if they both exist, whether or not they are equal. We will show that each of these completions is classically locally finite, because then T is classically locally finite. In the case that the endpoints are equal, the only model is the model with one element. So in any context there are precisely two equivalence classes of formulas. For the other four options, we note that each of the completions is ω-categorical (the only countable model of such a theory is $\mathbb{Q}$ with the necessary endpoints added). By Ryll-Nardzewski's theorem ([Mar02, Theorem 4.4.1]), this is equivalent to saying that each of these completions is classically locally finite.

Now that we have established that T is classically locally finite and not complete, we can apply Theorem 11.1.3 to see that $\mathbf{Set}^{\mathrm{b}}[T]$ is not two-valued.

Let us now turn towards the classifying extension of a theory, then we find the following result.

Theorem 11.1.7. *Let T be a complete finitary first-order theory, then $\mathbf{Set}^{\mathrm{b}}[\overline{T}]$ is two-valued.*

Proof. The proof is actually very similar to that of Theorem 11.1.3. Every subterminal object in $\mathbf{Set}^{\mathrm{b}}[\overline{T}]$ is the interpretation of some $\mathscr{L}_\infty$-sentence by Proposition 10.3.6. Then by the third fact about the classifying extension $\overline{T}$ in the introduction of this chapter, we have that every $\mathscr{L}_\infty$-sentence is $\overline{T}$-provably equivalent to a disjunction of $\mathscr{L}_\omega$-sentences. Now we can just repeat the argument of Theorem 11.1.3 again: let $\{\emptyset : \bigvee_i \varphi_i\}^{G_T}$ be any subterminal object in $\mathbf{Set}^{\mathrm{b}}[\overline{T}]$. Then $\bigvee_i \varphi_i$ or $\neg \bigvee_i \varphi_i$ must be classically T-provable (and hence classically $\overline{T}$-provable), so the subterminal object is either 0 or 1. $\square$

11.2 Quantifier elimination: a non-example

If a model-theoretic property can be characterized by the means of the Boolean classifying topos, it must be invariant under Morita-equivalence. Here by Morita-equivalence we mean restricted to Boolean topoi, that is

$$T\text{–}\mathbf{Mod}(\mathcal{E})_\infty \simeq T'\text{–}\mathbf{Mod}(\mathcal{E})_\infty,$$

natural in $\mathcal{E}$ where $\mathcal{E}$ is a Boolean topos. For the rest of this section, when we talk about Morita-equivalence, we will mean this version.

Unfortunately, not all properties that are considered in Model Theory are invariant under Morita-equivalence. An example of such a model-theoretic property is quantifier elimination.

Definition 11.2.1. A finitary first-order theory T has *quantifier eliminiation* if every $\mathscr{L}_\omega$-formula $\varphi(x)$ is T-provably equivalent to a quantifier-free formula $\psi(x)$, in the deduction-system for classical logic.

In this section we will give an example of a two theories that are Morita-equivalent, but where one has quantifier elimination and the other does not. For this we turn towards the Model Theory of real closed fields. Let us first define what a real closed field is (following [Mar02, Definition 3.3.4]).

Definition 11.2.2. A field is called *formally real* if -1 is not a sum of squares. A *real closed field* is a field F that is formally real with no proper formally real algebraic extensions.

Let us now consider the languages we will be working with in this section.

Definition 11.2.3. We define the *language of rings* $\mathscr{L}^r$ to be a one-sorted language with no relation symbols and the following five function symbols: binary function symbols $+$, $-$ and $\cdot$ and two 0-ary function symbols (i.e. constants) 0 and 1. We define the *language of ordered rings* $\mathscr{L}^{or}$ to be $\mathscr{L}^r$ with one binary relation symbol $<$ added.

By [Mar02, Corollary 3.3.6], we can define a theory in $\mathscr{L}^r$ such that its models are precisely the real closed fields. This theory is given in the following definition.

Definition 11.2.4. The *theory of real closed fields in the language of rings* RCF_r in $\mathscr{L}^r$ has the following axioms:

- the axioms for fields;
- for each $n \geq 1$, the axiom
$$\forall x_1 \ldots x_n (x_1^2 + \ldots + x_n^2 + 1 \neq 0);$$
- the axiom $\forall x \exists y (y^2 = x \vee y^2 + x = 0)$ and

- for each $n \geq 0$, the axiom

$$\forall x_0 \dots x_{2n} \exists y \left(y^{2n+1} + \sum_{i=0}^{2n} x_i y^i = 0 \right).$$

Next we consider orderings in a field. We want the ordering on a field to respect the structure the field already has. So we have the following definition (from [Mar02, Example 1.2.9]).

Definition 11.2.5. A field F is called an *ordered field* if it is equiped with a linear order $<$ such that for all $x, y, z \in F$:

1. if $x < y$, then $x + z < y + z$;
2. if $x < y$ and $0 < z$, then $x \cdot z < y \cdot z$.

If F is any real closed field, then given a nonzero $a \in F$, either a or $-a$ has to be a square (and not both). We can thus make F into an ordered field by defining $x < y$ if and only if $y - x$ is a nonzero square. In formulas:

$$x < y \leftrightarrow \exists z(z \neq 0 \wedge x + z^2 = y).$$

This ordering is the only possible ordering to make F into an ordered field. So we can see any real closed field as an ordered field in a unique way. This yields the following definition, which is [Mar02, Definition 3.3.7].

Definition 11.2.6. The *theory of real closed fields in the language of ordered rings* RCF_{or} in $\mathscr{L}^{or}$ extends RCF_r with the following axioms:

- the axioms for $<$ being a linear order;
- the axioms for ordered fields, that is:

$$\forall xyz(x < y \to x + z < y + z),$$
$$\forall xyz((x < y \wedge 0 < z) \to x \cdot z < y \cdot z).$$

So both RCF_r and RCF_{or} axiomatize the real closed fields, but in different languages. The only difference between the two is the order relation $<$, which is definable in RCF_r but for which we actually have a symbol in RCF_{or}.

Back to quantifier elimination. We have the following two results.

Proposition 11.2.7. *The theory* RCF_r *does not have quantifier elimination.*

Proof. By the argument provided in [Mar02, Section 3.3]. Suppose that RCF_r has quantifier elimination. Then consider $\mathbb{R}$ with its usual interpretation as real closed field. The set $\{x : \exists z(z \neq 0 \wedge x = z^2)\}$ must be definable by a quantifier-free formula $\varphi(x)$. Such a formula is (equivalent to) a Boolean combination of atomic formulas, and each atomic formula is of the form $P(x) = 0$ where $P(x)$ is a polynomial in x. Every such polynomial has finitely many roots or is trivial, so the set $\{x : \varphi(x)\}$ must be either finite or cofinite. However, $\{x : \exists z(z \neq 0 \wedge x = z^2)\}$ is both infinite and coinfinite, so we conclude that RCF_r cannot have quantifier elimination. □

We thus see that the order in a real closed field must be defined using a quantifier when we are in the language $\mathscr{L}^r$. This turns out to be the only obstacle for quantifier elimination, as can be seen in the following theorem.

Theorem 11.2.8. *The theory* RCF_{or} *has quantifier elimination.*

Proof. This can be found as [Mar02, Theorem 3.3.15]. □

Given a model M of RCF_r in some Boolean topos $\mathcal{E}$, we can make it into a model of RCF_{or} by giving an interpretation for the binary relation symbol $<$. There is only one possible choice, namely the subobject that is the interpretation of $\exists z(z \neq 0 \wedge x + z^2 = y)$. This operation is invertible by forgetting the extra structure. Also every ∞-elementary morphism of models of RCF_r is also an ∞-elementary morphism of models of RCF_{or}, and vice versa. So we have established a Morita-equivalence between RCF_r and RCF_{or}, but one has quantifier elimination and the other has not.

Even when considering the classifying extension of a theory, quantifier elimination of the original theory is not invariant under Morita-equivalence. We can extend our example of RCF_r and RCF_{or} as follows. We let T_r be the full $\mathscr{L}_\omega^r$-theory of $\mathbb{R}$. That is, T_r consists of all those $\mathscr{L}_\omega^r$-sentences that are true in $\mathbb{R}$. Likewise, we let T_{or} be the full $\mathscr{L}_\omega^{or}$-theory of $\mathbb{R}$. By the same argument as in Proposition 11.2.7 we have that T_r does not have quantifier elimination. Since $\mathrm{RCF}_{or} \subseteq T_{or}$ we do have that T_{or} has quantifier elimination by Theorem 11.2.8.

By replacing every instance of $x < y$ by $\exists z(z \neq 0 \wedge x + z^2 = y)$ in an $\mathscr{L}_\omega^{or}$-formula, we find an $\mathscr{L}_\omega^r$-formula that is T_{or}-provably equivalent to it. Here we can take intuitionistic provability, because the equivalence of the two formulas is expressed as a sentence in T_{or}. Furthermore, any two $\mathscr{L}_\omega^r$-formulas that are T_{or}-provably equivalent, must also be T_r-provably equivalent. Again, because the right equivalences will be in the respective theories already.

This means that the inclusion $\mathbf{Syn}_\omega^{\mathrm{fo}}(T_r) \hookrightarrow \mathbf{Syn}_\omega^{\mathrm{fo}}(T_{or})$ is an equivalence of categories. Thus $\mathbf{Sh}(\mathbf{Syn}_\omega^{\mathrm{fo}}(T_r), J_\omega)$ and $\mathbf{Sh}(\mathbf{Syn}_\omega^{\mathrm{fo}}(T_{or}), J_\omega)$ are equivalent and the theories $\overline{T_r}$ and $\overline{T_{or}}$ of the models $y(U_{T_r})$ and $y(U_{T_{or}})$ are Morita-equivalent. However, this means that we now have an example of two theories such that their classifying extensions are Morita-equivalent, but one of the theories has quantifier elimination while the other does not.

Chapter 12

Conclusion

In this thesis we have studied first-order classifying topoi, and which theories admit such a first-order classifying topos. After an extensive treatment of this characterization and the construction of such first-order classifying topoi, we developed the theory of Boolean classifying topoi. We then gave a simple application of Boolean classifying topoi to Model Theory. We shortly recall the results that were obtained in this thesis.

1. In section 8.1 we obtained a completeness result for infinitary classical logic, with respect to models in Boolean topoi.
2. We introduced the notion of a Boolean classifying topos in chapter 10, and characterized which infinitary first-order theories admit such a Boolean classifying topos (Theorem 10.1.8). We described a construction that can be performed on the first-order classifying topos of a theory, to obtain the Boolean classifying topos of that theory (Proposition 10.1.6).
3. We generalized a result from Blass and Ščedrov to characterize those (infinitary) geometric theories whose classifying topos is Boolean (Theorem 10.5.4).
4. We gave an example and a non-example of a connection between Boolean classifying topoi and Model Theory in chapter 11. In particular, we obtained the following results:
   (i) a characterization of classically locally finite complete theories (Theorem 11.1.3) and a necessary condition on $\mathbf{Set}^{\mathrm{b}}[\overline{T}]$ for when T is complete (Theorem 11.1.7);
   (ii) a counterexample to show that the property of quantifier elimination cannot be characterized in a similar way (section 11.2).

Naturally, one can ask a few more interesting questions regarding the topic of this thesis.

1. Can we obtain a notion of sub-first-order classifying topos? In the sense that, if we consider a theory T in infinitary sub-first-order logic, as described in Definition 8.1.1, can we find a topos $\mathcal{E}$ such that we have
$$\mathbf{SubOpen}(\mathcal{F}, \mathcal{E}) \simeq T\text{–}\mathbf{Mod}(\mathcal{F}),$$
natural in $\mathcal{F}$. Here $\mathbf{SubOpen}(\mathcal{F}, \mathcal{E})$ denotes the full subcategory of $\mathbf{Topos}(\mathcal{F}, \mathcal{E})$, that consists of sub-open geometric morphisms. One would hope that by using techniques similar to the construction of $\mathbf{Set}^{\mathrm{fo}}[T]$ such a topos $\mathcal{E}$ can be constructed, and that a characterization of theories which admit such a sub-first-order classifying topos can be obtained. In particular, one possible approach seems to be to find a theory of sub-open geometric morphisms, like we found a theory of open geometric morphisms in section 9.2.
2. Are there other model-theoretic properties that can be connected to topos-theoretic properties? It seems that looking into definable sets may very well yield some results, because the subobjects of X^{G_T} (where X is a type) in a Boolean classifying topos very much look like a formal form of definable sets. Perhaps some results can be obtained concerning elimination of imaginaries?
3. Stability is a very important topic in Model Theory, but it is also very abstract. Perhaps new insights can be obtained when looking at the property of being stable from a topos-theoretic side?

Index of symbols

Index of terms

Bibliography

[BJ98] Carsten Butz and Peter Johnstone. Classifying toposes for first-order theories. *Annals of Pure and Applied Logic*, 91:33–58, 1998.

[BŠ83] Andreas Blass and Andrej Ščedrov. Boolean classifying topoi. *Journal of Pure and Applied Algebra*, 28(1):15–30, 1983.

[But98] Carsten Butz. *Regular Categories and Regular Logic*. 1998.

[Joh77] Peter Johnstone. *Topos Theory*. Academic Press Inc. (London), 1977.

[Joh80] Peter Johnstone. Open maps of toposes. *Manuscripta Mathematica*, 31:217–247, 1980.

[Joh02a] Peter Johnstone. *Sketches of an Elephant: A Topos Theory Compendium*, volume 1. Oxford University Press, 2002.

[Joh02b] Peter Johnstone. *Sketches of an Elephant: A Topos Theory Compendium*, volume 2. Oxford University Press, 2002.

[Jon80] Dick de Jongh. A class of intuitionistic connectives. In Jon Barwise, H. Jerome Keisler, and Kenneth Kunen, editors, *The Kleene Symposium*, volume 101 of *Studies in Logic and the Foundations of Mathematics*, pages 103–111. Elsevier, 1980.

[Mar02] David Marker. *Model Theory: An Introduction*. Springer-Verlag New York Inc., 2002.

[ML71] Saunders Mac Lane. *Categories for the Working Mathematician*. Springer-Verlag New York Inc., second edition, 1971.

[MLM92] Saunders Mac Lane and Ieke Moerdijk. *Sheaves in Geometry and Logic*. Springer-Verlag New York Inc., 1992.

[Oos16] Jaap van Oosten. *Basic Category Theory and Topos Theory*. 2016.

[Sol66] Robert Solovay. New proof of a theorem of Gaifman and Hales. *Bulletin of the American Mathematical Society*, 72(2):282–284, 03 1966.