First Order Categorical Logic

Model-Theoretical Methods in the Theory of Topoi and Related Categories

Michael Makkai and Gonzalo E. Reyes

TO ZSUZSI AND MARIE

Preface

We arrived at the idea of writing a book on the connections of first order model theory and categories in early 1974 when we realized that it was impossible to communicate some of our fresh results and proofs because of the lack of a basic theory and a coherent terminology connecting logic and categories. As expected in such cases, it has taken us longer than promised to many people to complete the arduous task of writing this book.

The basic features of the work, some of them unusual, will be described in the Introduction below; here we make a few remarks only.

Primarily, the book is intended as a research monograph containing the exposition of the authors' results. On the other hand, it resembles a textbook because of the large amount of basic, sometimes even well-known, material that we have included. As a result, the book is essentially selfcontained reading. However, the reader should be warned that we have made no attempt to give a complete or even balanced account of the subject matter on the whole and it would be misleading to take the book as a faithful representation of the whole of categorical logic.

Our ideal goal in offering this work to the mathematical community is to help bring together two schools of thought in a fruitful collaboration. Logicians and category theorists seem to have resisted each others' ideas to a large extent. By building on very little in the way of prerequisites, and arriving at results which, besides being technically involved, have, we hope, some importance, we have attempted to show both logicians and category theorists some of the potentialities of a collaboration.

We would like to thank André Joyal and William Lawvere for many inspiring conversations. The stimulus they have given to our work goes much beyond the specific references we make to their papers.

The subject matter of this book was the topic of many sessions of the Séminaire de Logique, Université de Montréal, in the years 1973 to 1975. In the winter quarter of 1976, the first author gave a course on categorical logic at the University of California, Los Angeles. Both authors had numerous occasions to talk about the subject at meetings and seminars.

In 1975-76, the first author was visiting U.C.L.A., while the second author held a leave fellowship of the Canadian Council; much of the work on the book was done during this time.

Both authors hold operating grants of the National Research Council of Canada. The preparation of the type-script was supported by the authors' NRC grants.

We would like to express our heartfelt thanks to Mrs. Esther Massa who performed the difficult task of producing the type-script with great care and skill.

Michael Makkai
Gonzalo E. Reyes

Montreal, April 1977

Contents

Introduction

This work has grown out of efforts to write up some results the authors obtained a couple of years ago, most of which were announced in Makkai and Reyes [1976]. We soon realized that the background or folklore material we needed had never been written down, nor actually had been worked out in detail. It became clear that, even for the purposes of organizing the background material, new notions and terminology had to be introduced. We decided to include a complete treatment of all necessary preliminaries. The result is this book which, of course, was not originally intended to be this long.

In its present form, the book is intended to be a selfcontained introduction into "first order categorical logic". Several points have to be made to clarify this term, or more precisely, our interpretation of it.

1. First of all, we deal with first order logic only. The reader will not find any reference to higher order logic, which is, however, intimately related to elementary topoi (cf. e.g. Kock and Reyes [1977], Fourman [1977]) and thus it is very important for logically oriented category theory. On the other hand, we include a kind of infinitary logic under 'first order logic', thus departing somewhat from the traditional terminology. The infinitary first order logic we deal with is what has been called finite-quantifier infinitary logic, or $L_{\infty\omega}$, in the literature (cf. e.g. Keisler [1971], Barwise [1975]). It goes beyond traditional (or, as we will say: finitary) first order logic by allowing the formation of infinite conjunctions ("and") and disjunctions ("or"). From now on, 'logic' means 'possibly infinitary first order logic' for us.

2. Secondly, we obtain an algebraic, in fact categorical, formulation of first order logic, in the following sense. In this formulation *theories* (sets of axioms) are replaced by (certain) *categories* (our use of the adjective 'logical' in connection with categories will indicate that indeed, those categories can stand for theories in a sensible way), and *models* of theories will be (certain) *functors* (again, we will talk about 'logical' functors). E.g., an ordinary model (in the sense of Model Theory, cf. e.g. [CK]), will correspond to a logical functor into **Set**, the category of sets. At the same time, logical functors will replace and generalize the various, usually awkward notions of interpretations of theories in each other.

3. The replacement of the basic notions of logic by categorical notions as described above is, however, not a primary aim for us here, and indeed, it is not carried out systematically. On the contrary, we systematize our treatment on a conceptual basis which is both logical and categorical in character. This conceptual basis is a direct generalization of Tarski's foundation for Model Theory (cf. e.g. Tarski [1952] and [CK]). We have in mind Tarski's notion of a *structure* of a *similarity type* and his notion of *truth* in structures of *formulas* of the language associated with the similarity type; notions that are fundamental to Model Theory. First of all, the syntactical notions (formulas, etc.) of logic in our work are identical to the usual ones, except for the generalization to many-sorted logic which, however, is already familiar to logicians (cf. e.g. Feferman [1968]). In particular, a *language* (which is our equivalent to Tarski's 'similarity type')

is a collection of symbols called 'sorts' as well as finitary sorted relation and operation symbols. Tarski's notion of a structure is replaced by a category-valued structure, or an *interpretation of the language in a category.* For an $\mathcal{R}$-*structure of type* L where L is a language, $\mathcal{R}$ is a category, we also write (suggestively, as it turns out) "$M : L \to \mathcal{R}$". This notion is a simple generalization of the ordinary notion of structure that will be obtained as a special case as an interpretation $M : L \to \mathbf{Set}$, with **Set** the category of sets. To give an idea of the notion, we mention a few points. Given $M : L \to \mathcal{R}$, M interprets each sort s of L as an object $M(s)$ of $\mathcal{R}$; in the case of an ordinary structure $M : L \to \mathbf{Set}$, $M(s)$ is a set, one of the partial *domains* of M. If f is a unary operation symbol in L intended 'to map elements of sort s into elements of sort t' (which is specified by the *sorting* of f: the single argument of f is specified as of sorts s, the value of f is specified as of sort t, all of which is denoted by writing $f : s \to t$), then M will interpret f as a morphism $M(f) : M(s) \to M(t)$ in $\mathcal{R}$; this becomes an *operation*, indeed in accordance with the intention of the sorting, mapping elements of $M(s)$ into elements of $M(t)$, in case of a structure $M : L \to \mathbf{Set}$.

The next (and main) task of generalizing Tarski's setup consists in defining the interpretation of formulas in a category-valued structure $M : L \to \mathcal{R}$. This will be described in Chapter 2 below; here we only mention a few salient points. If the free variables of the formula ϕ are among $\vec{x} = \langle x_1, \ldots, x_n \rangle$, we will define $M_{\vec{x}}(\phi)$ and it will be a *subobject* of X, X being the product of $M(s_1) \times \cdots \times M(s_n)$, with x_i a variable of sort s_i, $i = 1, \ldots, n$. This will accomplish the generalization of Tarski's notion of truth because in case $\mathcal{R} = \mathbf{Set}$, $M_{\vec{x}}(\phi)$ turns out to be the *extension* of ϕ in the structure M, i.e., $M_{\vec{x}}(\phi) = \{\langle a_1, \ldots, a_n \rangle \in X : M \models \phi[a_1, \ldots, a_n]\}$; here '$M \models \phi[\vec{a}]$' stands for '$\vec{a}$ satisfies ϕ in M', as usual. Note that, of course, in general we do not have 'elements' of objects in our category, so the notion '$M \models \phi[\vec{a}]$' will not be available. We also note that in order for $M_{\vec{x}}(\phi)$ to be defined, the category $\mathcal{R}$ will have to satisfy certain conditions.

4. Finally, our treatment of categorical logic is geared towards establishing a link with Grothendieck's theory of (Grothendieck) topoi as it is exposed in SGA4. One of our main points is that some of the fundamental properties of some notions in this theory (notably the notions of *topos, coherence* of, and in, topoi and *pretopos*) are purely logical. Even more specifically, e.g. the notion of pretopos can be given a purely modeltheoretical characterization (among all theories or logical categories), cf. Theorem 7.1.8. In the description of the contents below, the reader will recognize our basic orientation towards Grothendieck's theory. It is a very interesting fact that notions originally developed for the purposes of (abstract) algebraic geometry turn out to be intimately related to logic and model theory. Compared to other existing versions of algebraic logic, categorical logic has the distinction of being concerned with objects that appear in mathematical practice.

If asked what is the most immediate point of contact between the Grothendieck theory and logic, we would point to the notion of a *site*, a category with a Grothendieck topology. It seems to us that it is most natural to identify a site with a theory, in the context of continuous functors from the site. The point is that the notion of a covering has the same *arbitrary nature* as an *axiom*; and in fact, each covering is considered an axiom in the precise identification we will consider below, explicitly first in Chapter 6, Section 1.

We note that Giraud's theorem (cf. Chapter 1, Section 4) can obviously be regarded as a logical characterization of Grothendieck topoi.

After the above characterization of the basic features of our approach, we should add that some equally natural approaches might offer themselves, even in the context of first order logic alone; we are planning to study some of these directions in the future. Also, some topics that would naturally fit into our context are omitted; perhaps the main one is A. Joyal generalization of forcing to sites and topoi, the Kripke-Joyal semantics, cf. e.g. Kock and Reyes [1977]. Thus, the book is far from being a complete treatment of the subject.

Let us briefly recall some of the historical background of this work. For more history and a more general context, we refer to Kock and Reyes [1977].

The program of doing algebraic logic via categories, i.e., with categories representing theories, is due to F. W. Lawvere. In Lawvere [1963], he introduced a categorical formulation of algebraic theories in which the basic idea is that substitution should be represented by composition of arrows. The second step was taken in Lawvere [1965] where he introduced the idea that quantifiers were 'adjoints to substitution'. More precisely, given a morphism $A \xrightarrow{f} B$ in a category, by pullback we obtain the functor $S(B) \xrightarrow{f^*} S(A)$, with $S(A)$ and $S(B)$ the subobjects categories of A and B, respectively; the quantifiers $\exists_f$, $\forall_f$ operating on elements of $S(A)$ are then defined as the left, respectively the right adjoint of f^*. The resulting 'elementary doctrines' are structures which are categories together with certain equivalents of the $S(A)$. The point of view taken in this book that theories correspond to categories without additional structure (but with additional properties) is due to Joyal and Reyes, and it appears in Reyes [1974].

In particular, the notion of a logical category, regarded as fundamental for categorical logic in this book, is due to Joyal and Reyes and it is the end product of several successive attempts at defining 'the right notion'. In the same paper, the existence of the classifying topos of coherent theories appears and together with this the realization that in connection with Grothendieck topoi, coherent first order logic has a distinctly important role.

Another important element in this work, the categorical interpretation of formulas first appears in Mitchell [1972] in a special context (for a similar work, cf. Osius [1973]). In Bénabou [1973], the substitution lemma and interpretation of formulas appear although in a context somewhat different from ours. Coste [1973] contains a categorical soundness theorem as well as a completeness theorem, in the context of intuitionistic logic and topos-valued models. Coste's soundness theorem is closely related to ours. Our work was independent of Coste's.

Next we give a description of the contents of the book.

Included solely for the convenience of the reader, Chapter 1 presents the basic theory of (Grothendieck) topoi. It follows SGA4 quite closely. It ends with an Appendix discussing some examples.

In Chapter 2, we define the basic notion, the interpretation of formulas in categories. Section 1 recalls the elementary concepts related to infinitary first order logic, $L_{\infty\omega}$. Section 2 introduces the (very elementary) categorical notions on which the interpretation is based, and Section 3 describes the interpretation itself. In Section 4 it is shown that certain properties of diagrams in categories can be expressed by formulas; this fundamental fact will be amplified and called the 'first main fact' in Chapter 3, Section 5.

Chapter 3 continues the study of the elementary properties of the categorical interpretation of formulas. The topic of Sections 1 and 2 is the *soundness* of certain rules of inference in the categorical interpretation: if a statement follows according to a specific formal rule from other statements that are all true in the category, then the original statement is true too. Here the *stability under pullback* of various notions play an important role just as it does in SGA4 where it is called *universality*. Section 3 is a detailed study of connections of some notions in SGA4 (such as effectiveness of equivalence relations, etc.) with logical formulas. Section 4 introduces the various kinds of 'logical' categories that can stand for theories of finitary and infinitary logic. The (simply) *logical* categories are related to finitary logic. More precisely, they turn out to be equivalent to so-called finitary *coherent* theories, axiomatized by Gentzen sequents with formulas built up using only $\wedge$, $\vee$ and $\exists$. A *pretopos* is a logical category with some additional conditions. We give a definition of pretopos in a logical spirit and show that this definition is equivalent to the definition given in SGA4, Exposé VI. We give infinitary generalizations of the above notions, arriving at κ-logical categories and κ-pretopoi, where κ is an infinite regular cardinal, or ∞. We show that ∞-pretopoi are almost the same as Grothendieck topoi. Finally, in Section 5 we outline the basic machinery of 'reducing' a categorical situation to a logical one and we give an example, Joyal's completeness theorem on logical categories and its proof via Gödel's completeness theorem. The technique of reduction is based on what we call the two main facts in connection with the relation of categories and first order logic. The first one says that a functor F from $\mathcal{R}$ preserves certain things (e.g., finite projective limits) in $\mathcal{R}$ if and only if F satisfies certain axioms. E.g., a logical functor from $\mathcal{R}$ will be one that (as an interpretation of a certain language) satisfies the axioms of a certain finitary coherent theory $T_{\mathcal{R}}$. The first main fact is based mainly on work done in Chapter 2. The second main fact is the soundness mentioned in connection with Sections 1 and 2; here it is formulated in a more general form. The two main facts are fundamental to our work later, especially in Chapters 6 and 7.

From the two main facts, the first one ('internal theories') is the one that seems to be a new contribution in this work; the second one was anticipated, although not quite in the form we need it here, by others, see especially Coste [1973].

Chapter 4 deals with elementary properties of Boolean- and Heyting-algebra-valued models. In Section 1, a notion is described which is the familiar one used by logicians except for small differences due to our use of many-sorted logic and possibly empty domains. In Section 2, D. Higgs' identification of the category $\mathrm{Sh}_{\mathcal{H}}$ of sheaves over a complete Heyting algebra $\mathcal{H}$ with the canonical topology on the one hand and the *category of $\mathcal{H}$-valued sets* on the other hand is stated without proof. (Unfortunately, D. Higgs' paper on the subject still exists only in preprint form.) Based on this identification, we describe how $\mathcal{H}$-valued models in the sense of Section 1 can be understood as $\mathrm{Sh}_{\mathcal{H}}$-structures, i.e., interpretations in the category $\mathrm{Sh}_{\mathcal{H}}$, in the sense of Chapter 2. Finally, we describe the well-known way of constructing 2-valued (**Set**-valued) models out of Boolean valued ones.

Chapter 5 is of a purely logical character without references to categories. In Section 1, we present the Boolean completeness theorem for $L_{\infty\omega}$. Not only do we give complete details but we also explain the (semantical) motivation for the (cut-free) Gentzen-type formal systems that we use. There are two versions of the Boolean completeness theorem in the literature. The first can be found in Karp [1964], the other one is the proof given in Mansfield [1972]. Karp's proof relies on a Lindenbaum-Tarski type construction of the Boolean value-algebra and is (therefore) related to what are called Hilbert-type formal systems. For our purposes Mansfield's approach is the natural one; indeed, this approach can be considered as a direct generalization of the two-valued completeness proof (for $L_{\omega\omega}$ and $L_{\omega_1\omega}$) for a Gentzen-type system (cf. e.g. Kleene [1967]) and also of the method of consistency properties (cf. Makkai [1969] and Keisler [1971]). Our detailed exposition is necessitated by the fact that the exact version we need cannot be found in the literature. This version has the features of applying to many-sorted logic, to possibly empty domains and of having a restricted cut-rule. In Section 2 we present an apparently new formal system that applies only to coherent logic. Here we also give a version of the method of consistency properties that will be used in Chapter 7.

The main part of the book consists of Chapters 6 and 7 where we describe our new logical proofs of some known theorems for categories as well as give several new results.

The topic of Chapter 6 is various embedding theorems for (Grothendieck) topoi. It turns out that they can be considered as more or less direct consequences of completeness theorems in logic. After some preliminaries in Section 1, Section 2 deals with embedding theorems with special Boolean 'target' topoi. The most general result in this area is Barr's theorem (Barr [1974]) that says that for any topos $\mathcal{E}$, there is a complete Boolean algebra $\mathcal{B}$ and a conservative geometric morphism $\mathrm{Sh}_{\mathcal{B}} \xrightarrow{u} \mathcal{E}$, with $\mathrm{Sh}_{\mathcal{B}}$ the category of sheaves over $\mathcal{B}$ with the canonical topology. Our proof of Barr's theorem, which is based on the Boolean completeness for $L_{\infty\omega}$, gives additional information such as a characterization of those infimums of subobjects in $\mathcal{E}$ that can be preserved by the inverse image functor $u^*:\mathcal{E} \to \mathrm{Sh}_{\mathcal{B}}$, etc. Another well known embedding theorem is Deligne's theorem (SGA4, Vol. 2, p. 173) that replaces $\mathrm{Sh}_{\mathcal{B}}$ by a Cartesian power $\mathbf{Set}^I$ of the category of sets in Barr's theorem for a coherent topos $\mathcal{E}$. This theorem turns out to be a 'consequence' of the original Gödel-Malcev completeness theorem for finitary logic. To the above, we add a new embedding theorem for what we call *separable* topoi. A topos is *separable* if it is equivalent to the category $\widetilde{\mathcal{C}}$ of sheaves over a site $\mathcal{C}$ having altogether countably many objects and morphisms and whose Grothendieck topology is generated by countably many covering families. Our result is that the conclusion of Deligne's theorem holds for separable topoi. It is surprising that this result was not noticed before; it certainly dispels a feeling one might have reading SGA4, namely, that the phenomenon of 'having enough points' is essentially related to the 'finitary' (quasi compact) character found in coherent topoi. Naturally, this result turns out to be related to the two-valued completeness theorem for countable fragments of $L_{\omega_1\omega}$, cf. e.g. Keisler [1971].

In Section 3, we consider embedding theorems with more general 'target' topoi; on the other hand, the inverse image functors of the geometric morphisms obtained here preserve more that before, namely the full power of the logic $L_{\infty\omega}$, including universal quantifiers and infinitary conjunctions. It would have been possible to do the work here on the basis of an *intuitionistically valid* formal system for $L_{\infty\omega}$, just like our extended version of Barr's theorem is based on a *classically valid* formal system; then the connections to (possibly infinitary) intuitionistic logic would have become clear. (At this point, the reader might profitably consult Rasiowa and Sikorski [1963] and Fitting [1969], although our exposition is selfcontained.) However, we have chosen a method of directly applying completeness for *coherent* logic. Our first two results seem to be simple-minded enough but we are unaware of their being stated in the literature. The first one (Theorem 6.3.1) says that every topos $\mathcal{E}$ has a complete Heyting algebra $\mathcal{H}$ and a conservative geometric morphism $\mathrm{Sh}_{\mathcal{H}} \xrightarrow{u} \mathcal{E}$ such that u^*, in addition, preserves intuitionistic implications, universal quantifiers and infinite conjunctions. We derive this as a corollary of our proof of Barr's theorem. The second result (Theorem 6.3.3) says that in the above theorem, $\mathrm{Sh}_{\mathcal{H}}$ can be replaced by the category of sheaves over a topological space whenever $\mathcal{E}$ has enough points, in particular, for coherent and separable topoi. The last result of this section, Theorem 6.3.5, is an elegant theorem of Joyal, which is a version of Kripke's completeness theorem for intuitionistic logic (cf. Kripke [1963] or Fitting [1969]). Joyal's theorem refers to a coherent topos $\mathcal{E}$ and starts with the category $\mathrm{Mod}(\mathcal{E})$ of all points (geometric morphisms) $\mathbf{Set} \to \mathcal{E}$. The theorem talks about the *evaluation functor* $\mathrm{ev}:\mathcal{E} \to \mathbf{Set}^{\mathrm{Mod}(\mathcal{E})}$ and it shows that ev is the inverse image functor of a geometric morphism, that ev is conservative and finally, that it preserves intuitionistic implication and $\forall$ on the level of subobjects of coherent objects. Joyal's theorem shows most clearly in what sense intuitionistic logic is *fully explained* in terms of coherent logic. We note that related work was done by Robitaille-Giguère [1975].

Chapter 7 contains our main new results. The basic situation we consider here is the following. Given a 'logical' (e.g., logical in the simple sense, or some infinitary sense) functor between 'logical' categories $I:\mathcal{R} \to \mathcal{S}$, we consider the categories $\mathrm{Mod}(\mathcal{R})$, $\mathrm{Mod}(\mathcal{S})$ of 'models' of $\mathcal{R}$ and $\mathcal{S}$, respectively, (e.g., in the simple logical case, the category of logical functors into **Set**) and the functor $I^*: \mathrm{Mod}(\mathcal{S}) \to \mathrm{Mod}(\mathcal{R})$ induced by I by composition. We prove results each saying that some property of I^* implies some other property of I. After proving some 'known' results and two rather technical theorems (which, however, have independent interest), in Section 1 we state and prove our characterization of pretopoi, Theorem 7.1.8. This refers to the above situation in the simply logical case and it says that if, in addition, $\mathcal{R}$ is a pretopos, I^* is an equivalence of categories, then so is I. This is indeed a characterization of pretopoi among logical categories because it is quite easy to see that it is false whenever $\mathcal{R}$ is not a pretopos and $\mathcal{S}$ is suitably chosen (take $\mathcal{S}$ to be the pretopos completion of $\mathcal{R}$, cf. Chapter 8). Paraphrasing the result, we might say this. Let us call $I:\mathcal{R} \to \mathcal{S}$ *strongly conservative* if I^* is an equivalence. This means that $\mathcal{S}$ is an 'extension' of $\mathcal{R}$ which, however, does not change the category of models. Now, the theorem says that pretopoi (and only pretopoi) are complete in the sense that they do not have *proper* strongly conservative extensions. The proofs in Section 1 are all applications of classical methods of finitary Model Theory, of course together with the basic reduction machinery.

The rest of Chapter 7 is devoted to infinitary generalizations of the results of Section 1. We succeeded in generalizing all results of Section 1 at the expense of introducing additional conditions on $\mathcal{R}$, $\mathcal{S}$ and I. The additional conditions turn out to be obviously true for the finitary case, so we have direct generalizations. The proofs of the infinitary versions are essentially different and more complicated than the finitary ones and they are patterned after Makkai [1969]. We actually formulate matters for *admissible fragments* of $L_{\infty\omega}$, but things are so arranged that the reader can ignore this level of generality and understand our proofs as referring to the simpler fragments $L_{\kappa\omega}$ of $L_{\infty\omega}$.

Up to and including Chapter 7, the basic point of view was that of replacing categories by theories. In Chapter 8, we perform the opposite step by constructing categories that can replace theories for all practical purposes, and we give some applications of this point of view. In particular, we describe the construction of the logical category associated to a finitary coherent theory, as well as the infinitary generalization of this construction. We explain that this construction (together with the 'first main fact') provides a basis for identifying finitary coherent theories, and actually, all theories in classical finitary first order logic, with logical categories. The application consist of descriptions of various kinds of 'completions' of categories. For instance, given a site $\mathcal{C}$, we give syntactical descriptions, or presentations in terms of logical operations, of the category $\widetilde{\mathcal{C}}$ of sheaves, and of the category (a pretopos) of coherent objects and morphisms in $\widetilde{\mathcal{C}}$ in case $\mathcal{C}$ is generated by finite coverings. Although these descriptions are quite elementary in nature, they contain information not immediately following from general arguments. They do not seem to appear in the literature except in the thesis Antonius [1975] in a somewhat different form. We feel that their knowledge should be an integral part of one's picture of topoi.

Chapter 9 touches on various topics. It discusses the notion of *classifying* topos of a theory. It gives a new proof of Grothendieck's theorem on the coherent objects of a coherent topos. It reformulates our characterization of pretopoi into a theorem on coherent topoi and coherent geometric morphisms, etc.

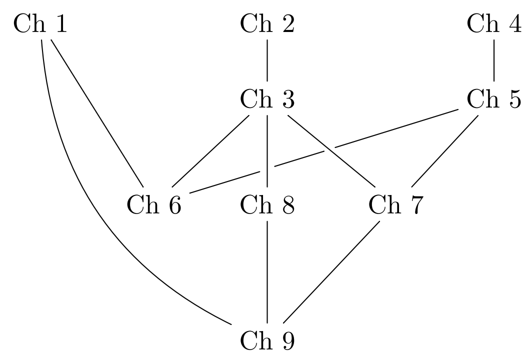
After the description of the contents, a few concluding remarks.

The book is essentially self-contained; it should be readable with a rudimentary knowledge of categories and with almost nothing in the way of a background in logic. Chapter 1 is a selfcontained exposition of topoi, using only material of e.g. Mac Lane [1971], except for a few minor places where we defer the work to Chapter 3. Chapter

2 and 3 do not rely on Chapter 1 and use only the most elementary category theory. All logical notions are explained (except things like free variables). Chapter 4 and 5 are likewise selfcontained expositions, entirely of a logical character, except that we state Higgs' theorem without proof and also another, actually related, result on Boolean valued models (Proposition 4.3.4). The first five chapters provide the ground-work for the rest of the book. Also, consult the chart of dependencies below.

The reader will find quite a few repetitions in the book; in general, we have tried to make the individual chapters readable by themselves as much as possible. E.g., although Chapter 6 relies on some things proved in Chapter 1, all what is needed from Chapter 1 is summarized in Section 1 of Chapter 6.

Leitfaden



Chapter 1

Grothendieck topoi

Introduction

In this chapter we will give the basic theory of Grothendieck topoi as it is exposed in SGA4, Volume 1. Our exposition follows SGA4 closely and it is selfcontained; the only prerequisite is some material in CWM. We will go only as far in the theory as our needs, especially in Chapter 6, dictate. In particular, no attempt is made to show what the algebraic geometrical motivation for topoi is. On the other hand, their logical context should gradually emerge in the later chapters.

A category in this book has either a set or a proper class of objects; with some exceptions noted below, the hom-sets $\mathrm{Hom}(A, B)$ are always sets. (In SGA4 terminology, every category is a $\mathcal{U}$-category, for a fixed universe $\mathcal{U}$.)

For emphasis, we may call a category *locally small* if each hom-set $\mathrm{Hom}(A, B)$ is a *set*. A locally small category is *small* if it has a *set* of objects.

§1 Sites and sheaves

In this book, every category will be assumed to have finite left limits.

A *site* is given by a category $\mathcal{C}$ (the *underlying category* of the site), together with a *Grothendieck topology on* $\mathcal{C}$ given by a class $\mathrm{Cov}(A)$ for each object A of $\mathcal{C}$. The elements of $\mathrm{Cov}(A)$ are families (*sets*) $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ of morphisms with codomain A; an element of $\mathrm{Cov}(A)$ is called a *covering family* of A. The Grothendieck topology has to satisfy the following four conditions:

1.1.1 (i) Every isomorphism $A' \xrightarrow{f} A$ gives a one-element covering family, $\{A' \xrightarrow{f} A\} \in \mathrm{Cov}(A)$.

1.1.1 (ii) ("Stability under pullback"). Whenever $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I} \in \mathrm{Cov}(A)$ and $B \xrightarrow{g} A$ is a morphism in $\mathcal{C}$, then $\langle A_i \times_A B \xrightarrow{f_i'} B\rangle_{i\in I}$ belongs to $\mathrm{Cov}(B)$; here

$$\begin{tikzcd} A_i \arrow[r, "f_i"] & A \\ A_i \times_A B \arrow[u] \arrow[r, "f_i'"'] & B \arrow[u, "g"'] \end{tikzcd}$$

is *any* pullback diagram, for each $i \in I$.
(Remark: Applying (ii) to the identity map $g = \mathrm{id}_A$, we obtain that in a covering family $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ any morphism $A_i \xrightarrow{f_i} A$ can be replaced by an 'isomorphic copy'

$A'_i \xrightarrow{f'_i} A$, (meaning that there is an isomorphism $A_i \xrightarrow{\alpha} A'_i$ with $f'_i\alpha = f_i$) such that the resulting family is still covering.)

1.1.1 (iii) ("Closure under composition"). Whenever $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I} \in \mathrm{Cov}(A)$ and $\langle A_{ij} \xrightarrow{g_{ij}} A_j\rangle_{j\in J_i} \in \mathrm{Cov}(A_i)$ for every $i \in I$, we have that $\langle A_{ij} \xrightarrow{f_i \circ g_{ij}} A\rangle_{j\in J_i, i\in I}$ belongs to $\mathrm{Cov}(A)$.

1.1.1 (iv) ("Monotonicity") If $\langle B_j \xrightarrow{g_j} A\rangle_{j\in J} \in \mathrm{Cov}(A)$ and $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ is such that for any $j \in J$ there is a $i \in I$ and a morphism $B_j \longrightarrow A_i$ with

$$\begin{tikzcd} B_j \arrow[r, "g_j"] \arrow[d] & A \\ A_i \arrow[ur, "f_i"'] & \end{tikzcd}$$

commutative, then $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I} \in \mathrm{Cov}(A)$.

In case the underlying category $\mathcal{C}$ is not small, only locally small, we impose another condition on the site $\mathcal{C}$: there should exist a *set* G of objects such that for each $A \in \mathrm{Ob}(\mathcal{C})$ there is a covering family $\langle A_i \xrightarrow{f_i} A\rangle$ with $A_i \in G$. G is called a *topologically generating set* for the site $\mathcal{C}$. For emphasis, the site $\mathcal{C}$ is said to be *locally small*.

The simplest example of a site is that derived from a topological space X. The open sets of X are made into a category $\mathcal{C}$; the objects of $\mathcal{C}$ are the open sets of X; with any domain U and codomain V there is at most one morphism $U \xrightarrow{f} V$ and there is one precisely when $U \subseteq V$. The Grothendieck topology is given by: $\langle U_i \to U\rangle_{i\in I} \in \mathrm{Cov}(A)$ iff $\bigcup_{i\in I} U_i = U$. The concept of a site is seen in the light of this example, as a reformulation and generalization of the notion of topology obtained by "eliminating" points of the space. From the point of view of topology, the success of the concept, of course, depends to what extent relevant constructions of topology can be formulated by it.

Conditions (i) to (iv) on Grothendieck topologies are *closure conditions*. In particular, if $\mathrm{Cov}_0(A)$ is *any* class of families $\langle A_i \to A\rangle_{i\in I}$, for each $A \in \mathrm{Ob}(\mathcal{C})$, we can talk about the Grothendieck topology on $\mathcal{C}$ generated by the *basic covering families* in the $\mathrm{Cov}_0(A)$, namely, we take the smallest Grothendieck topology containing the $\mathrm{Cov}_0(A)$.

Another formulation of the notion of Grothendieck topology uses the formalism of the Yoneda embedding, cf. CWM. We call any contravariant functor $F : \mathcal{C} \to \mathbf{Set}$ (**Set** is the category of sets) a *presheaf* over $\mathcal{C}$.

The category of all presheaves (with natural transformations as morphisms) is denoted by $\widehat{\mathcal{C}}$.

Remark. If $\mathcal{C}$ is a small category, $\widehat{\mathcal{C}}$ is a locally small category. If $\mathcal{C}$ is only locally small, $\widehat{\mathcal{C}}$ is not necessarily locally small. We use $\widehat{\mathcal{C}}$, nevertheless, for arbitrary locally small $\mathcal{C}$; we have to pass to another 'universe' $\mathcal{V}$ if we want $\widehat{\mathcal{C}}$ to be, say, a $\mathcal{V}$-category (cf. SGA4). This 'foundational' difficulty is not a serious one. On the other hand, one has to exercise caution in connection with $\widehat{\mathcal{C}}$ when applying certain constructions simply for reasons like, e.g., **Set**, the category of sets, has limits of *small* diagrams only.

For any $A \in \mathrm{Ob}(\mathcal{C})$, the functor $h_A = \mathrm{Hom}_{\mathcal{C}}(-, A)$ is a presheaf, called a *representable* presheaf (*represented* by A).

Given an object A of $\mathcal{C}$ and a presheaf F, an element $\alpha \in FA$ gives rise to a natural transformation

$$\overline{\alpha} : h_A \to F$$

defined thus: $\overline{\alpha}_B : \mathrm{Hom}(B, A) \to FB$ is the map $B \xrightarrow{g} A \mapsto (F(g))(\alpha)$. In particular, we have the functor $h_{(-)} : \mathcal{C} \to \widehat{\mathcal{C}}$ such that for $A \xrightarrow{f} A'$, $h_f = \overline{f}$; here $f \in FA$ for $F = h_{A'}(-) = \mathrm{Hom}(-, A')$. A part of Yoneda's lemma says that $h_{(-)}$ is full and faithful; because of this, it is customary to *identify* A and h_A, and also f with h_f when dealing with $\mathcal{C}$ and $\widehat{\mathcal{C}}$ at the same time. Returning to $\overline{\alpha} : h_A \to F$ for $\alpha \in F(A)$, we note that it is also customary to identify α and $\overline{\alpha}$. The full force of Yoneda's lemma then states that the arrows $A \to F$ (i.e., the arrows $h_A \to F$) in $\widehat{\mathcal{C}}$ where A is an object of $\mathcal{C}$, F is a presheaf, are precisely the elements of $F(A)$. Using these identifications, we can also say e.g. that a morphism $F \xrightarrow{\eta} G$ of presheaves is given if we know what the *compositions* $\eta \circ f$ are, for all

$$\begin{tikzcd} F \arrow[r, "\eta"] & G \\ A \arrow[u, "f"] \arrow[ur, "\eta\circ f"'] & \end{tikzcd}$$

objects A in $\mathcal{C}$ and morphisms $f : A \to F$ in $\widehat{\mathcal{C}}$.

A subobject *in* $\widehat{\mathcal{C}}$ $R \hookrightarrow A$ of an object A of $\mathcal{C}$ (identified with h_A) is called a sieve (French 'crible') in SGA4. (As usual, a subobject of an object is an isomorphism class of monomorphisms with codomain the given object.) A sieve on A is given if for any $B \in \mathrm{Ob}(\mathcal{C})$ a subset $R(B)$ of $\mathrm{Hom}(B, A)$ is given such that whenever $B \xrightarrow{f} B'$ is a morphism in $\mathcal{C}$, then $g \in R(B')$ implies that $g \circ f \in R(B)$. I.e., any subobject $R \hookrightarrow A$ is represented by a monomorphism $R \xrightarrow{i} A$ where each $R(B) \xrightarrow{i_B} \mathrm{Hom}(B, A)$ is a set-theoretic inclusion.

Now, the alternative definition of a Grothendieck topology on $\mathcal{C}$ is as follows. It is given by a set $J(A)$ of sieves on A, for any object A in $\mathcal{C}$, such that the following three conditions are satisfied.

1.1.2 (i) $A \xhookrightarrow{\mathrm{id}_A} A$ belongs to $J(A)$.

1.1.2 (ii) ("Stability under pullbacks"). If $R \in J(A)$, $f : B \to A$ is a morphism in $\mathcal{C}$, then $R \times_A B \hookrightarrow B$ in a pullback diagram

$$\begin{tikzcd} R \arrow[r, hook] & A \\ R \times_A B \arrow[u] \arrow[r, hook] & B \arrow[u, "f"'] \end{tikzcd}$$

belongs to $J(B)$.

1.1.2 (iii) ("Local character"). Let R, R' be sieves on A, $R \in J(A)$. Assume that for every object B of $\mathcal{C}$ and every morphism $B \to R$, the sieve $R' \times_A B$ belongs to $J(B)$. Then R' belongs to $J(A)$.

Note that using 1.1.2 (i) and (iii) we can infer that if $R \in J(A)$, and R' is a sieve such that $R \leq R'$ (in the partial ordering of subobjects of A), then $R' \in J(A)$. Similarly, if $R_1, R_2 \in J(A)$ then $R_1 \cap R_2 = R_1 \times_A R_2 \in J(A)$. In other words, $J(A)$ is a *filter* on the set of subobjects of A in $\widehat{\mathcal{C}}$.

We describe how to pass from one way of specifying a Grothendieck topology to the other. Given a family $\alpha = \langle A_i \xrightarrow{f_i} A \rangle_{i \in I}$, α *generates* the sieve $R \xhookrightarrow{i} A$ defined as follows. For each $B \in \mathrm{Ob}(\mathcal{C})$, $R(B) \xhookrightarrow{i_B} \mathrm{Hom}(B, A)$ is the inclusion. $R(B)$ consists of those morphisms $g : B \to A$ such that g factors through an f_i: there are $i \in I$ and $h : B \to A_i$ such that $g = f_i \circ h$. Denote this R by $R[\alpha]$. Then we have

Proposition 1.1.3 (i) *Given a Grothendieck topology on* $\mathcal{C}$ *according to* 1.1.1*, the equalities* $J(A) = \{R[\alpha] : \alpha \in \mathrm{Cov}(A)\}$ *for* $A \in \mathrm{Ob}(\mathcal{C})$ *define a Grothendieck topology according to* 1.1.2.

(ii) *Given a Grothendieck topology on* $\mathcal{C}$ *according to* 1.1.2, *the equalities* $\mathrm{Cov}(A) = \{\alpha : R[\alpha] \in J(A)\}$ *define one according to* 1.1.1.

(iii) *Moreover, the correspondences exhibited are inverses of each other.*

We leave the verifications to the reader.

Definition 1.1.4 *Given two sites* $\mathcal{C}, \mathcal{C}'$, *a* continuous functor $F : \mathcal{C} \to \mathcal{C}'$ *is one that preserves finite left limits (for this we also say: F is* left exact*) and preserves coverings. The latter means that if* $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I} \in \mathrm{Cov}_{\mathcal{C}}(A)$, *then* $\langle FA_i \xrightarrow{Ff_i} FA\rangle_{i\in I} \in \mathrm{Cov}_{\mathcal{C}'}(F(A))$.

Proposition 1.1.5 *Let the site* $\mathcal{C}$ *be given by basic covering families, i.e., a collection* $\mathrm{Cov}_0(A)$ *for each* $A \in \mathrm{Ob}(\mathcal{C})$ *(c.f. above). Suppose* $F : \mathcal{C} \to \mathcal{C}'$ *preserves finite left limits and it satisfies the continuity condition for families* $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ *in* $\mathrm{Cov}_0(A)$. *Then* F *is continuous.*

The proof is easy and left to the reader.

Returning to the example of sites defined by topological spaces, let $\mathcal{C}, \mathcal{C}'$ be the sites defined by the respective topological spaces X and X'. Let $f : X \to X'$ be a continuous map. Notice that the 'inverse image functor' $F : \mathcal{C}' \to \mathcal{C}$ defined by $F(U) = f^{-1}(U) \subset X$ is continuous according to 1.1.4. Under mild conditions on the spaces (that are satisfied if e.g. they are Hausdorff), any continuous $F : \mathcal{C}' \to \mathcal{C}$ corresponds to a unique continuous $f : X \to X'$.

Anticipating our "identification" of a site with a theory (explicitly in Chapter 6), we also call a continuous functor $\mathcal{C} \to \mathcal{D}$ between sites a $\mathcal{D}$-*model* of $\mathcal{C}$.

Definition 1.1.6 (i) *A* compatible family *of morphisms* from *a covering* $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ *to the presheaf F is a family* $\langle A_i \xrightarrow{\xi_i} F\rangle_{i\in I}$ *such that in*

$$\begin{tikzcd}
 & A_i \arrow[dr] \arrow[drr, "\xi_i"] & & \\
A_i \times_A A_j \arrow[ur, "p_1"] \arrow[dr, "p_2"'] & & A & F \\
 & A_j \arrow[ur] \arrow[urr, "\xi_j"'] & &
\end{tikzcd}$$

we have $\xi_i \circ p_1 = \xi_j \circ p_2$, *for any pair* i, j *of indices in* I.

(ii) *A presheaf F is a* (set-valued) sheaf *for the site* $\mathcal{C}$ *if whenever* $A \in \mathrm{Ob}(\mathcal{C})$, $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I} \in \mathrm{Cov}(A)$, *and* $\langle A_i \xrightarrow{\xi_i} F\rangle_{i\in I}$ *is a compatible family from* $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$, *then there is a unique morphism* $\xi : A \to F$ *such that* $\xi_i = \xi \circ f_i$ *for* $i \in I$.

(iii) *The presheaf F is called a* separated presheaf *if in* (ii), *there is* at most one ξ *as stated there.*

The reader who is familiar with the notion of sheaf over a topological space should check that that notion coincides with the one given here, taking the site to be the one derived from the topological space.

Another way of putting the definition of a sheaf is this. Consider a presheaf F and a family $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$. In **Set**, consider the following diagram

$$\mathrm{Hom}_{\widehat{\mathcal{C}}}(A, F) \xrightarrow{u} \prod_{i\in I} \mathrm{Hom}_{\widehat{\mathcal{C}}}(A_i, F) \underset{v_2}{\overset{v_1}{\rightrightarrows}} \prod_{i,j\in I} \mathrm{Hom}_{\widehat{\mathcal{C}}}(A_i \times_A A_j, F).$$

Here u is defined by

$$u(\xi) = \langle \xi \circ f_i : i \in I\rangle$$

(for $\xi : A \to F$), v_1 and v_2 are defined by

$$v_1(\langle g_i : i \in I\rangle) = \langle g_i \circ p_1^{i,j} : i, j \in I\rangle,$$

$$v_2(\langle g_i : i \in I \rangle) = \langle g_j \circ p_2^{i,j} : i, j \in I \rangle;$$

here $p_1^{i,j} = p_1$, $p_2^{i,j} = p_2$ are as in 1.1.6(i). Now the reader can check that F being a sheaf is equivalent to saying that u is the *equalizer* of the maps v_1, v_2, whenever $\langle A_i \xrightarrow{f_i} A \rangle_{i \in I}$ is a covering family. For this we say that the above diagram is *exact*.

To connect the notion of sheaf to the "sieve"-formulation of topologies, we note that compatible families from $\alpha = \langle A_i \xrightarrow{f_i} A \rangle_{i \in I}$ to F are in 1-1 correspondence with morphisms $R[\alpha] \to F$, where $R[\alpha]$ is the sieve generated by α (see above). Namely, if $\langle A_i \xrightarrow{\xi_i} F \rangle_{i \in I}$ is a compatible family, then the maps

$$\eta_B : R[\alpha](B) \to FB$$

defined by: $B \longrightarrow A_i \xrightarrow{f_i} A \quad \overset{\eta_B}{\longmapsto} \quad B \longrightarrow A_i \xrightarrow{\xi_i} F$ will combine to give the natural transformation $\eta : R[\alpha] \to F$. (Compatibility is used to show that the η_B are well-defined.) η is the unique morphism $R[\alpha] \xrightarrow{\eta} F$ in $\widehat{\mathcal{C}}$ such that

$$\begin{tikzcd}
A_i \arrow[r, "\widehat{f}_i"] \arrow[rr, bend left, "f_i"] \arrow[dr, "\xi_i"'] & R[\alpha] \arrow[r, hook, "\text{inclusion}"'] \arrow[d, "\eta"] & h_A = A \\
 & F &
\end{tikzcd}$$

is commutative for every $i \in I$.

Next, we mention a fact related to specifying topologies by "basic coverings". Let $\mathrm{Cov}_0(A)$ be a collection of families of morphisms with codomain A, for each $A \in \mathrm{Ob}(\mathcal{C})$, and assume that $\langle \mathrm{Cov}_0(A) : A \in \mathrm{Ob}(\mathcal{C}) \rangle$ satisfies 1.1.1(ii), stability under pullbacks. Let $\langle \mathrm{Cov}(A) \rangle_A$ be the Grothendieck topology generated by $\langle \mathrm{Cov}_0(A) \rangle_A$. Now, we can repeat the definition of sheaves using the $\mathrm{Cov}_0(A)$ instead of the $\mathrm{Cov}(A)$. But we have

Proposition 1.1.7 *F is a sheaf for the topology defined by the basic coverings in the $\mathrm{Cov}_0(A)$, iff it is one "relative to $\langle \mathrm{Cov}_0(A) \rangle_A$", if the latter is stable under pullbacks.*

PROOF. First we claim that the topology $\langle \mathrm{Cov}(A) \rangle_A$ generated by $\langle \mathrm{Cov}_0(A) \rangle_A$ is obtained as the 'smallest' collection $\langle \mathrm{Cov}(A) \rangle_A$ such that each $\mathrm{Cov}(A)$ contains $\mathrm{Cov}_0(A)$ and the conditions 1.1.1(i), 1.1.1(iii) and 1.1.1(iv) are satisfied. The thing to prove is that this $\langle \mathrm{Cov}(A) \rangle_A$ will then satisfy 1.1.1(ii) as well. This is proved by "induction" for families $\alpha = \langle A_i \xrightarrow{f_i} A \rangle_{i \in I} \in \mathrm{Cov}(A)$. For $\alpha \in \mathrm{Cov}_0(A)$, stability is true by hypothesis and one is left to show that it remains true for α "obtained" in clauses (i), (iii), (iv), once stability is true for the ones entering the construction of α in the clause. The details are easy and are omitted.

The second thing to show is that if for a given presheaf F, the 'sheaf property' for compatible families from coverings entering each of the clauses (i), (iii), (iv) holds, *then* it holds for compatible families from the coverings "obtained" by the clause. Using the first claim, by 'induction' again this will show that the 'sheaf property' is inherited from $\langle \mathrm{Cov}_0(A) \rangle_A$ to $\langle \mathrm{Cov}(A) \rangle_A$. Again, we omit the easy details.

The last proposition leads us to the notion of the *canonical topology* on a given $\mathcal{C}$. Remember that, for us, every category has finite limits, in particular pullbacks. In the following definition, $\mathcal{C}$ is fixed.

Definition 1.1.8 (i) *On a given category $\mathcal{C}$, for two topologies T_1, T_2, with respective classes $\mathrm{Cov}^{(T_1)}(A)$, $\mathrm{Cov}^{(T_2)}(A)$ of covering families for $A \in \mathrm{Ob}(\mathcal{C})$, we say that T_2 is* finer *than* T_1 *if* $\mathrm{Cov}^{(T_1)}(A) \subset \mathrm{Cov}^{(T_2)}(A)$ *for every* $A \in \mathrm{Ob}(\mathcal{C})$.

(ii) *Given a family* $\alpha = \langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ *of morphisms with common codomain* A*, and a presheaf* F*, we say that* F has the sheaf-property with respect to α *if the condition appearing in* 1.1.6(ii) *is satisfied, with the given* α *and* F *and for every compatible family from* α *to* F*. (Note that we can reformulate this condition by an exact diagram like the one after* 1.1.6.*)*

(iii) *A family* $\alpha = \langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ *is called* effective epimorphic *if every representable presheaf* h_B *(*$B \in \mathrm{Ob}(\mathcal{C})$*) has the sheaf property with respect to* α*. (The reader is invited to reformulate this condition in purely categorical terms in* $\mathcal{C}$*.)*

(iv) *A family* $\alpha = \langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ *is called* stable *(or:* universal*)* effective epimorphic *if for every morphism* $B \to A$*, the pullbacks*

$$\begin{tikzcd} B \arrow[r] & A \\ A_i \times_A B \arrow[u, "g_i"] \arrow[r] & A_i \arrow[u, "f_i"'] \end{tikzcd}$$

form an effective epimorphic family $\alpha = \langle A_i \times_A B \xrightarrow{g_i} B\rangle_{i\in I}$.

(v) *A family* $\alpha = \langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ *is* epimorphic *if for any pair* $A \overset{g}{\underset{h}{\rightrightarrows}} B$*, if* $g \circ f_i = h \circ f_i$ *for all* $i \in I$*, then* $g = h$*. (Note that any effective epimorphic family is epimorphic.)*

Proposition 1.1.9 *Given* $\mathcal{C}$ *with finite left limits, there is a unique finest topology in which every representable presheaf is a sheaf. The covering families in this topology are exactly the stable effective epimorphic families. This is called the* canonical topology *on* $\mathcal{C}$*.*

Proof. It is clear from the definitions that in every topology for which the representable presheaves are sheaves, every covering family α must be effective epimorphic, and since the topology is "stable under pullback" (1.1.1(ii)), α must be actually stable effective epimorphic. It remains to show that the stable effective epimorphic families form a topology. Let T be the topology *generated* by those families. Now apply 1.1.7 to conclude that every representable presheaf is actually a sheaf in the topology T. But that means that every covering in T is a stable effective epimorphic family. $\square$

§2. The associated sheaf

Let $\mathcal{C}$ be a locally small site. The *category of sheaves* $\widetilde{\mathcal{C}}$ over $\mathcal{C}$ is defined as the full subcategory of $\widehat{\mathcal{C}}$ whose objects are the sheaves over $\mathcal{C}$. Our aim is to show

Theorem 1.2.1 *The inclusion functor* $i : \widetilde{\mathcal{C}} \to \widehat{\mathcal{C}}$ *has a left adjoint* $a : \widehat{\mathcal{C}} \to \widetilde{\mathcal{C}}$*, which preserves finite left limits.*

Remark *For a presheaf* F*,* $a(F)$ *is called the sheaf* associated *to* F*.*

The proof of 1.2.1 will occupy the entire section.

First, let us note that the assertion of 1.2.1 is equivalent to saying that there is a functor $a : \widehat{\mathcal{C}} \to \widetilde{\mathcal{C}}$ preserving finite left limits and there is a natural transformation $\alpha : \mathrm{id}_{\widehat{\mathcal{C}}} \to i \circ a$ such that for every sheaf S in $\widetilde{\mathcal{C}}$, $\alpha_S : S \to a(S)$ (note that $S = iS$, $a(S) = iai(S)$) is an isomorphism and for every $f : F \to S$ where F is a presheaf and S

is a sheaf, *there is at most one* $g : a(F) \to S$ such that

$$\begin{tikzcd} F \arrow[r, "f"] \arrow[d, "\alpha_F"'] & S \\ a(F) \arrow[ur, "g"'] & \end{tikzcd}$$

commutes. (As the reader will see, from these facts it follows e.g. by Theorem 2, part (i) on page 81 in CWM that α is the unit of a suitable adjunction $\langle a, i, \phi\rangle$. The 'universal property' of $a(F)$ will be that for any f as above, *there is a unique* g making the previous diagram commute. Of course, g can be taken to be $g = \alpha_S^{-1} \circ a(f)$, from

$$\begin{tikzcd} F \arrow[r, "f"] \arrow[d, "\alpha_F"'] & S \arrow[d, "\alpha_S", "\cong"'] \\ a(F) \arrow[ur, "g"'] \arrow[r, "a(f)"'] & a(S) \end{tikzcd} \quad .)$$

We construct a and α by first constructing a functor $L : \widehat{\mathcal{C}} \to \widehat{\mathcal{C}}$ and a natural transformation $\ell : \mathrm{id}_{\widehat{\mathcal{C}}} \to L$ and finally putting $a = L^2$, $\alpha = \ell_{L(-)} \circ \ell$.

Definition 1.2.2 *Definition of L and ℓ.*
$L(F)$ is the presheaf defined by clauses (i) *and* (ii).
(i) *The effect of the functor $L(F) : \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ on objects of $\mathcal{C}$ is given by*

$$L(F)(A) = \varinjlim_{R \in J(A)^{\mathrm{op}}} \mathrm{Hom}_{\widehat{\mathcal{C}}}(R, F).$$

In other words, a typical morphism $A \to L(F)$ is of the form $\dot{\xi}$, for an arbitrary compatible family $\xi : R \to F$, for $R \in J(A)$, and we have $\dot{\xi}_1 = \dot{\xi}_2$, with $\xi_i : R_i \to F$, $R_i \in J(A)$ $(i = 1, 2)$ if and only if there is $R \leq R_1, R_2$, $R \in J(A)$, such that the two composites in
$\begin{tikzcd}[row sep=tiny] & R_1 \arrow[dr, "\xi_1"] & \\ R \arrow[ur] \arrow[dr] & & F \\ & R_2 \arrow[ur, "\xi_2"'] & \end{tikzcd}$ *coincide.*

Remark *In case $\mathcal{C}$ is a small site, $J(A)$ is a set, so the limit defining $L(F)(A)$ exists in* **Set**. *If $\mathcal{C}$ is only locally small as a site, we should note the following. We have a topologically generating set G of objects of $\mathcal{C}$. Denoting by $J_G(A)$ the covering sieves of A generated by coverings with morphisms having domains in G, it follows easily that* (a) *$J_G(A)$ is a set for each $A \in \mathrm{Ob}(\mathcal{C})$ and* (b) *$J_G(A)$ as a partially ordered subset of $J(A)$ is coinitial in $J(A)$, i.e. for every $R \in J(A)$ there is $R' \in J_G(A)$ such that $R' \leq R$. It follows that the limit defining $L(F)(A)$ is identical to the* small *limit obtained by replacing $J(A)^{\mathrm{op}}$ by $J_G(A)^{\mathrm{op}}$, so it exists in* **Set**, *hence the definition works for the locally small case as well.*

(ii) *The effect of $L(F)$ on a morphism $B \xrightarrow{f} A$ in $\mathcal{C}$, i.e. $g = L(F)(f)$, is as follows. For a typical element $\dot{\xi} : A \to L(F)$, with ξ a compatible family $R \to F$, $R \in J(A)$, $g(\dot{\xi})$ is defined as $\dot{\eta}$, where η is from the commutative diagram*

$$\begin{tikzcd} & F & \\ R \arrow[ur, "\xi"] \arrow[rr, hook] & & A \\ R \times_A B \arrow[u] \arrow[uur, bend left=60, "\eta"] \arrow[rr, hook] \arrow[urr, phantom, "\text{p.b.}"] & & B. \arrow[u, "f"'] \end{tikzcd}$$

(It is easy to see that this definition is correct. The dot on ξ refers to the definition of $L(F)(A)$, that on η refers to $L(F)(B)$.)

Remark *If ξ is given as $\langle A_i \xrightarrow{\xi_i} F\rangle_{i\in I}$, with the covering $\langle A_i \to A\rangle_{i\in I}$, then η is $\langle B_i \xrightarrow{\eta_i} F\rangle_{i\in I}$ from the commutative diagrams ($i \in I$)*

$$\begin{tikzcd}
 & & F \\
A_i \arrow[urr, "\xi_i"] \arrow[rr] & & A \\
B_i \arrow[u] \arrow[rr] \arrow[uurr, bend left, "\eta_i"] \arrow[urr, phantom, "\text{p.b.}"] & & B. \arrow[u, "f"']
\end{tikzcd}$$

(iii) *The effect of the functor L on a morphism $\nu : F \to G$ is defined as follows. For $\xi : R \to F$, $R \in J(A)$, we put $(L(\nu))_A(\dot{\xi}) = \dot{\eta}$ where η is the composite*

$$R \xrightarrow{\xi} F \xrightarrow{\nu} G.$$

Again, it is easy to see that the definition is correct.

(iv) *The morphism $\ell_F : F \to LF$ is defined as follows. For $A \in \mathrm{Ob}(\mathcal{C})$, and a morphism $\xi : A \to F$, ξ is a (one-element) compatible family, "from" the identical covering $A \xrightarrow{\mathrm{id}_A} A$. Hence $\dot{\xi}$ is a morphism $A \to L(F)$. Put $((\ell_F)_A)(\xi) = \dot{\xi}$.*

Remarks *The construction of $L(F)$ represents a natural attempt towards constructing the associated sheaf $a(F)$. Certainly, in $a(F)(A)$ there should be a "representative" $\dot{\xi}$ of each compatible family $R \xrightarrow{\xi} F \longrightarrow a(F)$* and hence *a representative of each compatible family $R \xrightarrow{\xi} F$ for $R \in J(A)$. This is done by introducing $\dot{\xi}$ as we did in $L(F)(A)$. Moreover, compatible families that, after refinement of the coverings, become the same, should clearly correspond to the same morphism $A \to a(F)$; hence the definition of equality $\dot{\xi}_1 = \dot{\xi}_2$. The difficulty is, however, that by this construction we have not taken care of* all *compatible families $R \xrightarrow{\xi} L(F)$, only those that factor through F; so, $L(F)$ is not necessarily a sheaf. On the other hand, we will see that* (a) *$L(F)$ is always a separated presheaf, and* (b) *if F is a separated presheaf, then $L(F)$* is *a sheaf, after all. This explains the construction of $a(F)$ as stated above.*

Proposition 1.2.3 *L is a functor $\widehat{\mathcal{C}} \to \widehat{\mathcal{C}}$ and ℓ is a natural transformation $\mathrm{id}_{\widehat{\mathcal{C}}} \to L$.*

The proof is by careful inspection of the definitions 1.2.2.

Proposition 1.2.4 (i) *If F is separated, the $F \xrightarrow{\ell_F} LF$ is a monomorphism.*

(ii) *If F is a sheaf, then the morphism $F \xrightarrow{\ell_F} LF$ in $\widehat{\mathcal{C}}$ is an isomorphism.*

The proofs are quite obvious on the basis of the definitions.

Proposition 1.2.5 *For any presheaf F, $L(F)$ is a separated presheaf.*

Proof. Let $A \overset{f}{\underset{g}{\rightrightarrows}} L(F)$ be two morphisms and assume that $R \overset{i}{\hookrightarrow} A$, $R \in J(A)$, is such that $fi = gi$. We want to conclude that $f = g$. By the definition of $L(F)$, $f = \dot{\xi}$ and $g = \dot{\eta}$ for some $\xi : R_1 \to F$, $\eta : R_2 \to F$ with $R_1, R_2 \in J(A)$. Using the filter property of $J(A)$ and the definition for equality of the $\dot{\xi}$, we can assume without loss of generality that $R = R_1 = R_2$. Consider an arbitrary morphism of the form $B \xrightarrow{\beta} R$. We compute

$\dot{\xi} i\beta = \dot{\xi}(i\beta)$ according to the definition of $L(F)$ as a functor. We take a pullback

$$\begin{tikzcd}
R \arrow[r, hook, "i"] & A \\
R \times_A B \arrow[u] \arrow[r] & B. \arrow[u, "i\beta"']
\end{tikzcd}$$

But for this we can take

$$\begin{tikzcd}
R \arrow[r, hook, "i"] & A \\
B \arrow[u, "\beta"] \arrow[r, "\mathrm{id}_B"'] & B. \arrow[u, "i\beta"']
\end{tikzcd}$$

Then $\dot{\xi}(i\beta)$ is defined as $(\dot{\xi\beta})$ where now the dot refers to the definition of $(LF)(B)$ and $\xi\beta$ is the compatible family $B \to F$ from the identity covering $B \xrightarrow{\mathrm{id}_B} B$.

We conclude that, since $\dot{\xi} i = \dot{\eta} i$, and thus $\dot{\xi} i\beta = \dot{\eta} i\beta$, we have $(\dot{\xi\beta}) = (\dot{\eta\beta})$ for an arbitrary morphism $B \xrightarrow{\beta} R$. Given such a β, according to the definition of "$(\dot{\xi\beta}) = (\dot{\eta\beta})$", this means that there is a covering $R_\beta \in J(B)$, $R_\beta \overset{i_\beta}{\hookrightarrow} B$, such that $\xi\beta i_\beta = \eta\beta i_\beta$. Choose and fix R_β for any $\beta : B \to R$.

Next, define the sieve $R' \xrightarrow{j} A$ as follows. We define R' such that $j_C : R'(C) \to \mathrm{Hom}(C, A)$ will be the inclusion, for every C. The morphism $C \xrightarrow{\gamma} A$ is put into $R'(C)$ if and only if there is a morphism $B \xrightarrow{\beta} R$ such that γ factors through i_β: there is $\gamma' : C \to R_\beta$ such that the following diagram commutes:

$$\begin{tikzcd}
C \arrow[rr, "\gamma"] \arrow[dd, "\gamma'"'] & & A \\
& R \arrow[ur, hook, "i"'] & \\
R_\beta \arrow[r, hook, "i_\beta"] & B \arrow[u, "\beta"'] &
\end{tikzcd}$$

It is easy to check that (i) this legitimately defines R', a sieve of A, (ii) $R' \le R$ and (iii) $R_\beta \le R' \times_B A$ for every $B \xrightarrow{\beta} R$. Hence $R' \times_A B \in J(B)$, with any $B \xrightarrow{\beta} R$, and hence by 1.1.2(iii) ("local character") $R' \in J(A)$.

Finally, we claim that, for $R' \xrightarrow{j'} R$, we have $\xi \circ j' = \eta \circ j'$. This will establish, according to the definition of $\dot{\xi} = \dot{\eta}$, that indeed $f = \dot{\xi} = \dot{\eta} = g$.

To prove the claim, it is enough to show that for any $C \xrightarrow{\gamma''} R'$ (with $C \in \mathrm{Ob}(\mathcal{C})$), we have $\xi \circ j' \circ \gamma'' = \eta \circ j' \circ \gamma''$. Consider the following diagram:

$$\begin{tikzcd}
C \arrow[rrr, "\gamma"] \arrow[dr, "\gamma''"'] \arrow[ddd, "\gamma'"'] & & & A \\
& R' \arrow[urr, "j"] \arrow[dr, hook, "j'"'] & & \\
& & R \arrow[ur, hook, "i"] \arrow[r, shift left, "\xi"] \arrow[r, shift right, "\eta"'] & F \\
R_\beta \arrow[rr, "i_\beta"'] & & B \arrow[u, "\beta"'] &
\end{tikzcd}$$

By definition, there are $B \xrightarrow{\beta} R$ and $\gamma' : C \to R_\beta$ such that $\gamma = i \circ \beta \circ i_\beta \circ \gamma'$ where $\gamma = j \circ \gamma''$. Hence $j' \circ \gamma'' = \beta \circ i_\beta \circ \gamma'$. Since $\xi \circ \beta \circ i_\beta = \eta \circ \beta \circ i_\beta$, it follows that $\xi \circ j' \circ \gamma'' = \eta \circ j' \circ \gamma''$ as desired. $\square$

Next we formulate a lemma that 'eliminates' the dots in the definition of LF.

Lemma 1.2.6 (i) *If $\xi : R \to F$ is a compatible family, $R \in J(A)$, $\dot{\xi}$ is the 'class' of ξ in the definition of $(LF)(A)$, then the diagram*

$$\begin{tikzcd} A \arrow[r, "\dot{\xi}"] & LF \\ R \arrow[u, hook, "i"] \arrow[r, "\xi"'] & F \arrow[u, "\ell_F"'] \end{tikzcd}$$

commutes.

(ii) *For every $R \in J(A)$ and $u : R \to F$, there is a morphism $v : A \to LF$ such that*

$$\begin{tikzcd} A \arrow[r, "v"] & LF \\ R \arrow[u, hook, "i"] \arrow[r, "u"'] & F \arrow[u, "\ell_F"'] \end{tikzcd}$$

commutes.

(iii) *For every morphism $v : A \to LF$ there is $R \in J(A)$ and $u : R \to F$ such that the diagram in* (ii) *commutes.*

(iv) *If $A \overset{u_1}{\underset{u_2}{\rightrightarrows}} F$ are two morphisms, $A \in \mathrm{Ob}(\mathcal{C})$ and $F \in \mathrm{Ob}(\widehat{\mathcal{C}})$ and $\ell_F \circ u_1 = \ell_F \circ u_2$, then the equalizer $R \longrightarrow A \overset{u_1}{\underset{u_2}{\rightrightarrows}} F$ is a covering, $R \hookrightarrow A \in J(A)$.*

Proof. (ii), (iii) and (iv) are immediate consequences of (i) and the definitions. To prove (i), we consider an arbitrary morphism $B \xrightarrow{\beta} R$ in $\widehat{\mathcal{C}}$ such that $B \in \mathrm{Ob}(\mathcal{C})$ and we prove that in

$$\begin{tikzcd} & A \arrow[r, "\dot{\xi}"] & LF \\ & R \arrow[u, hook, "i"] \arrow[r, "\xi"'] & F \arrow[u, "\ell_F"'] \\ B \arrow[ur, "\beta"] \end{tikzcd}$$

we have $\dot{\xi} \circ i \circ \beta = \ell_F \circ \xi \circ \beta$. By the definition of ℓ_F, we have that $\ell_F \circ \xi \circ \beta = (\xi \dot{\circ} \beta)$, where the dot now refers to the definition of $LF(B)$ and $\xi \circ \beta$ is meant as the compatible family $\xi \circ \beta : B \to F$ from the identical covering $B \xrightarrow{\mathrm{id}_B} B$. Turning to the left-side of the proposed equality, note that we have shown in the proof of 1.2.5 that $\dot{\xi}(i\beta)$ is $(\dot{\xi\beta})$ just as desired. $\square$

Proposition 1.2.7 *If $h_1 \circ \ell_F = h_2 \circ \ell_F$ in*

$$\begin{tikzcd} LF \arrow[r, shift left, "h_1"] \arrow[r, shift right, "h_2"'] & S \\ F \arrow[u, "\ell_F"] \end{tikzcd}$$

where S is a sheaf (or just a separated presheaf), then $h_1 = h_2$.

Proof. To show that $h_1 = h_2$, it is enough to show that $h_1 \circ v = h_2 \circ v$ for every morphism $v : A \to LF$. Take such a v. By 1.2.6(iii), we have $R \in J(A)$ and $u : R \to F$

such that

$$\begin{tikzcd}
A \arrow[r, "v"] & LF \arrow[r, shift left, "h_1"] \arrow[r, shift right, "h_2"'] & S \\
R \arrow[u, hook, "i"] \arrow[r, "u"'] & F \arrow[u, "\ell_F"']
\end{tikzcd}$$

commutes. From $h_1 \circ \ell_F = h_2 \circ \ell_F$ it follows that $h_1 \circ v \circ i = h_2 \circ v \circ i$. Since $R \in J(A)$ and S is separated $h_1 \circ v = h_2 \circ v$ as desired. □

Proposition 1.2.8 *If F is separated, then LF is a sheaf.*

Proof. Suppose that F is separated. By 1.2.4(i), ℓ_F is a monomorphism. To show that LF is a sheaf, let $R \xhookrightarrow{i} X \in J(X)$ $(X \in \mathrm{Ob}(\mathcal{C}))$ $u\colon R \to LF$ a 'compatible family'. Form the pullback

$$\begin{tikzcd}
F \arrow[r, hook, "\ell_F"] & LF \\
F \times_{LF} R = R' \arrow[u] \arrow[r, hook, "j"'] & R. \arrow[u, "u"']
\end{tikzcd}$$

Since ℓ_F is a monomorphism, so is j. We claim that actually, $R' \xhookrightarrow{i \circ j} X$ is a covering. By 1.1.2(iii), it is enough to show that for any $Y \to R$ $(Y \in \mathrm{Ob}(\mathcal{C}))$ we have $R'' = R' \times_R Y \hookrightarrow Y \in J(Y)$. Notice that $R'' = F \times_{LF} Y$.

By 1.2.6(iii), there is $R''' \in J(Y)$ and a morphism $R''' \to F$ such that

$$\begin{tikzcd}
F \arrow[r, "\ell_F"] & LF \\
 & R \arrow[u] \\
R''' \arrow[uu] \arrow[r] & Y \arrow[u]
\end{tikzcd}$$

commutes. Since $R'' = F \times_{LF} Y$, it follows that $R''' \leq R''$, hence $R'' \in J(Y)$ as desired.

Use 1.2.6(ii) to find v such that

$$\begin{tikzcd}
F \arrow[r, hook, "\ell_F"] & LF & \\
R' \arrow[u] \arrow[r, hook, "j"'] & R \arrow[r, hook, "i"'] & X \arrow[ul, "v"']
\end{tikzcd}$$

is commutative. We claim that $u = v \circ i$, which will suffice to show that LF is a sheaf (we already know that LF is separated). To establish the claim, take again an arbitrary morphism $Y \to R$ $(Y \in \mathrm{Ob}(\mathcal{C}))$ and form $R'' = R' \times_R Y \in J(Y)$.

$$\begin{tikzcd}
F \arrow[r, "\ell_F"] & LF & \\
R' \arrow[u] \arrow[r] \arrow[ur, phantom, "\circlearrowleft"] & R \arrow[u, "u"] \arrow[r, "i"] & X \arrow[ul, "v"'] \\
R'' \arrow[u] \arrow[r] \arrow[ur, phantom, "\circlearrowleft"] & Y \arrow[u] &
\end{tikzcd}$$

We have that the rectangles are commutative as shown as well as that the diagram preceding the present one is commutative. It follows that the composite morphisms

$R'' \to Y \to R \xrightarrow{u} LF$ and $R'' \to Y \to R \xrightarrow{i} X \xrightarrow{v} LF$ are equal. Since LF is separated, $Y \to R \xrightarrow{u} LF$ and $Y \to R \xrightarrow{i} X \xrightarrow{v} LF$ are equal. Since $Y \to R$ was arbitrary, it follows that $u = v \circ i$.

Proposition 1.2.9 *The functor $L : \widehat{\mathcal{C}} \to \widehat{\mathcal{C}}$ preserves finite left limits.*

Remark *The meaning of this statement is that whenever we have a "finite left limit diagram" in $\widehat{\mathcal{C}}$ then the* image *of this diagram (also in $\widehat{\mathcal{C}}$ in this case) is a "finite left limit diagram too". In more detail: let $\mathbf{J}$ be a fixed category (or even just a graph, c.f. CWM). A* diagram of type $\mathbf{J}$ in $\mathcal{C}$ *is a functor $\delta : \mathbf{J} \to \mathcal{C}$. Let, in particular, $\mathbf{J}$ be a category of the form of a "dual cone": with a distinguished object L, with a full subcategory $\mathbf{I}$ not containing L, but containing all other objects of $\mathbf{J}$, and with precisely one morphism ("projection") $\pi_A : L \to A$ for every $A \in \mathrm{Ob}(\mathbf{I})$. Now, the phrase: "the diagram $\delta : \mathbf{J} \to \mathcal{C}$* is a left limit diagram*" has the expected meaning: it means that $\delta(L)$ is the left (or: projective) limit of the diagram $\delta|_{\mathbf{I}} : \mathbf{I} \to \mathcal{C}$ with canonical projections $\delta(\pi_A) : \delta(L) \to \delta(A)$. To say that a functor $F : \mathcal{C} \to \mathcal{D}$ preserves* left limits of type $\mathbf{J}$ *means that whenever $\delta : \mathbf{J} \to \mathcal{C}$ is a left limit diagram, so is $F \circ \delta : \mathbf{J} \to \mathcal{D}$. This formulation has the advantage that it readily generalizes to situations such as preservation by a functor F of images, disjoint sums, etc.; in all these cases we have preservation under composition by F of a* property *relative to a category of a diagram indexed by a fixed graph $\mathbf{J}$. Finally, preservation of finite left limits means preservation of all left limit diagrams indexed by any finite graph $\mathbf{J}$.*

Proof of 1.2.9. This is an elementary argument based mainly on the fact that directed colimits can be interchanged with finite left limits, c.f. Theorem 1, p. 211 in CWM. Nonetheless, we state the main points of the argument.

Given categories Γ and $\mathcal{C}$, assume that all functors (diagrams) $\gamma : \Gamma \to \mathcal{C}$ have colimits: there is an object $R \in \mathcal{C}$ together with injections $\pi_c : \gamma(c) \to R$ for $c \in \mathrm{Ob}(\Gamma)$ with the well-known universal properties (the system $\langle \gamma(c) \xrightarrow{\pi_c} R \rangle_{c \in \mathrm{Ob}(\Gamma)}$ is called a *colimiting cone*, c.f. CWM). Then we can define the functor

$$\varinjlim : \mathcal{C}^\Gamma \to \mathcal{C}$$

($\mathcal{C}^\Gamma$ is the functor category of all functors $\Gamma \to \mathcal{C}$) as in Exercise 3, p. 110 in CWM as follows:

(i) for each $\gamma : \Gamma \to \mathcal{C} \in \mathrm{Ob}(\mathcal{C}^\Gamma)$ pick (using the axiom of choice) a *colimiting cone* $\langle \gamma(c) \xrightarrow{\pi_c} R \rangle_{c \in \mathrm{Ob}(\Gamma)}$ as above and put

$$\varinjlim(\gamma) = R$$

(ii) for each natural transformation $\nu : \gamma \to \gamma'$, define the morphism $f = \varinjlim(\nu)$ as follows. Let $\langle \gamma(c) \xrightarrow{\pi_c} R \rangle_{c \in \mathrm{Ob}(\Gamma)}$, $\langle \gamma(c) \xrightarrow{\pi'_c} R' \rangle_{c \in \mathrm{Ob}(\Gamma)}$ be the respective colimiting cones picked in (i). $f : R \to R'$ is defined to be the unique morphism such that $\pi'_c \circ \nu_c = f \circ \pi_c$ for all $c \in \mathrm{Ob}(\Gamma)$; we use of course the universal property of the colimiting cone $\langle \gamma(c) \xrightarrow{\pi_c} R \rangle_{c \in \mathrm{Ob}(\Gamma)}$.

It is easy to check that $\varinjlim$ is well defined.

Now, the main fact we need is as follows. Suppose Γ and $\mathcal{C}$ are as above, and in addition, Γ is filtered (c.f. CWM) and consider $\varinjlim : \mathcal{C}^\Gamma \to \mathcal{C}$. Let $\mathbf{J}$ be a category and let the functors $\Phi : \mathbf{J} \to \mathcal{C}^\Gamma$ and $\Phi' : \Gamma \to \mathcal{C}^{\mathbf{J}}$ be related in the obvious way: $(\Phi(j))(c) = (\Phi'(c))(j)$ for $j \in \mathrm{Ob}(\mathbf{J})$, $c \in \mathrm{Ob}(\Gamma)$, and similarly for morphisms. Suppose further that $\mathbf{J}$ is a finite category of the form of a dual cone, c.f. above. The claim is that if for all objects $c \in \Gamma$, $\Phi'(c) : \mathbf{J} \to \mathcal{C}$ is a left limit diagram of type $\mathbf{J}$ in $\mathcal{C}$, then $\varinjlim \circ \Phi : \mathbf{J} \to \mathcal{C}$ is one such too. This is seen to be a reformulation of Theorem 1, p. 211 in CWM, on "interchangeability" of limits.

Next, consider the definition of $L:\widehat{\mathcal{C}} \to \widehat{\mathcal{C}}$. Let A be a fixed object in the site $\mathcal{C}$, let Γ be the category $J(A)^{\mathrm{op}}$ (dual of the subcategory of $\widehat{\mathcal{C}}$ consisting of the covering sieves in $J(A)$ and morphisms of inclusion). Let $\varinjlim:\mathbf{Set}^{\Gamma} \to \mathbf{Set}$ be defined as above. Γ is directed (filtered). Let $\widehat{\mathcal{C}} \xrightarrow{\mathrm{ev}_A} \mathbf{Set}$ be the evaluation functor "$F \mapsto F(A)$" ($A \in \mathrm{Ob}(\mathcal{C})$) and let "$F \mapsto LF(A)$" be the composite $\mathrm{ev}_A \circ L:\widehat{\mathcal{C}} \to \mathbf{Set}$. Also, consider the Yoneda functor $Y = \mathrm{Hom}_{\widehat{\mathcal{C}}}(-,\cdot):\widehat{\mathcal{C}} \to \mathbf{Set}^{\widehat{\mathcal{C}}^{\mathrm{op}}}$, together with the *restriction* $\rho:\mathbf{Set}^{\widehat{\mathcal{C}}^{\mathrm{op}}} \to \mathbf{Set}^{\Gamma}$ (Γ being a subcategory of $\widehat{\mathcal{C}}^{\mathrm{op}}$) and let $\Psi = \rho \circ Y:\widehat{\mathcal{C}} \to \mathbf{Set}^{\Gamma}$. Inspection of the definition of L shows that, actually, the functor "$F \mapsto LF(A)$"$:\widehat{\mathcal{C}} \to \mathbf{Set}$ is nothing but the composite $\varinjlim \circ \Psi$.

We claim that the functor "$F \mapsto LF(A)$" preserves finite left limits. We use the above formulation, with $\varinjlim:\mathbf{Set}^{\Gamma} \to \mathbf{Set}$, of interchangeability of limits. Let $\delta:\mathbf{J} \to \widehat{\mathcal{C}}$ be a finite left limit diagram in $\widehat{\mathcal{C}}$. Define $\Phi:\mathbf{J} \xrightarrow{\delta} \widehat{\mathcal{C}} \xrightarrow{\Psi} \mathbf{Set}^{\Gamma}$, $\Phi = \Psi \circ \delta$. Inspection shows that $\Phi'(R)$ (with the above meaning for Φ') is $\Phi'(R):\mathbf{J} \to \mathbf{Set} = \mathbf{J} \xrightarrow{\delta} \widehat{\mathcal{C}} \xrightarrow{\mathrm{Hom}_{\widehat{\mathcal{C}}}(R,\cdot)} \mathbf{Set}$, for $R \in J(A) = \mathrm{Ob}(\Gamma)$. Hence, by Theorem 1, p. 112 in CWM ("hom functors preserve limits"), we have that $\Phi'(R)$ is a left limit diagram in $\mathbf{Set}$. Therefore, by "interchangeability", $\varinjlim \circ \Phi = (\varinjlim \circ \Psi) \circ \delta$ is also a left limit diagram, which shows that the functor "$F \mapsto LF(A)$"$= \varinjlim \circ \Psi$ preserves finite left limits.

The las fact we need is Theorem 1, p. 111 in CWM on "pointwise computability of limits in a functor category". According to that theorem, for $\delta':\mathbf{J} \to \widehat{\mathcal{C}}$ to know that δ' is a left limit diagram, it is enough to know that for each object A of $\mathcal{C}$, $\mathrm{ev}_A \circ \delta':\mathbf{J} \xrightarrow{\delta'} \widehat{\mathcal{C}} \xrightarrow{\mathrm{ev}_A} \mathbf{Set}$ is such. Let $\delta:\mathbf{J} \to \widehat{\mathcal{C}}$ be a finite left limit diagram. Let $\delta':\mathbf{J} \xrightarrow{\delta} \widehat{\mathcal{C}} \xrightarrow{L} \widehat{\mathcal{C}}$ be $\delta' = L \circ \delta$. Then $\mathrm{ev}_A \circ \delta' = \mathbf{J} \to \mathbf{Set} = \delta \circ$"$F \mapsto LF(A)$" is a left limit diagram, according to what we said above. Hence $\delta' = L \circ \delta$ is a left limit diagram, proving that L preserves finite left limits. $\square$

We can now summarize the work of this section as follows.

Define the functor $a:\widehat{\mathcal{C}} \to \widehat{\mathcal{C}}$ as $a = L \circ L:\widehat{\mathcal{C}} \to \widehat{\mathcal{C}} \to \widehat{\mathcal{C}}$. By 1.2.5 and 1.2.8, $a(F)$ is always a sheaf. Since $\widetilde{\mathcal{C}}$ is a full subcategory of $\widehat{\mathcal{C}}$, we hence regard a as a functor

$$a:\widehat{\mathcal{C}} \to \widetilde{\mathcal{C}}.$$

Define the natural transformation $\alpha:\mathrm{id}_{\widehat{\mathcal{C}}} \to i \circ a$ (with $i:\widetilde{\mathcal{C}} \to \widehat{\mathcal{C}}$ the inclusion) as follows: for a presheaf F, let

$$\alpha_F : F \xrightarrow{\ell_F} LF \xrightarrow{\ell_{LF}} LLF, \qquad \alpha_F = \ell_{LF} \circ \ell_F.$$

Clearly, α is a natural transformation as required (since ℓ is). By 1.2.4(ii), if S is a sheaf, ℓ_S is an isomorphism, hence LS is a sheaf and hence again, ℓ_{LS} is an isomorphism and finally, α_S is an isomorphism. Also, by 1.2.7, if in

$$\begin{tikzcd} F \arrow[r, "f"] \arrow[d, "\alpha_F"'] & S \\ a(F) \arrow[ur, shift left, "g_1"] \arrow[ur, shift right, "g_2"'] & \end{tikzcd}$$

we have $g_1 \circ \alpha_F = g_2 \circ \alpha_F = f$, where S is a sheaf, then $g_1 = g_2$. Finally, by 1.2.9, the functor $L \circ L:\widehat{\mathcal{C}} \to \widehat{\mathcal{C}}$ preserves finite left limits. We leave it to the reader to verify the easy

Lemma 1.2.10 *For any (finite of infinite) diagram of sheaves and morphisms of sheaves, the* left *limit of the diagram computed in $\widehat{\mathcal{C}}$ is actually a sheaf and hence it is the left limit computed in $\widetilde{\mathcal{C}}$.*

It follows that $a : \widehat{\mathcal{C}} \to \widetilde{\mathcal{C}}$ preserves finite left limits. Accordingly to what we said above, this concludes the proof of Theorem 1.2.1.

§3 Grothendieck Topoi

Definition 1.3.1 *A (Grothendieck) topos is a category that is equivalent to the category of sheaves, $\widetilde{\mathcal{C}}$, over a* small *site $\mathcal{C}$.*

In this section we derive some properties of topoi, culminating in a 'universal property' of $\widetilde{\mathcal{C}}$ for a given site $\mathcal{C}$ among all topoi (1.3.15 below). In the next section, we finally give an 'abstract' characterization of topoi (Giraud's theorem).

We will need here the following lemma whose proof we defer to Chapter 3, 3.4.11.

Lemma 1.3.2 *In a Grothendieck topos, every epimorphic family is a* stable effective *epimorphic family (c.f. Definition* 1.1.8*).*

Let $\mathcal{C}$ be any site. Recall the associated sheaf functor $a(= a_{\mathcal{C}}) : \widehat{\mathcal{C}} \to \widetilde{\mathcal{C}}$. The *representable sheaf* functor (for the lack of a better name) $\varepsilon(= \varepsilon_{\mathcal{C}}) : \mathcal{C} \to \widetilde{\mathcal{C}}$ is the composition $\mathcal{C} \xrightarrow{h_{(\cdot)}} \widehat{\mathcal{C}} \xrightarrow{a} \widetilde{\mathcal{C}}$; it takes every object $A \in \mathrm{Ob}(\mathcal{C})$ into the sheaf associated to h_A.

The category $\widetilde{\mathcal{C}}$ has finite left limits (c.f. 1.2.10). Hence, we can consider $\widetilde{\mathcal{C}}$ a site with its canonical topology. In what follows, whenever $\widetilde{\mathcal{C}}$ is regarded as a site, the canonical topology is the topology intended.

Proposition 1.3.3 *$\varepsilon : \mathcal{C} \to \widetilde{\mathcal{C}}$ is a continuous functor between sites; actually we have*

(i) *ε preserves finite left limits.*

(ii) *$\langle \varepsilon A_i \xrightarrow{\varepsilon f_i} \varepsilon A \rangle_{i \in I}$ is a covering in $\widetilde{\mathcal{C}}$ if,* and actually only if, *$\langle A_i \xrightarrow{f_i} A \rangle_{i \in I}$ is a covering in $\mathcal{C}$.*

Proof. (i) is a consequence of the facts that a preserves finite left limits and all inductive limits and that the Yoneda functor $h_{(\cdot)}$ preserves all projective limits (c.f. CWM).

Next we turn to the proof of part (ii). Using also 1.3.2, we see that in $\widetilde{\mathcal{C}}$, a family $\langle X_i \xrightarrow{g_i} X \rangle_{i \in I}$ is a covering iff it is an effective epimorphic family (stability being a consequence). Hence the condition

$$\langle \varepsilon A_i \xrightarrow{\varepsilon f_i} \varepsilon A \rangle_{i \in I} \quad \text{is a covering in } \widetilde{\mathcal{C}} \qquad (*)$$

is equivalent to saying that the diagram

$$\mathrm{Hom}(\varepsilon A, F) \longrightarrow \textstyle\prod_i \mathrm{Hom}(\varepsilon A_i, F) \rightrightarrows \textstyle\prod_{i,j} \mathrm{Hom}(\varepsilon A_i \times_{\varepsilon A} \varepsilon A_j, F)$$

in **Set**, for any $F \in \mathrm{Ob}(\widetilde{\mathcal{C}})$, with the natural arrows (c.f. the remark after 1.1.6) is exact. From the adjoint functors $a \dashv i$,

$$\begin{tikzcd} \widetilde{\mathcal{C}} \arrow[r, shift right, "i"'] & \widehat{\mathcal{C}}, \arrow[l, shift right, "a"'] \end{tikzcd}$$

using the unit $\alpha : \mathrm{id}_{\widehat{\mathcal{C}}} \to i \circ a$ of the adjunction, we have that the morphisms $G \xrightarrow{\alpha_G} aG$ for $G = A$, A_i and $A_i \times_A A_j$ induce the vertical arrows

$$\begin{tikzcd} \mathrm{Hom}(A, F) \arrow[r, "u"] \arrow[d, "f"'] & \prod_i \mathrm{Hom}(A_i, F) \arrow[r, shift left, "v_2"] \arrow[r, shift right, "v_1"'] \arrow[d, "g"'] & \prod_{i,j} \mathrm{Hom}(A_i \times_A A_j, F) \arrow[d, "h"'] \\ \mathrm{Hom}(\varepsilon A, F) \arrow[r, "u'"] & \prod_i \mathrm{Hom}(\varepsilon A_i, F) \arrow[r, shift left, "v_2'"] \arrow[r, shift right, "v_1'"'] & \prod_{i,j} \mathrm{Hom}(\varepsilon(A_i \times_A A_j), F). \end{tikzcd}$$

By adjointness, f, g and h are isomorphisms and in fact, they establish an isomorphism of the upper and lower halves of the diagram, i.e.

$$g \circ u = u' \circ f, \quad h \circ v_k = v'_k \circ g \quad (k = 1, 2).$$

Now, recall that we also have $\varepsilon(A_i \times_A A_j) \simeq \varepsilon A_i \times_{\varepsilon A} \varepsilon A_j$ (part (i) of the proposition). We can draw the conclusion that $(*)$ above holds if and only if *every sheaf F over the site $\mathcal{C}$ has the sheaf property with respect to the given family $\langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$* (c.f. 1.1.8(ii)).

The 'if' direction of (ii) now follows immediately. For the other direction we have to do some more work.

A morphism $X \xrightarrow{f} Y$, in any category, is an *effective epimorphism* if the family consisting of the single element f is an effective epimorphic family. This is equivalent to saying that for any morphism $X \xrightarrow{g} Z$ such that in

$$\begin{tikzcd}
 & X \arrow[dr, "f"] \arrow[drr, bend left, "g"] & & \\
X \times_Y X \arrow[ur, "p_1"] \arrow[dr, "p_2"'] & & Y \arrow[r, dashed, "k"'] & Z \\
 & X \arrow[ur, "f"'] \arrow[urr, bend right, "g"'] & &
\end{tikzcd}$$

we have $g \circ p_1 = g \circ p_2$, there is a unique $k : Y \to Z$ such that $g = k \circ f$. Next we formulate two simple lemmas whose proofs are left to the reader.

1.3.4 *An effective epimorphism which is a monomorphism is an isomorphism.*

1.3.5 *Given the site $\mathcal{C}$, the family $\alpha = \langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ of morphisms in $\mathcal{C}$ and R, the sieve $R \overset{i}{\hookrightarrow} A$ generated by α, if $\langle \varepsilon A_i \xrightarrow{\varepsilon f_i} \varepsilon A\rangle_{i\in I}$ is an effective epimorphic family in $\widetilde{\mathcal{C}}$, then $aR \xrightarrow{ai} \varepsilon A$ is an effective epimorphism, hence (by* 1.3.4*) ai is an isomorphism.*

Suppose $\phi = \langle A_i \xrightarrow{f_i} A\rangle_{i\in I}$ in $\mathcal{C}$ is such that $\langle \varepsilon A_i \xrightarrow{\varepsilon f_i} \varepsilon A\rangle_{i\in I}$ is a covering in $\widetilde{\mathcal{C}}$. With $R \overset{i}{\hookrightarrow} A$ the sieve generated by ϕ, we have the isomorphism $aR \xrightarrow[ai]{\simeq} \varepsilon A$.

Start with the following commutative diagram referring to the construction of the associated sheaf in Section 2:

$$\begin{tikzcd}
A \arrow[r, "\ell_A"] & LA \arrow[r, "\ell_{LA}"] & LLA \\
R \arrow[u, hook, "i"'] \arrow[r, "\ell_R"'] & LR \arrow[u, hook, "Li"'] \arrow[r, "\ell_{LR}"'] & LLR \arrow[u, hook, "ai"']
\end{tikzcd}$$

Apply 1.2.6(iii) to the morphism $(ai)^{-1} \circ \ell_{LA} \circ \ell_A : A \to L(LR)$ in place of v there and for the identical covering $A \xrightarrow{\mathrm{id}} A$ for R there. We obtain a covering $S_1 \overset{i_1}{\hookrightarrow} A \in J(A)$ and a morphism $S_1 \xrightarrow{u_1} LR$ such that in

$$\begin{tikzcd}
S_2 \times_{S_1} B \arrow[r, hook] \arrow[d] & B \arrow[d, "\beta"'] & & & \\
S_2 \arrow[r, hook] & S_1 \arrow[r, "i_1"'] \arrow[rr, bend left, "m"] \arrow[rr, bend right=60, "u_1"'] \arrow[rr, bend right=90, "n"'] & A \arrow[r, "\ell_A"'] & LA \arrow[r, "\ell_{LA}"] & LLA \\
 & & R \arrow[u, hook, "i"'] \arrow[r, "\ell_R"] & LR \arrow[u, hook, "Li"] \arrow[r, "\ell_{LR}"'] & LLR \arrow[u, hook, "LLi=ai"']
\end{tikzcd}$$

we have

$$LLi \circ \ell_{LR} \circ u_1 = \ell_{LA} \circ \ell_A \circ i_1.$$

Put $m = \ell_A \circ i_1$, $n = Li \circ u_1$. Let $S_2 \to S_1$ be the equalizer (in $\widehat{\mathcal{C}}$) of the arrows m and n. We claim that $\begin{tikzcd}[column sep=small] & S_1 \arrow[dr, hook, "i_1"] & \\ S_2 \arrow[ur, hook] \arrow[rr, "i_2"'] & & A \end{tikzcd}$ is a covering $\in J(A)$. To show this, let $B \xrightarrow{\beta} S_1$ be an arbitrary morphism with $B \in \mathrm{Ob}(\mathcal{C})$. Then $S_2 \times_{S_1} B \hookrightarrow B$ is the equalizer of $m \circ \beta$ and $n \circ \beta$. But we have $\ell_{LA} \circ m \circ \beta = \ell_{LA} \circ n \circ \beta$, hence by 1.2.6(iv) $S_2 \times_{S_1} B \hookrightarrow B \in J(B)$. Since $B \xrightarrow{\beta} S_1$ was arbitrary, it follows from 1.1.2(iii) ("local character") that $S_2 \xhookrightarrow{i_2} A \in J(A)$ indeed.

Hence, we have $S_2 \xhookrightarrow{i_2} A \in J(A)$ and a morphism $u_2 : S_2 \to LR$ ($u_2 = S_2 \hookrightarrow S_1 \xrightarrow{u_1} LR$) such that the following commutes

$$\begin{tikzcd}
S_2 \arrow[rr, hook, "i_2"] \arrow[dd, "u_2"'] & & A \arrow[dd] \\
 & R \arrow[ur, hook, "i"'] \arrow[dl, "\ell_R"] & \\
LR \arrow[rr, "Li"'] & & LA.
\end{tikzcd}$$

With an eye on applying 1.1.2(iii) to show that $S_2 \times_A R \in J(A)$, let $B \xrightarrow{\beta} S_2$ be an arbitrary morphism with $B \in \mathrm{Ob}(\mathcal{C})$. Apply again 1.2.6(iii) to obtain $Q_1 \hookrightarrow B \in J(B)$ and $v_1 : Q_1 \to R$

$$\begin{tikzcd}
S_2 \arrow[rrrr, hook, "i_2"] \arrow[dddd, "u_2"'] & & & & A \arrow[dddd, "\ell_A"] \\
 & B \arrow[ul, "\beta"'] & Q_2 \arrow[d, hook] & & \\
 & & Q_1 \arrow[ul, hook, "j_1"] \arrow[uurr, bend left, "f"] \arrow[uurr, "g"'] \arrow[dr, "v_1"'] & & \\
 & & & R \arrow[uuur, hook, "i"'] \arrow[dlll, "\ell_R"'] & \\
LR \arrow[rrrr, "Li"'] & & & & LA
\end{tikzcd}$$

such that $u_2 \circ \beta \circ j_1 = \ell_R \circ v_1$. Put $f = i_2 \circ \beta \circ j_1$, $g = i \circ v_1$. Let $Q_2 \hookrightarrow Q_1$ be the equalizer of f and g. Just as we showed above that $S_2 \xhookrightarrow{i_2} A \in J(A)$ we can show that $Q_2 \xhookrightarrow{j_2} B = Q_2 \hookrightarrow Q_1 \xhookrightarrow{j_1} B$ belongs to $J(B)$. Let now $S_3 \hookrightarrow A$ be the 'intersection' $S_2 \times_A R \hookrightarrow A$. By the definition of Q_2, we have $Q_2 \leq S_3 \times_{S_2} B$, hence $S_3 \times_{S_2} B \in J(B)$. Since $B \xrightarrow{\beta} S_2$ was arbitrary, this shows that $S_3 \in J(A)$ and a fortiori $R \in J(A)$. This is equivalent to saying that the family $\langle A_i \xrightarrow{f_i} A \rangle_{i \in I}$ is a covering. $\square$

Remark *According to what we said above, the proof shows that if every sheaf over $\mathcal{C}$ has the sheaf property with respect to $\phi = \langle A_i \xrightarrow{f_i} A \rangle_{i \in I}$, then ϕ is a covering. This is a 'completeness' property of the notion of sheaf with respect to Grothendieck topologies.*

Definition 1.3.6 *For an arbitrary category $\mathcal{R}$, a set G of objects of $\mathcal{R}$ is said to be a* set of generators *for $\mathcal{R}$ if for every $A \in \mathrm{Ob}(\mathcal{R})$ the family of all morphisms with domains in G and codomain A,*

$$\langle B \xrightarrow{g} A \rangle_{B \in G}$$

is an epimorphic family.

Proposition 1.3.7 *For $\mathcal{C}$ a* small *site, the category $\widetilde{\mathcal{C}}$ of sheaves has a set of generators,* *namely the set of objects of the form εA, for $A \in \mathrm{Ob}(\mathcal{C})$ and $\varepsilon : \mathcal{C} \to \widetilde{\mathcal{C}}$ the representable sheaf functor.*

Proof. Suppose $F \underset{g}{\overset{f}{\rightrightarrows}} G$ are two morphisms in $\widetilde{\mathcal{C}}$ such that for any $A \in \mathrm{Ob}(\mathcal{C})$ and any $h : \varepsilon A = aA \to F$, we have $f \circ h = g \circ h$. To show that $f = g$, let $A \xrightarrow{k} F$ be an arbitrary morphism in $\widehat{\mathcal{C}}$ (!) with $A \in \mathrm{Ob}(\mathcal{C})$. By the 'universal property' of aA there is a unique h_k such that

$$\begin{tikzcd}
aA \arrow[r, "h_k"] & F \\
& A \arrow[u, "k"'] \arrow[ul, "\alpha_A"]
\end{tikzcd}$$

commutes. It follows that $f \circ k = g \circ k$. Since $k : A \to F$ is arbitrary, $f = g$. $\square$

We will need the following lemma in Chapter 6.

Lemma 1.3.8 (i) *Given $\mathcal{C} \xrightarrow{\varepsilon} \widetilde{\mathcal{C}}$ as above and a morphism $\varepsilon A \xrightarrow{f} \varepsilon B$ in $\widetilde{\mathcal{C}}$, there is a covering $\langle A_i \xrightarrow{f_i} A\rangle_{i \in i} \in \mathrm{Cov}(A)$ and morphisms $A_i \xrightarrow{g_i} B$ in $\mathcal{C}$ such that $f \circ \varepsilon f_i = \varepsilon g_i$ for all $i \in I$.*

(ii) *Given a monomorphism $X \xrightarrow{\xi} \varepsilon A$ in $\widetilde{\mathcal{C}}$, $A \in \mathrm{Ob}(\mathcal{C})$, there is a covering family $\langle \varepsilon A_i \xrightarrow{g_i} X\rangle_{i \in I}$ in $\widetilde{\mathcal{C}}$ such that the compositions $\xi \circ g_i$ are of the form $\xi \circ g_i = \varepsilon(f_i)$ for some*

$$A_i \xrightarrow{f_i} A \text{ in } \mathcal{C}, \text{ for every } i \in I.$$

Proof of (i):

$$\begin{tikzcd}
LLA_{ij} \arrow[rr, "LLh_{ij}"'] \arrow[rrr, bend left, "LLg_{ij}"] & & \varepsilon A = LLA \arrow[r, "f"'] & LLB = \varepsilon B \\
LA_{ij} \arrow[u, "\ell_{LA_{ij}}"'] & & LA \arrow[u, "\ell_{LA}"'] & LB \arrow[u, "\ell_{LB}"'] \\
A_{ij} \arrow[u, "\ell_{A_{ij}}"'] \arrow[r, "f_{ij}"] \arrow[rr, bend right, "h_{ij}=f^k"] \arrow[rrr, bend right=40, "g_{ij}=g^k"'] & A_i \arrow[r, "f_i"] \arrow[rru, bend right, "g_i"'] & A \arrow[u, "\ell_A"'] & B \arrow[u, "\ell_B"']
\end{tikzcd}$$

By 1.2.6(iii), for $f \circ \ell_{LA} \circ \ell_A$ as u there, there is a covering $\langle A_i \xrightarrow{f_i} A\rangle_{i \in I} \in \mathrm{Cov}(A)$ and morphisms $A_i \xrightarrow{g_i} LB$ such that $f \circ \ell_{LA} \circ \ell_A \circ f_i = \ell_{LB} \circ g_i$ for every $i \in I$. With given $i \in I$, similarly, there is a covering $\langle A_{ij} \xrightarrow{f_{ij}} A_i\rangle_{j \in J_i}$ and morphisms $A_{ij} \xrightarrow{g_{ij}} B$, ($j \in J_i$) such that $g_i \circ f_{ij} = \ell_B \circ g_{ij}$. Denote the covering $\langle A_{ij} \xrightarrow{f_i \circ f_{ij}} A\rangle_{j \in J_i, i \in I}$ by $\langle A^k \xrightarrow{f^k} A\rangle_{k \in K}$ and accordingly, g_{ij} by g^k. We have

$$f \circ \ell_{LA} \circ \ell_A \circ f^k = \ell_{LB} \circ \ell_B \circ g^k,$$

hence

$$f \circ LLf^k \circ \ell_{LA^k} \circ \ell_{A^k} = \ell_{LB} \circ \ell_B \circ g^k.$$

Applying the functor LL to g^k, we also have

$$LLg^k \circ \ell_{LA^k} \circ \ell_{A^k} = \ell_{LB} \circ \ell_B \circ g^k,$$

hence

$$f \circ LLf^k \circ \ell_{LA^k} \circ \ell_{A^k} = LLg^k \circ \ell_{LA^k} \circ \ell_{A^k}.$$

By 1.2.7, it follows that $f \circ \varepsilon f^k = \varepsilon g^k$. In other words, the covering $\langle A^k \xrightarrow{f^k} A \rangle_{k\in K}$ and the morphism $A^k \xrightarrow{g^k} B$ satisfy the requirements.

Proof of (ii): By 1.3.7, find an epimorphic family of the form $\langle \varepsilon B_i \xrightarrow{f_i} X\rangle_{i\in I}$. By 1.3.2 (as yet unproved), the same family is a covering in the canonical topology of $\widetilde{\mathcal{C}}$. Now, apply part (i) to each morphism $\varepsilon B_i \xrightarrow{\xi\circ f_i} \varepsilon A$ separately. By 1.3.3(ii), the resulting coverings of the B_i become (through ε) coverings of the εB_i. The thus resulting coverings 'add up' to form a covering of X. □

Given topoi $\mathcal{E}_1$ and $\mathcal{E}_2$, we regard them as sites with the respective canonical topologies. A continuous functor $F : \mathcal{E}_1 \to \mathcal{E}_2$ is also called an $\mathcal{E}_2$-model of $\mathcal{E}_1$. We will show that, for two sites $\mathcal{C}$ and $\mathcal{D}$, any $\mathcal{D}$-model of $\mathcal{C}$ can be 'lifted' to a $\widetilde{\mathcal{D}}$-model of $\widetilde{\mathcal{C}}$ in a natural way. However, instead of an $\mathcal{E}_2$-model of $\mathcal{E}_1$, we first will talk about an apparently different notion.

Definition 1.3.9 *Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be two (Grothendieck) topoi. A* morphism U from $\mathcal{E}_1$ *to $\mathcal{E}_2$,*

$$U : \mathcal{E}_1 \to \mathcal{E}_2$$

is a triple $U = (u_*, u^*, \phi)$ where u_*, u^* are functors

$$\begin{tikzcd} \mathcal{E}_1 \arrow[r, shift right, "u_*"'] & \mathcal{E}_2 \arrow[l, shift right, "u^*"'] \end{tikzcd}$$

such that u^* preserves finite left limits, moreover u^* is a left adjoint to u_*, and in fact (u_*, u^*, ϕ) is an adjunction form $\mathcal{E}_2$ to $\mathcal{E}_1$ in the sense of CWM.

Remark *This means that ϕ is a function which assigns to each pair of objects $A \in \mathrm{Ob}(\mathcal{E}_2)$, $B \in \mathrm{Ob}(\mathcal{E}_1)$ a bijection $\phi_{B,A} : \mathrm{Hom}_{\mathcal{E}_1}(u^*A, B) \simeq \mathrm{Hom}_{\mathcal{E}_2}(A, u_*B)$ which is natural in A and B.*

Theorem 1.3.10 *Let $\mathcal{C}$ be a small site, $\mathcal{D}$ a locally small site and let $u : \mathcal{C} \to \mathcal{D}$ be a $\mathcal{D}$-model of $\mathcal{C}$. Then u can be lifted to a geometric morphism*

$$U : \widetilde{\mathcal{D}} \to \widetilde{\mathcal{C}}$$

in the following sense: there is $U = (u_*, u^*, \phi)$ such that U is a geometric morphism $\widetilde{\mathcal{D}} \to \widetilde{\mathcal{C}}$ and the diagram

$$\begin{tikzcd} \mathcal{C} \arrow[r, "u"] \arrow[d, "\varepsilon_{\mathcal{C}}"'] & \mathcal{D} \arrow[d, "\varepsilon_{\mathcal{D}}"] \\ \widetilde{\mathcal{C}} \arrow[r, "u^*"'] & \widetilde{\mathcal{D}} \end{tikzcd}$$

commutes.

Proof. We will use the concept of Kan-extension, c.f. CWM, Chapter 10. Given the categories $\mathcal{C}$, $\mathcal{D}$ (the underlying categories of the sites) and the functor $u : \mathcal{C} \to \mathcal{D}$, we have that $\widehat{\mathcal{C}} = \mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$, $\widehat{\mathcal{D}} = \mathbf{Set}^{\mathcal{D}^{\mathrm{op}}}$ are functor categories of the kind treated in loc. cit., with $\mathcal{A} = \mathbf{Set}$. We denote by $u_\bullet : \widehat{\mathcal{D}} \to \widehat{\mathcal{C}}$ what CWM denotes by $\mathbf{Set}^u$. This is the functor such that for $F \in \mathrm{Ob}(\widehat{\mathcal{D}})$, $u_\bullet F = F \circ u$ and for $F \xrightarrow{\nu} G$ in $\widehat{\mathcal{D}}$, $(u_\bullet(\nu))_C = \nu_{uC}$ for

$C \in \mathrm{Ob}(\mathcal{C})$. According to loc. cit. (c.f. dual of Corollary 2, p. 235) $u_\bullet$ has a left adjoint, denoted by $u^\bullet$:

$$u^\bullet : \widehat{\mathcal{C}} \to \widehat{\mathcal{D}}$$

$$u^\bullet \dashv u_\bullet$$

$$\widehat{\mathcal{C}} \underset{u_\bullet}{\overset{u^\bullet}{\rightleftarrows}} \widehat{\mathcal{D}}$$

$u^\bullet F$ for $F \in \mathrm{Ob}(\widehat{\mathcal{C}})$ is called in CWM the *(left) Kan-extension of* F. We will also need to know the way $u^\bullet$ is actually constructed.

Let D be an object in $\mathcal{D}$. Define the comma category $D{\downarrow}u$ as follows. An object of $D{\downarrow}u$ is a pair (f, C) where $C \in \mathrm{Ob}(\mathcal{C})$ and $f : D \to uC$ is a morphism in $\mathcal{D}$. We also write $D \xrightarrow{f} uC$ for an object of $D{\downarrow}u$. A morphism between (f, C) and (f', C') is a morphism $g : C \to C'$ such that

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

commutes. Composition in $D{\downarrow}u$ is defined in the obvious way. Notice that since $\mathcal{C}$ is small and $\mathcal{D}$ is locally small, $D{\downarrow}u$ is a small category for each $D \in \mathrm{Ob}(\mathcal{D})$.

Let now F be a presheaf $\in \mathrm{Ob}(\widehat{\mathcal{C}})$. To define $G = u^\bullet F$, we put

$$G(D) = \varinjlim_{(f:D\to uC)\in(D\downarrow u)^{\mathrm{op}}} F(C).$$

In other words $G(D)$ is the right limit of the composite functor $(D{\downarrow}u)^{\mathrm{op}} \xrightarrow{p} \mathcal{C}^{\mathrm{op}} \xrightarrow{F} \mathbf{Set}$ where $p(D \xrightarrow{f} uC) = C$ and $p(g) = g$ for g a morphism in $D{\downarrow}u$. Also, for a morphism $D \xrightarrow{\delta} D'$ in $\mathcal{D}$, we define $G(\delta)$ by the universal property of the limit defining $G(D')$; we omit the obvious description.

Finally, to define the effect of $u^\bullet$ on a natural transformation $\nu : F \to F'$, we have to define morphisms

$$(u^\bullet \nu)_D : (u^\bullet F)(D) \to (u^\bullet F')(D)$$

for $D \in \mathrm{Ob}(\mathcal{D})$. This again is a canonical map between limits, based on the maps

$$F(C) \xrightarrow{\nu_C} F'(C).$$

For a representable presheaf $h^{\mathcal{C}}_{C_0} \in \mathrm{Ob}(\widehat{\mathcal{C}})$, we invite the reader to check that we have a canonical isomorphism

$$\varinjlim_{(f:D\to uC)\in(D\downarrow u)^{\mathrm{op}}} h^{\mathcal{C}}_{C_0}(C) \simeq h^{\mathcal{D}}_{uC_0}(D).$$

Actually, since the exact choice of the limit objects is irrelevant, we can define $u^\bullet(h^{\mathcal{C}}_{C_0})$ such that we have

$$(u^\bullet(h^{\mathcal{C}}_{C_0}))(D) = h^{\mathcal{D}}_{uC_0}(D).$$

This, and the appropriate choice for the morphisms $(u^\bullet(h^{\mathcal{C}}_{C_0}))(\delta)$ for $\delta : D \to D'$, will make sure that the diagram

$$\begin{tikzcd} \mathcal{C} \arrow[r, "u"] \arrow[d, "h^{\mathcal{C}}"'] & \mathcal{D} \arrow[d, "h^{\mathcal{D}}"] \\ \widehat{\mathcal{C}} \arrow[r, "u^\bullet"'] & \widehat{\mathcal{D}} \end{tikzcd} \qquad (1)$$

commutes (with $h^{\mathcal{C}}$, $h^{\mathcal{D}}$ the Yoneda functors).

Next, we have to verify that $u^\bullet$ preserves finite left limits. The proof of this fact is very similar to the proof of 1.2.9 and it is based on the fact that the categories $(D{\downarrow}u)^{\mathrm{op}}$, for $D \in \mathrm{Ob}(\mathcal{D})$, are filtered (c.f. CWM). This latter fact will be seen to be a consequence of the fact that u preserves finite left limits. E.g., we want to see that for objects $A_1 : D \xrightarrow{f_1} uC_1$, $A_2 : D \xrightarrow{f_2} uC_2$ in $D{\downarrow}u$, there are morphisms $A_3 \xrightarrow{g_1} A_1$, $A_3 \xrightarrow{g_2} A_2$ with some $A_3 : D \xrightarrow{f_3} uC_3$. To this end, define $C_3 = C_1 \times C_2$, with projections π_1 and π_2. Then uC_3 is $uC_1 \times uC_2$ with projections $u\pi_1$ and $u\pi_2$. Hence, there is $f_3 : D \to uC_3$ such that the following is commutative:

$$\begin{tikzcd} & uC_3 \arrow[dl, "u\pi_1"'] \arrow[dr, "u\pi_2"] & \\ uC_1 & & uC_2 \\ & D \arrow[ul, "f_1"] \arrow[ur, "f_2"'] \arrow[uu, "f_3"] & \end{tikzcd}$$

Put $g_1 = \pi_1$, $g_2 = \pi_2$. These choices will obviously work.

The rest of the proof that $u^\bullet$ preserves finite left limits is left to the reader.

Next, we are going to 'lift' $u_\bullet$, $u^\bullet$ to u_*, u^* as follows. Consider the following diagram:

$$\begin{tikzcd} \mathcal{C} \arrow[r, "h^{\mathcal{C}}"] \arrow[d, "u"'] & \widehat{\mathcal{C}} \arrow[r, shift left, "a_{\mathcal{C}}"] \arrow[d, shift right, "u^\bullet"'] & \widetilde{\mathcal{C}} \arrow[l, shift left, "i_{\mathcal{C}}"] \arrow[d, shift right, "u^*"'] \\ \mathcal{D} \arrow[r, "h^{\mathcal{D}}"'] & \widehat{\mathcal{D}} \arrow[r, shift left, "a_{\mathcal{D}}"] \arrow[u, shift right, "u_\bullet"'] & \widetilde{\mathcal{D}} \arrow[l, shift left, "i_{\mathcal{D}}"] \arrow[u, shift right, "u_*"'] \end{tikzcd}$$

Here, $h^{\mathcal{C}}$, $h^{\mathcal{D}}$ are the Yoneda functors, $i_{\mathcal{C}}$, $i_{\mathcal{D}}$ are the inclusions and $a_{\mathcal{C}}$, $a_{\mathcal{D}}$ are the associated sheaf functors. Define $u_* \underset{\mathrm{df}}{=} a_{\mathcal{C}} \circ u_\bullet \circ i_{\mathcal{D}}$, $u^* \underset{\mathrm{df}}{=} a_{\mathcal{D}} \circ u^\bullet \circ i_{\mathcal{C}}$.

Since each of $i_{\mathcal{C}}$, $u^\bullet$, $a_{\mathcal{D}}$ preserve finite left limits (c.f. 1.2.10 for $i_{\mathcal{D}}$) the same is true of u^*.

Lemma 1.3.11 *For a sheaf $F \in \mathrm{Ob}(\widetilde{\mathcal{D}}) \subset \mathrm{Ob}(\widehat{\mathcal{C}})$, $u_\bullet F$ is a sheaf $\in \mathrm{Ob}(\widetilde{\mathcal{C}})$.*

This is an immediate consequence of the definition of $u_\bullet$. The diagram

$$(u_\bullet F)(A) \longrightarrow \prod_{i\in I}(u_\bullet F)(A_i) \rightrightarrows \prod_{i,j\in I}(u_\bullet F)(A_i \times_A A_j)$$

is identical to

$$F(uA) \longrightarrow \prod_{i\in I} F(uA_i) \rightrightarrows \prod_{i,j\in I} F(u(A_i \times_A A_j))$$

for a covering $\langle A_i \longrightarrow A\rangle_{i\in I}$ of A in $\mathcal{C}$. Since u preserves finite left limits, $u(A_i \times_A A_j) \simeq uA_i \times_{uA} uA_j$, and the required exactness is a consequence of F being a sheaf.

Lemma 1.3.11 has the effect that $i_{\mathcal{C}} u_* F \simeq u_\bullet F = u_\bullet i_{\mathcal{D}} F$ for $F \in \mathrm{Ob}(\widetilde{\mathcal{D}})$. This permits us to prove that u^* is left adjoint to u_*, as follows:

$$\begin{aligned} \mathrm{Hom}_{\widetilde{\mathcal{D}}}(u^*G, F) &\simeq \mathrm{Hom}_{\widetilde{\mathcal{D}}}(a_{\mathcal{D}} u^\bullet i_{\mathcal{C}} G, F) \\ &\simeq \mathrm{Hom}_{\widehat{\mathcal{D}}}(u^\bullet i_{\mathcal{C}} G, i_{\mathcal{D}} F) \\ \text{by } a_{\mathcal{D}} \dashv i_{\mathcal{D}};\ &\simeq \mathrm{Hom}_{\widehat{\mathcal{C}}}(i_{\mathcal{C}} G, u_\bullet i_{\mathcal{D}} F) \\ \text{by } u^\bullet \dashv u_\bullet;\ &\simeq \mathrm{Hom}_{\widehat{\mathcal{C}}}(i_{\mathcal{C}} G, i_{\mathcal{C}} u_* F) \\ \text{by the above remark};\ &\simeq \mathrm{Hom}_{\widetilde{\mathcal{C}}}(G, u_* F) \end{aligned}$$

by the fact that $i_{\mathcal{C}}$ is full and faithful. The reader can check that the required naturalities are present to ensure that indeed,

$$u^* \dashv u_* \quad \text{as claimed.}$$

In fact, the required adjunction ϕ can be read off the above sequence of isomorphisms.

Furthermore, we have the following isomorphisms, for $K \in \mathrm{Ob}(\widehat{\mathcal{C}})$, $F \in \mathrm{Ob}(\widetilde{\mathcal{D}})$:

$$\begin{aligned}
\mathrm{Hom}_{\widetilde{\mathcal{D}}}(u^* a_{\mathcal{C}} K, F) &\simeq \mathrm{Hom}_{\widetilde{\mathcal{C}}}(a_{\mathcal{C}} K, u_* F) \\
\text{by } u^* \dashv u_*; \quad &\simeq \mathrm{Hom}_{\widehat{\mathcal{C}}}(K, i_{\mathcal{C}} u_* F) \\
\text{by } a_{\mathcal{C}} \dashv i_{\mathcal{C}}; \quad &\simeq \mathrm{Hom}_{\widehat{\mathcal{C}}}(K, u_\bullet i_{\mathcal{D}} F) \\
\text{by } i_{\mathcal{C}} u_* \simeq u_\bullet i_{\mathcal{D}}; \quad &\simeq \mathrm{Hom}_{\widehat{\mathcal{D}}}(u^\bullet K, i_{\mathcal{D}} F) \\
\text{by } u^\bullet \dashv u_\bullet; \quad &\simeq \mathrm{Hom}_{\widetilde{\mathcal{D}}}(a_{\mathcal{D}} u^\bullet K, F) \\
\text{by } a_{\mathcal{D}} \dashv i_{\mathcal{D}}. \quad &
\end{aligned}$$

This implies that for any $K \in \mathrm{Ob}(\widehat{\mathcal{C}})$,

$$u^* a_{\mathcal{C}} K \simeq a_{\mathcal{D}} u^\bullet K.$$

In fact, since the above isomorphisms are natural in K and F, we have the isomorphism

$$u^* \circ a_{\mathcal{C}} \simeq a_{\mathcal{D}} \circ u^\bullet$$

of functors. Combining this with the commutative diagram (1) above, we obtain that

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "u"] \arrow[d, "\varepsilon_{\mathcal{C}}"'] & \mathcal{D} \arrow[d, "\varepsilon_{\mathcal{D}}"] \\
\widetilde{\mathcal{C}} \arrow[r, "u^*"'] & \widetilde{\mathcal{D}}
\end{tikzcd}$$

commutes up to isomorphism:

$$u^* \circ \varepsilon_{\mathcal{C}} \simeq \varepsilon_{\mathcal{D}} \circ u.$$

Having constructed u_*, u^* as above, now it is easy to modify u^* to a functor isomorphic to u^* so that the last diagram commutes literally and still $u^* \dashv u_*$, and u^* is left exact. $\square$

Theorem 1.3.12 *Given $\mathcal{C} \xrightarrow{u} \mathcal{D}$, a $\mathcal{D}$-model of $\mathcal{C}$, consider the following properties of functors $F : \widetilde{\mathcal{C}} \to \widetilde{\mathcal{D}}$:*

(a)

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "u"] \arrow[d, "\varepsilon_{\mathcal{C}}"'] & \mathcal{D} \arrow[d, "\varepsilon_{\mathcal{D}}"] \\
\widetilde{\mathcal{C}} \arrow[r, "F"'] & \widetilde{\mathcal{D}}
\end{tikzcd}$$

commutes up to isomorphism.

(b) *F preserves all (small) inductive limits in $\widetilde{\mathcal{C}}$ (that exist in $\widetilde{\mathcal{C}}$).*

(c) *F is a $\widetilde{\mathcal{D}}$-model of $\widetilde{\mathcal{C}}$, i.e., it preserves finite left limits and epimorphic families.*

Then u^* as determined in 1.3.10 *has properties* (a), (b), (c). *Also, each of the pairs*

(a) *and* (b)

(a) *and* (c)

determines F uniquely up to isomorphism.

Remark *We will state and prove in the next section that $\widetilde{\mathcal{C}}$ has* all *small inductive limits.*

Proof. We will use the following two Lemmas whose proofs we defer to Chapter 3 and to the next section.

Lemma 1.3.13 *For a functor $F : \widetilde{\mathcal{C}} \to \widetilde{\mathcal{D}}$ between topoi, if F preserves inductive limits and finite left limits, then F preserves epimorphic families (F is continuous).*

Lemma 1.3.14 *Considering $\widetilde{\mathcal{C}}$ a site with the canonical topology and forming $\widetilde{\widetilde{\mathcal{C}}}$, the category of sheaves over the site $\widetilde{\mathcal{C}}$, the canonical functor*

$$\varepsilon_{\widetilde{\mathcal{C}}} : \widetilde{\mathcal{C}} \to \widetilde{\widetilde{\mathcal{C}}}$$

is an equivalence.

Remark *It is not hard to show* 1.3.14 *directly. Also, by* 1.3.2 *and* 1.3.7*, $\mathcal{E}$ is a locally small site.*

Since u^* from 1.3.10 has a right adjoint u_*, u^* preserves inductive limits (c.f. CWM). Also, by 1.3.13, u^* has all the properties (a), (b), (c). Next we show that if F has (a) and (c) then it has (b) as well.

Use 1.3.10 to $u = F : \widetilde{\mathcal{C}} \to \widetilde{\mathcal{D}}$. We obtain a commutative diagram

$$\begin{tikzcd}
\widetilde{\mathcal{C}} \arrow[r, "F"] \arrow[d, "\varepsilon_{\widetilde{\mathcal{C}}}"'] & \widetilde{\mathcal{D}} \arrow[d, "\varepsilon_{\widetilde{\mathcal{D}}}"] \\
\widetilde{\widetilde{\mathcal{C}}} \arrow[r, "F^*"'] & \widetilde{\widetilde{\mathcal{D}}}
\end{tikzcd}$$

with a right adjoint F_* to F^*. Since by 1.3.14 $\varepsilon_{\widetilde{\mathcal{C}}}$ and $\varepsilon_{\widetilde{\mathcal{D}}}$ are equivalences, $F_* : \widetilde{\widetilde{\mathcal{D}}} \to \widetilde{\widetilde{\mathcal{C}}}$ will be transferred to some $G : \widetilde{\mathcal{D}} \to \widetilde{\mathcal{C}}$ such that $F \dashv G$. It follows that F preserves inductive limits as claimed.

Finally, it remains to show that properties (a) and (b) determine F up to isomorphisms. In the diagram

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "u"] \arrow[d, "h^{\mathcal{C}}"'] & \mathcal{D} \arrow[d, "h^{\mathcal{D}}"] \\
\widehat{\mathcal{C}} \arrow[r, "u^\bullet"'] \arrow[d, "a_{\mathcal{C}}"'] & \widehat{\mathcal{D}} \arrow[d, "a_{\mathcal{D}}"] \\
\widetilde{\mathcal{D}} \arrow[r, "F"'] & \widetilde{\mathcal{D}}
\end{tikzcd}$$

the upper and outer rectangles commute. We leave it to the reader to verify the existence of the following 'functorial' isomorphism, for any given $K \in \mathrm{Ob}(\widehat{\mathcal{C}})$ (and actually to make precise sense out of the phrase 'functorial'):

$$\varinjlim_{h^{\mathcal{C}}(X) \to K \in \mathrm{Ob}(\mathcal{C}/K)} h^{\mathcal{C}}(X) \xrightarrow{\quad \simeq \quad} K$$

(**Remark** *The category $\mathcal{C}/K$ has objects: morphisms of the form $f : h^{\mathcal{C}}(X) \to K$; a morphism between $h^{\mathcal{C}}(X) \to K$ and $h^{\mathcal{C}}(Y) \to K$ is a morphism $X \to Y$ such that*
$\begin{tikzcd}[cramped] h^{\mathcal{C}}(X) \arrow[r] \arrow[d] & K \\ h^{\mathcal{C}}(Y) \arrow[ur] & \end{tikzcd}$ *commutes. The inductive limit is take in the category $\widehat{\mathcal{C}}$.*)

Since $u^\bullet$, F, $a_{\mathcal{C}}$, $a_{\mathcal{D}}$ commute with inductive limits, we have

$$\varinjlim_{h^{\mathcal{C}}(X)\to K} F a_{\mathcal{C}} h^{\mathcal{C}}(X) \xrightarrow{\simeq} F a_{\mathcal{C}} K$$

$$\varinjlim_{h^{\mathcal{C}}(X)\to K} a_{\mathcal{D}} u^\bullet h^{\mathcal{C}}(X) \xrightarrow{\simeq} a_{\mathcal{D}} u^\bullet K$$

We have that $u^\bullet \circ h^{\mathcal{C}} = h^{\mathcal{D}} \circ u$, and, by hypothesis, $a_{\mathcal{D}} \circ h^{\mathcal{D}} \circ u \simeq F \circ a_{\mathcal{C}} \circ h^{\mathcal{C}}$. Hence it follows that

$$F a_{\mathcal{C}} K \simeq a_{\mathcal{D}} u^\bullet K.$$

Actually,

$$F \circ a_{\mathcal{C}} \simeq a_{\mathcal{D}} \circ u^\bullet$$

(where the meaning of the phrase "functorial" above plays a role). Since

$$a_{\mathcal{C}} \circ i_{\mathcal{C}} \simeq \in_{\mathcal{C}}$$

it follows that $F \simeq a_{\mathcal{D}} \circ u^\bullet \circ i_{\mathcal{C}} = u^*$. □

Next we formulate a variant of the above results.

Corollary 1.3.15 *Given a small site $\mathcal{C}$, a Grothendieck topos $\mathcal{E}$, and a continuous $u : \mathcal{C} \to \mathcal{E}$ (an $\mathcal{E}$-model of $\mathcal{C}$), there is $\widetilde{u} : \widetilde{\mathcal{C}} \to \mathcal{E}$ an $\mathcal{E}$-model of $\widetilde{\mathcal{C}}$, unique up to isomorphism, such that the diagram*

$$\begin{tikzcd} \mathcal{C} \arrow[r, "u"] \arrow[d, "\varepsilon_{\mathcal{C}}"'] & \mathcal{E} \\ \widetilde{\mathcal{C}} \arrow[ur, "\widetilde{u}"'] & \end{tikzcd}$$

commutes.

Proof. $\mathcal{E}$ is $\widetilde{\mathcal{D}}$ for a (small) site $\mathcal{D}$; also $\varepsilon_{\widetilde{\mathcal{D}}} : \widetilde{\mathcal{D}} \to \widetilde{\widetilde{\mathcal{D}}}$ is an equivalence (1.3.14). Apply 1.3.10 and 1.3.12 to obtain $u^* : \widetilde{\mathcal{C}} \to \widetilde{\mathcal{E}}$, a $\widetilde{\mathcal{E}}$-model of $\widetilde{\mathcal{C}}$, unique up to isomorphism, such that

$$\begin{tikzcd} \mathcal{C} \arrow[r, "u"] \arrow[d, "\varepsilon_{\mathcal{C}}"'] & \mathcal{E} \arrow[d, "\simeq"] \\ \widetilde{\mathcal{C}} \arrow[r, "u^*"'] & \widetilde{\mathcal{E}} \end{tikzcd}$$

commutes. Now the assertion is clear. □

Finally, we look at a special case of a $\mathcal{D}$-model u of $\mathcal{C}$ in case the pair of adjoint functors $\widetilde{\mathcal{C}} \underset{u^*}{\overset{u^*}{\rightleftarrows}} \widetilde{\mathcal{D}}$ actually give an equivalence: $u^* \circ u_* \simeq \mathrm{id}_{\widetilde{\mathcal{D}}}$, $u_* \circ u^* \simeq \mathrm{id}_{\widetilde{\mathcal{C}}}$.

Theorem 1.3.16 *Let $\mathcal{D}$ be a site. Let $\mathcal{C}$ be a full subcategory of $\mathcal{D}$ such that for any finite diagram with objects all in $\mathcal{C}$, there is a left limit in the sense of $\mathcal{D}$ in which the limit object belongs to $\mathcal{C}$. (Briefly, $\mathcal{C}$ is closed under finite left limits in $\mathcal{D}$.) Make $\mathcal{C}$ into a site by taking those families in $\mathcal{C}$ to be covering which are covering in the sense of $\mathcal{D}$. By the above, the inclusion functor $u : \mathcal{C} \to \mathcal{D}$ is continuous.*

Assume furthermore (the main hypothesis) that every object D in $\mathcal{D}$ has a covering $\langle C_i \to D \rangle_{i \in I}$ with objects $C_i \in \mathrm{Ob}(\mathcal{C})$.

In this case the functor $u_* : \widetilde{\mathcal{D}} \to \widetilde{\mathcal{C}}$

$$u_* : F \mapsto F \circ u$$

is an equivalence.

Remark *The statement of the theorem is equivalent to saying that, with u_*, u^* of* 1.3.10, *we have the isomorphisms before the statement of the Theorem. This is an obvious consequence of the uniqueness up to isomorphism of the left adjoint of the functor u_* (c.f.* CWM*). In* SGA4*, the proof of* 1.3.15 *("Lemme de comparaison",* vol. 1, p. 288*) is given by directly checking these isomorphisms. However, under the special conditions of the theorem, there is a simpler direct description of the quasi-inverse of u_*, i.e. u^*, than the one resulting from the general theory. We found that the proof of* 1.3.16 *using this direct approach is simpler than the one of the kind given in* SGA4.

PROOF. In order to describe a quasi-inverse u^* to u_*, we have to introduce some special terminology. A $\mathcal{C}$*-covering* of $D \in \mathrm{Ob}(\mathcal{D})$ is one of the form $\gamma = \langle C_i \xrightarrow{f_i} D\rangle_{i\in I} \in \mathrm{Cov}(D)(= \mathrm{Cov}_{\mathcal{D}}(D))$ such that $C_i \in \mathrm{Ob}(\mathcal{C})$

The $\mathcal{C}$-coverings generate the topology on $\mathcal{D}$; in fact, for every $R \in J(\mathcal{D})$ there is a $\mathcal{C}$-covering γ of D such that $R[\gamma] \le R$. This is a consequence of the (main) hypothesis of the theorem.

Let F be a sheaf over $\mathcal{C}$, $F \in \mathrm{Ob}(\widetilde{\mathcal{C}})$. A *morphism* $\xi : \gamma \to F$ is a family $\xi_i : C_i \to F$ $(i \in I)$ of morphisms in $\widehat{\mathcal{C}}$ (!) such that for any $i, j \in I$ there is a $\mathcal{C}$-covering $\langle C^k \xrightarrow{h_k} C_i \times_D C_j\rangle_k$ that makes, for every k, the following diagram commute:

$$\begin{tikzcd} & & C_i \arrow[dr] \arrow[drr, "\xi_i"] & & \\ C^k \arrow[r, "h_k"'] & C_i \times_D C_j \arrow[ur, "p_1"] \arrow[dr, "p_2"'] & & D & F \\ & & C_j \arrow[ur] \arrow[urr, "\xi_j"'] & & \end{tikzcd}$$

i.e., $\xi_i \circ (p_i \circ h_k) = \xi_j \circ (p_2 \circ h_k)$. *Notice* that the morphisms $p_1 \circ h_k$, $p_2 \circ h_k$ are between two objects in $\mathcal{C}$ and *therefore*, F being a sheaf over $\mathcal{C}$, the composites $\xi_i \circ (p_1 \circ h_k)$, $\xi_j \circ (p_2 \circ h_k)$ make sense.

As a first remark, we note that if $D \in \mathrm{Ob}(\mathcal{C})$, then a morphism $\xi : \gamma \to F$ is *exactly* what is called a compatible family from γ to F, in the sense of the site $\mathcal{C}$. (The reader will see that here there is something to check; the fact that F is a separated presheaf over $\mathcal{C}$ will be used.) Secondly, if in the above definition one $\mathcal{C}$-covering $\langle C^k \to C_i \times_D C_j\rangle_k$ works, then *any* $\mathcal{C}$-covering of $C_i \times_D C_j$ equally works. This again is true because F is a sheaf over $\mathcal{C}$.

$\mathrm{Hom}(\gamma, F)$ denotes the set of all morphisms $\xi : \gamma \to F$, for γ and F as above.

For $\mathcal{C}$-coverings γ, γ' of D, we write $\gamma \le \gamma'$ if $R_{\mathcal{D}}[\gamma] \le R_{\mathcal{D}}[\gamma']$ where $R_{\mathcal{D}}[\gamma]$ is the sieve $R \hookrightarrow D$ in $\widehat{\mathcal{D}}$ generated by γ. The reader is invited to write out a direct definition. Denoting the set of $\mathcal{C}$-coverings of D by $J_{\mathcal{C}}(D)$, $\le$ is a partial ordering of $J_{\mathcal{C}}(D)$ which is directed downward. In the familiar way, $J_{\mathcal{C}}(D)$ can then be considered a category.

Given $\gamma \le \gamma'$, $\mathcal{C}$-coverings of D, we define a natural map $\rho : \mathrm{Hom}(\gamma', F) \to \mathrm{Hom}(\gamma, F)$. Given a morphism $\xi' : \gamma' \to F$, the morphism $\xi = \rho\xi'$ is defined as follows. Using $\gamma \le \gamma'$, for every $C_i \to D$ in γ, we fix a $C'_{i'} \to D$ in γ' and an arrow $C_i \to C'_{i'}$, such that

$$\begin{tikzcd} C'_{i'} \arrow[r] & D \\ & C_i \arrow[ul] \arrow[u] \end{tikzcd}$$

commutes. We define $\xi_i : C_i \to F$ as the composite $C_i \to C'_{i'} \to F$, and put $\xi = \langle \xi_i \rangle_i$. It actually requires checking that this definition is correct (the result does not depend on the choice of i', etc) and that ξ is a morphism in the required sense. We are omitting the details.

Having defined de functor

$$(J_{\mathcal{D}}(D))^{\mathrm{op}} \xrightarrow{\mathrm{Hom}(\cdot,F)} \mathbf{Set}$$

now we can imitate the definition of the functor L in Section 2, to define the required quasi-inverse u^*. Given $F \in \mathrm{Ob}(\widetilde{\mathcal{C}})$, $H = u^*F : \mathcal{D}^{\mathrm{op}} \to \mathbf{Set}$ defined as follows.

(i) For $D \in \mathrm{Ob}(\mathcal{D})$,

$$H(D) = \varinjlim_{\gamma \in (J_{\mathcal{D}}(D))^{\mathrm{op}}} \mathrm{Hom}(\gamma, F).$$

(ii) For $f : D \to D'$, $H(f) : H(D') \to H(D)$ is defined as follows. Given $\gamma' \in J_{\mathcal{C}}(D')$ and $\xi' : \gamma' \to F \in \mathrm{Hom}(\gamma', F)$, first define $\gamma \in J_{\mathcal{C}}(D)$ and $\xi : \gamma \to F$ as follows. Form the pullback

$$\begin{tikzcd} D_i = C_i \times_{D'} D \arrow[r] \arrow[d] & D \arrow[d, "f"] \\ C_i' \arrow[r] & D' \end{tikzcd}$$

for $C_i' \to D'$ in γ', and choose a $\mathcal{C}$-covering $\langle C_{ik} \to D_i \rangle_k$ of D_i. Put $\gamma = \langle C_i k \to D_i \to D \rangle_{i,k}$ and $\xi = \langle C_{ik} \longrightarrow D_i \longrightarrow C_i' \xrightarrow{\xi_i'} F \rangle_{i,k}$.

Finally, the maps $\mathrm{Hom}(\gamma', F) \to \mathrm{Hom}(\gamma, F)$ thus defined induce a natural map $H(f) : H(D') \to H(D)$ by properties of colimits.

(iii) Given a natural transformation $\nu : F \to G$, $F, G \in \mathrm{Ob}(\widetilde{\mathcal{C}})$, we define $u^*(\nu) = \mu$, by defining $\mu_D : (u^*F)(D) \to (u^*G)(D)$ as follows. Given $\gamma \in J_{\mathcal{C}}(D)$ and $\xi : \gamma \to F$, by composition we can directly define $\xi' : \gamma \to G$: $\xi_i' = \nu \circ \xi_i$. The maps $\xi \mapsto \xi' : \mathrm{Hom}(\gamma, F) \to \mathrm{Hom}(\gamma, G)$ for all $\gamma \in J_{\mathcal{C}}(D)$ thus defined induce the required map $\mu_D : (u^*F)(D) \to (u^*G)(D)$.

This completes the description of the functor $u^* : \widetilde{\mathcal{C}} \to \widetilde{\mathcal{D}}$, a quasi-inverse of u_*. There is a host of things to check; e.g., that $u^*(F)$ for $F \in \mathrm{Ob}(\widetilde{\mathcal{C}})$ is a sheaf over $\mathcal{D}$. Since these details are quite similar to our previous work, we feel we can omit them. □

Corollary 1.3.17 *For any locally small site $\mathcal{D}$, $\widetilde{\mathcal{D}}$ is a Grothendieck topos, i.e., $\widetilde{\mathcal{D}}$ is equivalent to $\widetilde{\mathcal{C}}$ for a small site $\mathcal{C}$.*

Proof. Apply 1.3.16 to a suitable chosen small full subcategory $\mathcal{C}$ containing a topologically generating set for $\mathcal{D}$.

§4. Characterization of Grothendieck topoi: Giraud's theorem

In this section we will show that Grothendieck topoi have certain 'exactness' properties and that, in fact, these properties actually characterize topoi.

The following definition takes place in a fixed category.

Definition 1.4.1 (i) *An* initial object $\varnothing$ *in a category is an inductive limit of the empty diagram, i.e., $\varnothing$ is such that for every object A, there is exactly one morphism $\varnothing \to A$. $\varnothing$ is a* strict initial object *if, as the limit of the empty diagram, is 'stable under pullback', or equivalently, for any $f : B \to \varnothing$, f is an isomorphism.*

(ii) *Let A_i $(i \in I)$ be a family of objects. A* disjoint sum *of the A_i, $\coprod_i A_i$, with* canonical injections *$j_i : A_i \to \coprod_i A_i$, is such that, first of all, $\coprod_i A_i$ is the colimit (coproduct) of the A_i, with canonical injections j_i, and in addition, we have that each j_i is a monomorphism and for $i \neq j$ $A_i \times_{\coprod_i A_i} A_j$ is an initial object. The disjoint sum $\coprod_i A_i$ is* stable (under pullback) *if for any morphism* $B \to \coprod_i A_i$ *("change of basis"), the fibered products* $B \times_{\coprod A_i} A_i$ *have* B *as their disjoint sum, with canonical injections pullbacks of the* j_i*.*

(iii) *A diagram* $A \underset{p_2}{\overset{p_1}{\rightrightarrows}} B$ *is an* equivalence relation*, if* for every *object* C*, the functor* $\mathrm{Hom}(C, -)$ *transforms the given diagram into a real equivalence relation in* **Set***:*

$$\mathrm{Hom}(C, A) \underset{\circ p_2}{\overset{\circ p_1}{\rightrightarrows}} \mathrm{Hom}(C, B),$$

if denoted by

$$X \underset{q_2}{\overset{q_1}{\rightrightarrows}} Y$$

is such that $\langle x_1, x_2 \rangle \mapsto \langle q_1(x_1), q_2(x_2) \rangle \in Y \times Y$ *is a bijection of* $X \times X$ *onto an equivalence relation* $\subset Y \times Y$ *on* Y*. The equivalence relation* $A \underset{p_2}{\overset{p_1}{\rightrightarrows}} B$ *is* effective *if there is an effective epimorphism* $\pi : B \to C$ *such that*

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

is a pullback diagram. If in addition π *is stable effective, then the equivalence relation is* stable effective*.*

Remark *The notion of an equivalence relation in a category can be described entirely in terms of finite left limits in the category. Instead of giving this definition here, we refer ahead to Definition* 3.3.6 *containing the alternative definition in this spirit, but already using the logical notation, of equivalence relations. A consequence is that a left exact functor preserves equivalence relations.*

Another remark is that effective epimorphisms (hence effectivity of equivalence relations) can be described by inductive limits, viz.: a morphism $f : A \to B$ *is an effective epimorphism iff the pullback diagram*

$$\begin{tikzcd} A \arrow[r, "f"] & B \\ A \times_B A \arrow[u] \arrow[r] & A \arrow[u, "f"'] \end{tikzcd}$$

is a pushout as well.

(iv) *We say that the colimit of a diagram* D *of the category is* stable under pullbacks *if, for* R *the colimit and* $F \xrightarrow{j_F} R$ *the canonical injections, we have that for any morphism* $S \xrightarrow{\alpha} R$*,* S *is the colimit of the diagram whose objects are* $F' = F \times_S R$*, for* $F \in \mathrm{Ob}(D)$ *and whose morphisms are* $f' : F' \to G'$ *from the commutative diagram*

$$\begin{tikzcd} & R & \\ F \arrow[ur, "j_F"] \arrow[rr, "f"] & & G \arrow[ul, "j_G"'] \\ & S \arrow[uu, "\alpha"] & \\ F' \arrow[uu] \arrow[ur, "\text{p.b.}"] \arrow[rr, "f'"'] & & G'. \arrow[uu] \arrow[ul, "\text{p.b.}"'] \end{tikzcd}$$

Theorem 1.4.2 *The category of sheaves $\mathcal{E} = \widetilde{\mathcal{C}}$ over any site has the following properties:*

(i) *All (small) projective limits exist in $\mathcal{E}$.*

(ii) *All inductive limits exist in $\mathcal{E}$ and they are stable under pullback.*

(iii) *$\mathcal{E}$ has disjoint sums and all disjoint sums are stable.*

(iv) *The equivalence relations in $\mathcal{E}$ are stable effective.*

PROOF. All these facts, ultimately, are consequence of the fact that they hold for $\mathcal{E} = \mathbf{Set}$. We leave it to the reader to convince himself of this latter fact. We are going to transfer these properties of $\mathbf{Set}$ to $\widetilde{\mathcal{C}}$ in two steps: first to $\widehat{\mathcal{C}}$, then from $\widehat{\mathcal{C}}$ to $\widetilde{\mathcal{C}}$.

In CWM, Theorem 1 on page 111 (with its dual) says that the projective, as well as the inductive, limit of any (small) diagram $\mathbf{J} \xrightarrow{P} \widehat{\mathcal{C}}$ exists and can be computed pointwise, viz., if $\mathbf{J}' \xrightarrow{P'} \widehat{\mathcal{C}}$ is such that for any $A \in \mathrm{Ob}(\mathcal{C})$, the composite $\mathbf{J}' \xrightarrow{P'} \widehat{\mathcal{C}} \xrightarrow{F \mapsto F(A)} \mathbf{Set}$ is a left limit diagram, then $\mathbf{J}' \xrightarrow{P'} \widehat{\mathcal{C}}$ is a left limit diagram too; with similar statements for colimits. E.g. if

$$\begin{tikzcd} & H \arrow[dl, "p"'] \arrow[dr, "q"] & \\ F & & G \end{tikzcd}$$

is a diagram of presheaves in $\widehat{\mathcal{C}}$ such that for every $A \in \mathrm{Ob}(\mathcal{C})$,

$$\begin{tikzcd} & H(A) \arrow[dl, "p_A"'] \arrow[dr, "q_A"] & \\ F(A) & & G(A) \end{tikzcd}$$

is a product diagram, then the previous one was a product diagram in $\widehat{\mathcal{C}}$.

In particular, we have (i) and the first part of (iii) for $\mathcal{E} = \widehat{\mathcal{C}}$. The second part of (ii) and properties (iii) and (iv) for $\widehat{\mathcal{C}}$ will be seen to follow because all notions involved are defined in terms of projective and inductive limits as well as because the properties hold for $\mathbf{Set}$. E.g., let us check that the coproduct $\coprod_i F_i$, with canonical injections $j_i : F_i \to \coprod_i F_i$ is a disjoint sum. First of all, for any $A \in \mathrm{Ob}(\mathcal{C})$, $(\coprod_i F_i)(A)$ is a coproduct of the $F_i(A)$, with canonical injections $(j_i)_A$. Since the property in question holds in $\mathbf{Set}$, we have that for any $i, j \in I$, $i \neq j$, in

$$\begin{tikzcd} F_i(A) \arrow[r, "(j_i)_A"] & (\coprod_i F_i)(A) = X \\ Y = F_i(A) \times_X F_j(A) \arrow[u] \arrow[r] & F_j(A) \arrow[u, "(j_j)_A"'] \end{tikzcd}$$

Y is an initial object in $\mathbf{Set}$ (i.e., $Y = \varnothing$). But Y is $(F_i \times_{\coprod_i F_i} F_j)(A)$, by the above; since it is an initial object in $\mathbf{Set}$, for every A, so is $F_i \times_{\coprod_i F_i} F_j$ in $\widehat{\mathcal{C}}$, for the same reason.

We are leaving the rest of our claims about $\widehat{\mathcal{C}}$ to the reader to check.

Secondly, we invoke the fact of the existence of the pair of adjoint functors $a \dashv i$

$$\begin{tikzcd} \widehat{\mathcal{C}} \arrow[r, shift left, "a"] & \widetilde{\mathcal{C}} \arrow[l, shift left, "i"] \end{tikzcd}$$

such that a preserves finite left limits and i is full and faithful. It is seen in a sequence of straightforward steps that this fact alone is enough to infer properties (i)-(iv) for $\widetilde{\mathcal{C}}$, knowing them for $\widehat{\mathcal{C}}$. First of all, it follows (as we know it anyway) that $a \circ i \simeq \mathrm{id}_{\widetilde{\mathcal{C}}}$, more precisely, if ϕ is the adjunction

$$\phi_{R,F} : \mathrm{Hom}_{\widehat{\mathcal{C}}}(R, iF) \xrightarrow{\simeq} \mathrm{Hom}_{\widetilde{\mathcal{C}}}(aR, F)$$

then

$$\phi_{iF,F}(\mathrm{id}_{iF}) : aiF \xrightarrow{\simeq} F$$

for $F \in \mathrm{Ob}(\widetilde{\mathcal{C}})$.

Also, a preserves all inductive limits, i preserves all projective limits (c.f. CWM). It also follows that a left limit of an arbitrary diagram of sheaves, computed in $\widehat{\mathcal{C}}$, is already a sheaf:

$$i(\varprojlim{}^{(\widetilde{\mathcal{C}})}_{F \in \mathrm{Ob}(D)} F) = \varprojlim{}^{(\widehat{\mathcal{C}})}_{F \in \mathrm{Ob}(iD)} iF$$

but we checked this directly in 1.2.10. This shows (i).

Given a diagram D of sheaves, we claim that

$$\varinjlim{}^{(\widetilde{\mathcal{C}})}_{F \in \mathrm{Ob}(D)} F = a(\varinjlim{}^{(\widehat{\mathcal{C}})}_{F \in \mathrm{Ob}(iD)} iF)$$

$$\begin{tikzcd}
iH & & \\
R \arrow[u, dashed, "p"'] & & \\
iF \arrow[u, "j_{iF}"] \arrow[r, "if"'] \arrow[uu, bend left=70, "i(h_F)"] & iG \arrow[uul, bend right, "i(h_G)"'] & iD
\end{tikzcd}
\qquad
\begin{tikzcd}
 & H & & \\
 & aR \arrow[u, dashed, "\phi(p)"'] & & \\
aiF \arrow[ur, "a(j_{iF})"] \arrow[r, "\simeq", "\phi(\mathrm{id}_{iF})"'] & F \arrow[u, "j_F"'] \arrow[r, "f"'] \arrow[uu, bend left=40, "h_F"] & G \arrow[uul, bend right, "h_G"'] & D
\end{tikzcd}$$

Let R be the colimit of iD in $\widehat{\mathcal{C}}$, with canonical injections $j_{iF} : iF \to R$. Define for $F \in \mathrm{Ob}(D)$ the morphism

$$j_F : F \to aR$$

so that $j_F \circ \phi(\mathrm{id}_{iF}) = a(j_{iF})$. To check that in this way we indeed have a colimiting cone, let $H \in \mathrm{Ob}(\widetilde{\mathcal{C}})$ and $h_F : F \to H$, for $F \in \mathrm{Ob}(D)$, for a 'dual cone' (i.e., $h_G \circ f = h_F$ for $F, G \in \mathrm{Ob}(D)$, $f \in D$). Then i transforms this into a dual cone in $\widehat{\mathcal{C}}$, hence we will have $p : R \to iH$ such that $p \circ j_{iF} = ih_F$ ($F \in \mathrm{Ob}(D)$). $\phi(p) : aR \to H$ ($\phi = \phi_{R,H}$) will be such that $\phi(p) \circ j_F = h_F$ ($F \in \mathrm{Ob}(D)$). This follows from the commutative diagram

$$\begin{tikzcd}
aiF \arrow[r, "aj_{iF}"] \arrow[d, "\phi(\mathrm{id}_{iF})"'] & aR \arrow[d, "\phi(p)"] \\
F \arrow[r, "h_F"'] & H
\end{tikzcd}$$

The latter diagram is commutative because the two morphisms $aiF \to H$ are identical to $\phi_{iF,H}(i(h_F)) = \phi_{iF,H}(p \circ j_{iF})$, by the naturality of ϕ. Similarly, it is seen that $\phi(p)$ is the unique morphism $aR \to H$ with the required property.

This proves our claim about how inductive limits are computed in $\widetilde{\mathcal{C}}$. We leave the rest of the proof of 1.4.2 to the reader with the only remark that it should be based on our computations of projective and inductive limits in $\widetilde{\mathcal{C}}$.

Definition 1.4.3 *We call a category $\mathcal{E}$ (temporarily) a* Giraud topos *if the following are satisfied*

(i) *$\mathcal{E}$ has finite left limits.*

(ii) *$\mathcal{E}$ has disjoint sums of arbitrary sets of objects; the disjoint sums are stable under pullback.*

(iii) *The equivalence relations of $\mathcal{E}$ are stable effective.*

(iv) *$\mathcal{E}$ has a* set *of generators (c.f.* 1.3.6*).*

Theorem 1.4.5 *The following conditions on a category $\mathcal{E}$ are equivalent:*

(i) *There is a small category $\mathcal{R}$ with finite left limits such that, when $\mathcal{R}$ is considered a site with the canonical topology on $\mathcal{R}$ (c.f.* 1.1.9*), the category of sheaves on $\mathcal{R}$, $\widetilde{\mathcal{R}}$, is equivalent to $\mathcal{E}$.*

(ii) *$\mathcal{E}$ is a Grothendieck topos (i.e., $\mathcal{E} \simeq \widetilde{\mathcal{C}}$ for a small site $\mathcal{C}$).*

(iii) *There is a small category $\mathcal{C}$ and a pair of adjoint functors $a \dashv i$:*

$$\begin{tikzcd} \widehat{\mathcal{C}} \arrow[r, shift left, "a"] & \mathcal{E} \arrow[l, shift left, "i"] \end{tikzcd}$$

such that a is left exact and i is full and faithful. ($\widehat{\mathcal{C}}$ is the category of presheaves over $\mathcal{C}$).

(iv) *$\mathcal{E}$ is a Giraud topos.*

(v) *$\mathcal{E} \xrightarrow{\varepsilon_{\mathcal{E}}} \widetilde{\mathcal{E}}$ is an equivalence, where $\widetilde{\mathcal{E}}$ is the category of sheaves over $\mathcal{E}$; and $\mathcal{E}$ has a set of generators.*

Proof. (i) $\Rightarrow$ (ii) is trivial. (ii) $\Rightarrow$ (iii) was established in Section 2. As we emphasized in the proof of 1.4.2, that proof actually establishes that (iii) implies the first three conditions of 1.4.3. Similarly, the proof of 1.3.7 clearly establishes that (iii) implies that $\mathcal{E}$ has a set of generators. So, we have (iii) $\Rightarrow$ (iv).

We have the canonical functor

$$\varepsilon = \varepsilon_{\mathcal{E}} : \mathcal{E} \to \widetilde{\mathcal{E}}.$$

By 1.3.3, we have that ε is continuous, with both categories $\mathcal{E}$ and $\widetilde{\mathcal{E}}$ equipped with their canonical topologies. Moreover, since $\mathcal{E}$ as a site has the canonical topology, the representable presheaves $h_E \in \mathrm{Ob}(\widehat{\mathcal{E}})$ are already sheaves, $a(h_E) \simeq h_E$. Since ε is the composite $\mathcal{E} \xrightarrow{h} \widehat{\mathcal{E}} \xrightarrow{a} \widetilde{\mathcal{E}}$ with h the Yoneda functor which is full and faithful, it follows that ε is full and faithful.

Assume that $\mathcal{E}$ is a Giraud topos. We have the following lemma whose proof we defer to Chapter 3.

Lemma 1.4.6 *In a Giraud topos, every epimorphic family is a stable effective epimorphic family.*

Remark *Compare* 1.3.2*. From what we already know,* 1.3.2 *will be a consequence of* 1.4.6.

Let $\mathcal{R}$ be a *small* full subcategory of $\mathcal{E}$, containing a *set* of generators for $\mathcal{E}$ as well as "closed under finite left limits in $\mathcal{E}$", c.f. 1.3.16. It is easy to construct such an $\mathcal{R}$. Let $\mathcal{R} \xrightarrow{u} \mathcal{E}$ be the inclusion functor. By 1.3.10 we have the commutative diagram

$$\begin{tikzcd} \mathcal{E} \arrow[r, "\varepsilon = \varepsilon_{\mathcal{E}}"] & \widetilde{\mathcal{E}} \\ \mathcal{R} \arrow[u, hook, "u"] \arrow[r, "\varepsilon_{\mathcal{R}}"'] & \widetilde{\mathcal{R}}. \arrow[u, "u^*"', "\simeq"] \end{tikzcd}$$

From 1.3.16, we also know that u^* is an equivalence. By 1.3.7, the set $\{\varepsilon_{\mathcal{R}}(R) : R \in \mathrm{Ob}(\mathcal{R})\}$ of objects is a set of generators for $\widetilde{\mathcal{R}}$. It follows that for any sheaf $X \in \mathrm{Ob}(\widetilde{\mathcal{E}})$ there is a *set* of morphisms of the form $\langle \varepsilon(E_i) \xrightarrow{f_i} X\rangle_{i\in I}$ which is an epimorphic family. By 1.4.6, $\langle \varepsilon(E_i) \xrightarrow{f_i} X\rangle_{i\in I}$ is an effective epimorphic family.

Lemma 1.4.7 *Given any category with finite left limits, suppose that $\langle Y_i \xrightarrow{f_i} X\rangle_{i\in I}$ is an effective epimorphic family. Assume the coproduct $\coprod_{i\in I} Y_i$ exists, with canonical injections $Y_i \xrightarrow{\alpha_i} \coprod_{i\in I} Y_i$. Then $f : \coprod_{i\in I} Y_i \to X$ induced by f_i such that $f \circ \alpha_i = f_i$ is an effective epimorphism.*

Proof. This is an elementary computation

$$\begin{tikzcd}
& Y_i \arrow[rr, "\alpha_i"] \arrow[dr, phantom, "\circlearrowleft"] & & \coprod Y_i \arrow[dr, "f"] \arrow[drr, bend left, "g"] & & \\
Y_i \times_X Y_j \arrow[ur, "q_1"] \arrow[dr, "q_2"'] \arrow[rr, dotted, "\beta"'] & & \coprod Y_i \times_X \coprod Y_i = U \arrow[ur, "p_1"] \arrow[dr, "p_2"'] & & X \arrow[r, dotted, "h"'] & Z \\
& Y_j \arrow[rr, "\alpha_j"'] \arrow[ur, phantom, "\circlearrowright"] & & \coprod Y_i \arrow[ur, "f"'] \arrow[urr, bend right, "g"'] & &
\end{tikzcd}$$

Assume that $\coprod Y_i \xrightarrow{g} Z$ is such that $g \circ p_1 = g \circ p_2$ for p_1, p_2 the canonical projections $U = \coprod Y_i \times_X \coprod Y_i \to \coprod Y_i$. Form the fibered products

$$\begin{tikzcd}
& Y_i \arrow[dr, "f_i"] & \\
Y_i \times_X Y_j \arrow[ur, "q_i"] \arrow[dr, "q_j"'] & & X \\
& Y_j \arrow[ur, "f_j"'] &
\end{tikzcd}$$

and use the fact that $f_i = f \circ \alpha_i$. It follows that for any given i and j, there is $\beta : Y_i \times_X Y_j \to Y$ such that $p_1 \circ \beta = \alpha_i \circ q_1$, $p_2 \circ \beta = \alpha_j \circ q_2$. Putting $g_i = g \circ \alpha_i$, it follows that the g_i form a compatible family: $g_i \circ q_1 = g_j \circ q_2$. Since $\langle Y_i \xrightarrow{f_i} X\rangle_i$ is effective epimorphic, there is a unique $h : X \to Z$ such that $h \circ f_i = g_i$ for every $i \in I$. It follows that $(h \circ f) \circ \alpha_i = g \circ \alpha_i$ for every $i \in I$. Since $\coprod Y_i$ is a coproduct, α_i are canonical, we have that $h \circ f = g$ as required for f being effective epimorphic. The uniqueness of h with $h \circ f = g$ follows from the uniqueness of h with $h \circ f_i = g_i$ $(i \in I)$. $\square$

Returning to the proof of the theorem, we look at the effective epimorphic family $\langle Y_i \xrightarrow{f_i} X\rangle_{i\in I}$, for $Y_i = \varepsilon(E_i)$ in $\widetilde{\mathcal{E}}$. Consider the disjoint sum $Y = \coprod Y_i$ in $\widetilde{\mathcal{E}}$. Since ε preserves disjoint sums (c.f. 3.4.10 and 3.4.13), we have that $Y = \varepsilon(E)$ for $E = \coprod_i E_i$, and the canonical injections $\varepsilon(E_i) \to \varepsilon(E)$ are $\varepsilon(\alpha_i)$ where $\alpha_i : E_i \to \coprod_i E_i$ are the canonical injections in $\mathcal{E}$. It follows by 1.4.7 that *we have an effective epimorphism of the form $\varepsilon(E) \to X$ for every $X \in \mathrm{Ob}(\widetilde{\mathcal{E}})$.*

To move on from here to being able to say that $X \simeq \varepsilon(E')$ for some $E' \in \mathrm{Ob}(\mathcal{E})$ we have to invoke a general lemma, c.f. 1.4.9 below. First some terminology. A functor $F : \mathcal{R} \to \mathcal{S}$ is called *conservative (with respect to monomorphisms)* if whenever $R_1 \xrightarrow{i} R_2$ is a monomorphism in $\mathcal{R}$ such that $F(R_1) \xrightarrow{F(i)} F(R_2)$ is an isomorphism, then i is an isomorphism.

Lemma 1.4.8 *$\varepsilon : \mathcal{E} \to \widetilde{\mathcal{E}}$ is conservative.*

Proof. Let $E_1 \xrightarrow{i} E_2$ be a monomorphism in $\mathcal{E}$ such that $\varepsilon(E_1) \xrightarrow{\varepsilon(i)} \varepsilon(E_2)$ is an isomorphism, hence an effective epimorphism in $\widetilde{\mathcal{E}}$. By 1.3.3, i is an effective epimorphism. By 1.3.4, i is an isomorphism.

Lemma 1.4.9 *Let $I:\mathcal{R}\to\mathcal{S}$ be a functor between categories with finite left limits and assume the following:*

(i) *I is full and conservative with respect to monomorphisms.*

(ii) *I preserves effective epimorphisms and finite left limits.*

(iii) *The equivalence relations in $\mathcal{R}$ are effective.*

(iv) *For every object S of $\mathcal{S}$ there is $R\in\mathrm{Ob}(\mathcal{R})$ and effective epimorphism $I(R)\overset{p}{\twoheadrightarrow} S$.*

Then I is an equivalence of categories.

Proof. We use the following simple remarks which, nevertheless, we will verify in Chapter 3 only.

Sublemma 1.4.10 *I is faithful; moreover, if $I(P)\underset{I(p_2)}{\overset{I(p_1)}{\rightrightarrows}}I(R)$ is an equivalence relation (in $\mathcal{S}$), then so is $P\underset{p_2}{\overset{p_1}{\rightrightarrows}}R$ (in $\mathcal{R}$).*

It is enough to verify that for every $S\in\mathrm{Ob}(\mathcal{S})$, there is $R\in\mathrm{Ob}(\mathcal{R})$ such that $S\simeq I(R)$.

Let $S\in\mathrm{Ob}(\mathcal{S})$ be arbitrary. Let $IR\xrightarrow{p}S$ be an effective epimorphism and $S'\underset{q_2}{\overset{q_1}{\rightrightarrows}}IR$ the kernel-pair of p:

$$\begin{tikzcd} & IR \arrow[dr, "p"] & \\ S' \arrow[ur, "q_1"] \arrow[dr, "q_2"'] \arrow[rr, phantom, "\text{p.b.}"] & & S. \\ & IR \arrow[ur, "p"'] & \end{tikzcd}$$

Consider the product $R\xleftarrow{\alpha_1}R\times R\xrightarrow{\alpha_2}R$ and use the fact that I preserves products. We obtain $I(R\times R)\simeq I(R)\times I(R)$ and the following diagram

$$\begin{tikzcd} & IR \arrow[dr, "p"] & \\ S' \arrow[ur, "q_1"] \arrow[r, "i"] \arrow[dr, "q_2"'] & I(R\times R) \arrow[u, "I(\alpha_1)"'] \arrow[d, "I(\alpha_2)"] & S \\ & IR \arrow[ur, "p"'] & \end{tikzcd}$$

such that i is a monomorphism. Applying hypothesis (iv) to the object S', we have an effective epimorphism $IR''\xrightarrow{p'}S'$. The fullness of I applied to the morphism

$$IR''\xrightarrow{p'}S'\xrightarrow{i}I(R\times R)$$

gives us $\beta:R''\to R\times R$ such that $I(\beta)=i\circ p'$. Consider the kernel-pair

$$\begin{tikzcd} & R'' \arrow[dr, "\beta"] & \\ R''' \arrow[ur, "r_1"] \arrow[dr, "r_2"'] \arrow[rr, phantom, "\text{p.b.}"] & & R\times R \\ & R'' \arrow[ur, "\beta"'] & \end{tikzcd}$$

Then $R'''\underset{r_2}{\overset{r_1}{\rightrightarrows}}R''$ is an equivalence relation in $\mathcal{R}$, hence by (iii), there is a pullback diagram

$$\begin{tikzcd} & R'' \arrow[dr, "p''"] & \\ R''' \arrow[ur, "r_1"] \arrow[dr, "r_2"'] & & R' \\ & R'' \arrow[ur, "p''"'] & \end{tikzcd}$$

where p'' is an effective epimorphism. Hence

$$\begin{tikzcd} & IR'' \arrow[dr, "Ip''"] & \\ IR''' \arrow[ur, "Ir_1"] \arrow[dr, "Ir_2"'] & & IR' \\ & IR'' \arrow[ur, "Ip''"'] & \end{tikzcd} \tag{1}$$

is a pullback diagram and Ip'' is an effective epimorphism, by assumption (ii). Also

$$\begin{tikzcd} & IR'' \arrow[dr, "I(\beta)"] & \\ IR''' \arrow[ur, "Ir_1"] \arrow[dr, "Ir_2"'] & & I(R \times R) \\ & IR'' \arrow[ur, "I(\beta)"'] & \end{tikzcd}$$

is a pullback diagram, and since i is a monomorphism, $I(\beta) = i \circ p'$, so is

$$\begin{tikzcd} & IR'' \arrow[dr, "p'"] & \\ IR''' \arrow[ur, "Ir_1"] \arrow[dr, "Ir_2"'] & & S'. \\ & IR'' \arrow[ur, "p'"'] & \end{tikzcd} \tag{2}$$

Comparing diagrams (1) and (2), both of which being pullback diagrams and Ip'', p' being effective epimorphisms, by the definition of "effective epimorphism" it follows that $IR' \simeq S'$.

Returning to the diagram defining S', we conclude that we have a pullback diagram of the form

$$\begin{tikzcd} & IR \arrow[dr, "p"] & \\ IR' \arrow[ur, "q_1"] \arrow[dr, "q_2"'] & & S \\ & IR \arrow[ur, "p"'] & \end{tikzcd} \tag{3}$$

with p an effective epimorphism. By the fullness of I, $q_i = I(\rho_i)$, $i = 1, 2$, for some ρ_1, ρ_2.

Since $IR' \overset{q_1}{\underset{q_2}{\rightrightarrows}} IR$ is a kernel pair, it is an equivalence relation. By the conservativeness of I and 1.4.10, $R' \overset{\rho_2}{\underset{\rho_2}{\rightrightarrows}} R$ is an equivalence relation. By (iii), let $\pi : R \to \overline{R}$ be an effective epimorphism such that

$$\begin{tikzcd} & R \arrow[dr, "\pi"] & \\ R' \arrow[ur, "\rho_1"] \arrow[dr, "\rho_2"'] & & \overline{R} \\ & R \arrow[ur, "\pi"'] & \end{tikzcd}$$

is a pullback diagram. These properties are inherited to

$$\begin{tikzcd}[column sep=small] & IR \arrow[dr, "I\pi"] & \\ IR' \arrow[ur, "q_1=I\rho_1"] \arrow[dr, "q_2=I\rho_2"'] & & I\overline{R}. \\ & IR \arrow[ur, "I\pi"'] & \end{tikzcd}$$

Comparing the last diagram with (3) having similar properties, it follows that $I\overline{R} \simeq S$, proving the lemma. □

Returning to the proof of the theorem, let us remind the reader that under the hypothesis that $\mathcal{E}$ is a Giraud topos, above we have shown that $\varepsilon_{\mathcal{E}} : \mathcal{E} \to \widetilde{\mathcal{E}}$ satisfies the hypothesis of Lemma 1.4.9. Hence $\varepsilon_{\mathcal{E}}$ is an equivalence as stated in (v).

At the same time, we have shown that $\mathcal{E} \simeq \widetilde{\mathcal{E}} \simeq \widetilde{\mathcal{R}}$ for a small subcategory $\mathcal{R}$ of $\mathcal{E}$, proving that (i) also follows from (iv). The same argument can be repeated under the hypothesis (v) to prove that (v) implies (i).

This completes the proof of the main Theorem 1.4.5.

Notice that, in a roundabout way, we have established Lemma 1.3.14.

Finally, for later reference we state a version of Lemma 1.4.9 whose proof is contained in that of 1.4.9.

Lemma 1.4.11 *Let I, $\mathcal{R}$, $\mathcal{S}$ be as in* 1.4.9 *and assume conditions* (i), (ii), (iii) *of* 1.4.9 *but drop* (iv). *Let S be a fixed object of $\mathcal{S}$. Assume (in place of* (iv)*) that there is an effective epimorphism of the form $I(R) \xrightarrow{p} S$, $R \in \mathrm{Ob}(\mathcal{R})$, moreover, with the same R, there is another epimorphism of the form $I(R'') \to S' = I(R) \times_S I(R)$. Then there is an object $\overline{R}$ of $\mathcal{R}$ such that $I(\overline{R}) \simeq S$.*

For the proof of the lemma, notice that the condition replacing (iv) contains exactly what is needed in the proof of 1.4.9 to show the existence of the required $\overline{R}$. We note that we also say 'S is exactly covered via I' for the condition in 1.4.11 replacing (iv) of 1.4.9.

Appendix to Chapter 1.
Concepts of local character, examples.

The aim of this appendix is to present a few examples of topoi, trying to motivate the notions of site and topos.

Just as set theory formalizes the notion of "collection", one can consider that topos theory formalizes the notion of "concept". In fact, a concept may be considered as a "variable extension", parametrized by the domains of applications of the concept. When one disposes of a notion of "localization" at the level of the domains of application, one arrives to the basic notion of "concept of local character".

A few examples will hopefully clarify these remarks.

1. The concept of "real-valued continuous function defined on an open set of a topological space X" has, as domains of applications the open sets $\mathrm{Open}(X)$ of X. The family

$$\langle C_{\mathbb{R}}(U) \rangle_{U \in \mathrm{Open}(X)}$$

where

$$C_{\mathbb{R}}(U) = \text{ all real-valued continuous functions defined on } U$$

is a "variable extension" parametrized by $\mathrm{Open}(X)$. Notice, however, that $\mathrm{Open}(X)$ is a category (and not just a set) whose objects are the open sets of X and its morphisms are the inclusions. In this context, $C_{\mathbb{R}}$, rather than the family parametrized by $\mathrm{Open}(X)$, is a functor

$$C_{\mathbb{R}} : \mathrm{Open}(X)^{\mathrm{op}} \to \mathbf{Set}$$

which acts on inclusions by restrictions.

The concept in question may be identified with this functor.

Let us notice that we have a notion of localization at the level of the domains of application (of the concept), namely every open covering of an open set U is a localization of U. Furthermore, the concept in question (i.e., "real valued continuous function") is of local character in the following sense: if $\langle U_i \rangle_{i\in I}$ covers U, then

(i) Every $f \in C_{\mathbb{R}}(U)$ may be recovered from its localizations, i.e. if $g \in C_{\mathbb{R}}(U)$ is such that $f|_{U_i} = g|_{U_i}$, for all $i \in I$, then $f = g$.

(ii) (Glueing condition for compatible families.) If $\langle f_i \rangle_{i\in I}$ is a family such that $f_i \in C_{\mathbb{R}}(U_i)$, for all $i \in I$ and

$$(*) \qquad f_i|_{U_i\cap U_j} = f_j|_{U_i\cap U_j}, \qquad \text{for every } i, j \in I,$$

then there is $f \in C_{\mathbb{R}}(U)$ such that $f|_{U_i} = f_i$, for all $i \in I$.

The reader will notice that (i) and (ii) are equivalent to require that the diagram

$$C_{\mathbb{R}}(U) \longrightarrow \prod_{i\in I} C_{\mathbb{R}}(U_i) \rightrightarrows \prod_{i,j\in I} C_{\mathbb{R}}(U_i \cap U_j) \qquad \text{is exact.}$$

2. The concept of "solutions of a finite system of polynomial equations $p_1 = 0, \ldots, p_m = 0$ with integer coefficients in the indeterminates $X_1, \ldots, X_n$" may be identified with the functor

$$S : \mathcal{R} \to \mathbf{Set}$$

such that

$$S(A) = \{\langle a_1, \ldots, a_n \rangle \in A^n : \langle a_1, \ldots, a_n \rangle \text{ is a common root of } p_1, \ldots, p_m \text{ in } A\}.$$

By $\mathcal{R}$ we mean the category of commutative rings with 1. Sometimes one considers, instead of $\mathcal{R}$, the full subcategory $\mathcal{R}_f \hookrightarrow \mathcal{R}$ of finitely presented rings, i.e. of the form $\mathbb{Z}[X_1, \ldots, X_n]/\langle f_1, \ldots, f_m \rangle$. The reason is that $\mathcal{R}_f$ is (equivalent to a) small category and we can talk e.g. of $\mathbf{Set}^{\mathcal{R}_f}$.

At first sight, no notion of localization hits the eye. However, several are available, in particular the Zariski localization in $\mathcal{R}^{\mathrm{op}}$ defined as follows:

if $A \xrightarrow{I_a} A\left[\frac{1}{a}\right] \in \mathcal{R}$ is the solution of the universal problem of inverting $a \in A$, (see e.g. Atiyah-MacDonald [1969]), the family

$$\begin{tikzcd}[row sep=small] & A\left[\frac{1}{a_1}\right] \\ A \arrow[ur] \arrow[dr] & \vdots \\ & A\left[\frac{1}{a_n}\right] \end{tikzcd}$$

where $a_1 + \cdots + a_n = 1$ will be considered as a localization in $\mathcal{R}^{\mathrm{op}}$ (hence a "co-localization" in $\mathcal{R}$). Furthermore, the empty family is considered as a co-localization of the null ring.

Intuitively, one can think of $\left(A\left[\frac{1}{a_i}\right]\right)_{i\leq n}$ as an "open cover" of A in $\mathcal{R}^{\mathrm{op}}$. (This is just that if one describes $\mathcal{R}^{\mathrm{op}}$ as the category of affine schemes.)

Let us answer the question whether our concept is of local character.

To make the analogy with our previous example closer, let us first notice that the functor S is representable by the ring $B = \mathbb{Z}[X_1, \ldots, X_n]/\langle p_1, \ldots, p_m\rangle$.

Indeed

$$h^B(A) = (\text{the set of ring homomorphisms from } B \text{ to } A) \simeq S(A).$$

(Since every "solution" $\langle a_1, \ldots, a_n\rangle \in S(A)$ gives a unique homomorphism of B to A via $X_i \mapsto a_i$.)

But this means that $S \simeq h^B$, i.e., S is representable by the ring B.

Let us now formulate the analogues of (i) and (ii) of our previous examples.

(i)' Every $f \in h^B(A)$ may be recovered from its localizations, i.e., if $g \in h^B(A)$ is such that

$$I_{a_i}(f) = I_{a_i}(g) \in h^B\left(A\left[\frac{1}{a_i}\right]\right) \quad \text{for every } i \leq n,$$

then $f = g$.

(ii)' if $\langle f_i\rangle_{i\leq n}$ is a family such that $f_i \in h^B\left(A\left[\frac{1}{a_i}\right]\right)$ for all $i \leq n$ and furthermore,

$$(*)' \qquad B \xrightarrow{f_i} A\left[\tfrac{1}{a_i}\right] \longrightarrow A\left[\tfrac{1}{a_ia_j}\right] \;=\; B \xrightarrow{f_j} A\left[\tfrac{1}{a_j}\right] \longrightarrow A\left[\tfrac{1}{a_ia_j}\right]$$

then there is $f \in h^B(A)$ such that all the diagrams

$$\begin{tikzcd} B \arrow[r, "f"] \arrow[dr, "f_i"'] & A \arrow[d, "I_{a_i}"] \\ & A\left[\frac{1}{a_i}\right] \end{tikzcd}$$

commute.

We notice that this condition is the exact analogue of $(*)$ in our first example, thinking of $A\left[\frac{1}{a_i}\right]$ as an open subset of A (in $\mathcal{R}^{\mathrm{op}}$), since,

$$U_i \cap U_j = U_i \times_U U_j \quad \text{in the category } \mathrm{Open}(X) \text{ and}$$

$$A\left[\frac{1}{a_ia_j}\right] = A\left[\frac{1}{a_i}\right] \times_A A\left[\frac{1}{a_j}\right] \quad \text{in } \mathcal{R}^{\mathrm{op}}.$$

Now (i)' and (ii)' are consequences of the following

Proposition. *In $\mathcal{R}$, the push-out,*

$$\begin{tikzcd} & A\left[\frac{1}{a}\right] \arrow[dr] & \\ A \arrow[ur] \arrow[dr] & & A\left[\frac{1}{ab}\right] \\ & A\left[\frac{1}{b}\right] \arrow[ur] & \end{tikzcd}$$

is a pull-back, provided that $a + b = 1$.

PROOF. We shall only prove that elements of $A\left[\frac{1}{a}\right]$, $A\left[\frac{1}{b}\right]$ having the same image in $A\left[\frac{1}{ab}\right]$ come from an element of A. The rest of the proof is left to the reader.

We first note that $a+b=1$ implies that the ideal $\langle a^n, b^n\rangle$ is the unit ideal A, for all $n \geq 0$, since (using the binomial expansion)

$$1 = (a+b)^{2n-1} = \lambda a^n + \mu b^n, \quad \text{for some } \lambda, \mu \in A.$$

(This argument was pointed out to us by E. Dubuc.)

Let $\frac{s}{a^n} \in A\left[\frac{1}{a}\right]$ and $\frac{t}{b^m} \in A\left[\frac{1}{b}\right]$ be such that $\frac{s}{a^n} = \frac{t}{b^m}$ in $A\left[\frac{1}{ab}\right]$. We may obviously assume that $m = n$.

Therefore, there is $p \geq 0$ such that

$$(ab)^p(sb^n - ta^n) = 0 \quad \text{in } A, \text{ i.e.,}$$

$$a^p b^m s = a^m b^p t \text{ for } m = p + n.$$

By the observation at the beginning

$$1 = \lambda a^m + \mu b^m, \quad \text{for some } \lambda, \mu \in A$$

and we can define

$$z = \lambda a^p s + \mu b^p t \in A.$$

Then,

$$\begin{aligned} a^m z &= \lambda a^p s a^m + \mu a^m b^p t \\ &= \lambda a^p s a^m + \mu a^p b^m s \\ &= a^p s(\lambda a^m + \mu b^m) = a^p s \end{aligned}$$

and this shows that the canonical homomorphism $A \to A\left[\frac{1}{a}\right]$ sends z into $\frac{s}{a^n} \in A\left[\frac{1}{a}\right]$.

Similarly, $A \to A\left[\frac{1}{b}\right]$ sends z into $\frac{t}{b^n} \in A\left[\frac{1}{b}\right]$.

To check uniqueness of z, assume that z' is such that $z = z'$ in both $A\left[\frac{1}{a}\right]$ and $A\left[\frac{1}{b}\right]$. Then there are $p, q \geq 0$ such that $a^p u = b^q u = 0$, where $u = z - z'$. We may assume that $p = q$. From our observation $1 = \lambda a^p + \mu b^p$ for some $\lambda, \mu \in A$ and this implies that $u = \lambda a^p u + \mu b^p u = 0$.

Let us notice that this Proposition implies, more generally, that any representable functor

$$F : \mathcal{R} \to \mathbf{Set}$$

is local for the Zariski localization on $\mathcal{R}^{\mathrm{op}}$.

In particular, the concepts of "being invertible" and "being an element" are of local character, since they may be identified with the representable functors

$$h^{\mathbb{Z}[X,Y]/\langle XY-1\rangle} \text{ and } h^{\mathbb{Z}[X]}, \text{ respectively.}$$

3. The concept "partial element of a set X" may be analyzed as follows: the domains of applications from a complete Heyting Algebra $\mathcal{H}$ whose elements may be thought as "degrees of existence". The concept itself may be identified with a certain functor

$$\underline{X} : \mathcal{H}^{\mathrm{op}} \to \mathbf{Set}$$

such that, intuitively,

$$\underline{X}(h) = \{x \in X : \text{degree of existence of } x \text{ is at least } h\}.$$

Turning the tables, the degree of existence of x could be defined (having $\underline{X}$ available) as the largest $h \in \mathcal{H}$ such that $x \in \underline{X}(h)$. Unfortunately, such an x does not always exist for an arbitrary functor.

It does exist, however, precisely when the following condition is satisfied:

$$\text{if} \;\; h_0 = \bigvee \{h : x \in \underline{X}(h)\}, \;\text{then}\; x \in \underline{X}(h_0).$$

On easily checks that this condition may be expressed as follows:

$$\text{if} \;\; h = \bigvee_{i \in I} h_i, \;\text{then}\; \underline{X}(h) = \bigcap_{i \in I} \underline{X}(h_i).$$

These remarks suggest to consider the family $\langle h_i \rangle_{i \in I}$ as a localization of h, whenever $h = \bigvee_{i \in I} h_i$. Our condition expresses the local character of the concept "partial element of X".

4. (This example will be discussed more fully and from a syntactical point of view in Chapter 9.)

Let T be a first order finitary theory in a countable language L. Let $\mathrm{Mod}(T)$ be the category of countable models of T with *algebraic* homomorphisms (i.e., preserving relations and operations in the following sense: $M \xrightarrow{f} N$ is algebraic if $\langle a_1, \ldots, a_n \rangle \in R^M \Rightarrow \langle fa_1, \ldots, fa_n \rangle \in R^N$ for every primitive n-ary relation symbol R, with a similar clause for operations).

If $\phi(x_1, \ldots, x_n)$ is a *coherent* formula of L, i.e., obtained from the atomic formulas by using $\vee$, $\wedge$, $\exists$, $\uparrow$ (true), $\downarrow$ (false) as the only logical operators, then ϕ gives rise to a functor

$$\phi^{(\;)} : (\mathrm{Mod}(T)^{\mathrm{op}})^{\mathrm{op}} \to \mathbf{Set}$$

defined by

$$\phi^{(M)} = \{\langle a_1, \ldots, a_n \rangle \in |M|^n : M \models \phi[a_1, \ldots, a_n]\}.$$

Indeed all coherent formulas are obviously preserved by algebraic homomorphisms. An exercise of [CK] tell us that the converse is true.

Proposition. *Let $\phi(x_1, \ldots, x_n) \in L$. Then $\phi^{(\;)}$ is a subfunctor of the n^{th} power of the forgetful functor $|\;|$ iff ϕ is T-equivalent to a coherent formula.*

We shall call a functor *definable* if it is of this form (i.e., of the form $\phi^{(\;)}$, for some coherent formula $\phi \in L$).

A natural transformation $\phi(x_1, \ldots, x_n)^{(\;)} \xrightarrow{\eta} \psi(x_1, \ldots, x_m)^{(\;)}$ is *definable* if there is a coherent formula $\Phi(x_1, \ldots, x_n; y_1, \ldots, y_m)$ such that $\Phi^{(M)}$ is the graph of the function η_M, for all $M \in \mathrm{Mod}(T)$.

We let $\mathcal{D}(T)$ be the subcategory of $\mathbf{Set}^{\mathrm{Mod}(T)}$ consisting of definable functors and definable natural transformations.

Proposition. *$\mathcal{D}(T)$ has finite $\varprojlim$, images which are stable under pull-backs and supremum (of finite families of sub-objects of a given object) which are also stable under pull-backs.*

Proof (Sketch). Let us recall some definitions: a finite sup $A = \bigvee_{i \in I} A_i$ is stable (under pull-backs) if for every $B \to A \in \mathcal{D}(T)$ $B \simeq \bigvee_{i \in I} A_i \times_A B$.

An image $A \stackrel{f}{\twoheadrightarrow} B$ (i.e., such that f does not factor thru a proper sub object of B) is stable under pull-backs if for every $C \to B \in \mathcal{D}(T)$, the horizontal lower arrow is again an image in the diagram

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \uparrow & & \uparrow \\ A \times_B C & \longrightarrow & C. \end{array}$$

For the proof, we use the fact that every category of functors, in particular $\mathbf{Set}^{\mathrm{Mod}(T)}$, satisfies the conclusion of the Proposition. (Indeed $\varprojlim$, images and sups are computed point-wise and so we are dealing essentially with $\mathbf{Set}$.)

Let $F, G \in |\mathcal{D}(T)|$. Then $\theta^{(\;)}$ is the product $F \times G$, where

$$\theta(x_1, \ldots, x_n, x_{n+1}, \ldots, x_m) = \phi(x_1, \ldots, x_n) \wedge \psi(x_{n+1}/x_1, \ldots, x_{n+m}/x_m)$$

and $F = \phi^{(\;)}$, $G = \psi^{(\;)}$. Notice that the canonical projections $F \times G \xrightarrow{\pi_1} F$, $F \times G \xrightarrow{\pi_2} G$ are also definable, e.g.,

$$\pi_1(x_1, \ldots, x_{n+m}; y_1, \ldots, y_n) = \theta \wedge y_1 = x_1 \wedge \cdots \wedge y_n = x_n.$$

The final object of $\mathcal{D}(T)$ is the functor $\uparrow^{(\;)}$ which is also the final object of $\mathbf{Set}^{\mathrm{Mod}(T)}$.

Let us indicate the existence of images. Assume that

$$\phi(x_1, \ldots, x_n)^{(\;)} \xrightarrow{\eta} \psi(x_1, \ldots, x_m)^{(\;)}$$

is definable by a formula $\Phi(x_1, \ldots, x_n; y_1, \ldots, y_m)$. Then the functor image of ϕ under η, in $\mathbf{Set}^{\mathrm{Mod}(T)}$, is definable by the formula

$$\exists x_1 \cdots \exists x_n \Phi(x_1, \ldots, x_n; y_1, \ldots, y_m).$$

Details are left to the reader.

The category $\mathcal{D}(T)$ can be made into a site by defining a localization of $A \in |\mathcal{D}(T)|$ as a finite family $\langle A_i \xrightarrow{f_i} A \rangle_{i \in I}$ such that $A = \bigvee_{i \in I} \mathrm{Im}(f_i)$, where $\mathrm{Im}(f_i)$ denotes the image of A_i under f_i. This localization will be called the *precanonical localization*. (One should notice that the stability of images and sups assures us that this is indeed a localization.)

Definition. A category satisfying the conclusion of the Proposition will be called *logical*.

A functor between logical categories is *logical* if it preserves finite $\varprojlim$, finite sups and images.

Corollary. *The inclusion functor $\mathcal{D}(T) \to \mathbf{Set}^{\mathrm{Mod}(T)}$ is logical.*

Chapter 2

Interpretation of the logic $L_{\infty\omega}$ in categories

Introduction

Categorical formulations of logic were initiated by W. Lawvere, c.f. Lawvere [1965] and other references in Kock and Reyes [1977]. In particular, the categorical (or "functorial") interpretation of quantifiers is due to Lawvere, loc. cit. In the present work, the fundamental notion is the interpretation of formulas of the logic $L_{\infty\omega}$ in categories. This device is necessary to make the connections between ordinary formulations in logic on the one hand, and categorical logic on the other, explicit. The categorical interpretation of logic appears first in the work of Mitchell [1972]. It was Joyal and Reyes (c.f. Reyes [1974]) who isolated the notion of a logical category (in loc. cit. "regular category with stable sups") as the basic notion for categorical logic.

§1. The logic $L_{\infty\omega}$

In this section we briefly describe the basic terminology related to the infinitary logic $L_{\infty\omega}$. For more details, c.f. e.g. Barwise [1975]. For the many-sorted formulation we need, c.f. also Feferman [1968].

A *language* L is a collection of symbols falling into the following disjoint classes:

(1) a non-empty set of elements which are called *sorts*,

(2) a set of *finitary sorted predicate symbols*,

(3) a set of *finitary sorted operation symbols.*

In more detail: a predicate symbol R in class (2) is equipped with a natural number n, the *number of places of* R; R is called an n-ary predicate symbol. Also, R is equipped with an assignment of a sort s_i (a symbol in group (1)) to each $i = 1, \dots, n$; s_i is the *sort of the* i^{th} *place* of R. The operative effect of this assignment will be that only variables of the right sort can occupy a given place when forming formulas using R. We write "$R \subset s_1 \times \cdots \times s_n$" to indicate the "sorting" of R. Similarly, an operation symbol f in class (3) is equipped with a natural number n, the *number of places of* f; f is called an n*-ary operation symbol.* Moreover, f is equipped with sorts s_i for $i = 1, \dots, n$, s_i being called *the sort of the* i^{th} *place of* f and finally, also with an additional sort s, called the *sort of the value of* f. We write $f : s_1 \times \cdots \times s_n \to s$, anticipating the "intended meaning" (c.f. below).

A 0-ary operation symbol is called an *individual constant.*

Given a language L, we form the logic $L_{\infty\omega}$ based on L by using some additional symbols. These additional symbols are as follows:

(4) a set of *free (individual) variables of sort* s, for each sort s in L; this set should be infinite but can be taken to be a countable set irrespective of the cardinality of L,

(5) a set of *bound (individual) variables* of sort s, for each sort s in L; similarly as under (4), this set can be taken to be a countably infinite set for each s.

(6) $\approx$, the symbol for *identity*; $\bigvee$, (infinitary) *disjunction* symbol; $\bigwedge$, (infinitary) *conjunction* symbol; $\neg$, *negation* symbol; $\rightarrow$, *implication* symbol; $\exists$, *existential quantifier* symbol; $\forall$, *universal quantifier* symbol.

On the basis of the symbols, we define *terms, atomic formulas* and formulas of $L_{\infty\omega}$ as follows.

Every free variable and every individual constant is a term, in fact, a *term of* the sort originally assigned to it. If f is an n-ary operation symbol, $n > 0$, and in particular $f : s_1 \times \cdots \times s_n \to s$, and if $t_1, \ldots, t_n$ are *terms of* the sorts $s_1, \ldots, s_n$, respectively, then $ft_1 \cdots t_n$ is a *term of* sort s.

We remark that by the foregoing description we meant to give an inductive definition of the set of all terms; in particular, every term is one that is obtained in the way described. As is familiar from many analogous situations, the definition could be phrased as an explicit definition of the set of all terms as the smallest set satisfying certain obvious closure conditions. (Actually, the more complicated notion "t is a term of sort s" is being defined by induction.) As another remark, note that the exact identity of the object $ft_1 \cdots t_n$ is largely irrelevant except that we should be able to recover each of $f, t_1, \ldots, t_n$ from $ft_1 \cdots t_n$ in a unique fashion ("unique readability"). These remarks apply, mutatis mutandis, to the definition of formulas below.

The *atomic formulas* of $L_{\infty\omega}$ are the expressions of the form $Pt_1 \cdots t_n$, with P an n-ary predicate symbol and with $t_1, \ldots, t_n$ terms, or of the form $t_1 \approx t_2$, with $\approx$ the symbol for identity, and t_1, t_2 terms, subject to the following restrictions on sorts. If $P \subset s_1 \times \cdots \times s_n$, then $t_1, \ldots, t_n$ must have sorts $s_1, \ldots, s_n$, respectively. In $t_1 \approx t_2$, t_1 and t_2 must have the same (but otherwise arbitrary) sort.

The formulas of $L_{\infty\omega}$ are formed by repeatedly applying the logical operators to formulas and sets of formulas. We also stipulate that our formulas should contain finitely many free variables only. We take for granted the notion of *substitution*: $\phi(x/w)$ denotes the result of substituting w for x at each occurrence of the free variable in ϕ. Accordingly, the class of formulas is the least class X (actually, a proper class) such that

(i) X contains all atomic formulas;

(ii) X contains $\neg\phi$, $\phi \to \psi$, $\bigwedge \Theta$, $\bigvee \Theta$ whenever $\phi, \psi \in X$, $\Theta \subset X$ is a *set* (as opposed to being merely a subclass of X) and there are altogether finitely many free variables occurring in the formulas in Θ; and

(iii) X contains $\exists w\phi(x/w)$ and $\forall w\phi(x/w)$ whenever x is a variable actually occurring as a free variable in $\phi \in X$ and w is a bound variable not occurring in ϕ.

The requirement in (iii) is only for the sake of convenience. We can circumvent this restriction by considering e.g. $\exists x(\phi \wedge x \approx x)$ instead of $\exists x\phi$.

We will usually suppress our distinction between the two classes of variables, namely the free and bound variables. Accordingly, we refer to $\exists x\phi$, x being a free variable in ϕ, meaning $\exists w\phi(x/w)$.

A few more remarks on the formalities of quantification. When forming $\exists x\phi$, i.e. $\exists w\phi(x/w)$, it is irrelevant what w we use as long as w does not occur in ϕ. Taking now a *free* variable y not occurring in ϕ, we can form $\phi' = \phi(x/y)$. Since $\phi'(y/w) = \phi(x/w)$, what we denote by $\exists y\phi'(y)$, i.e. $\exists w\phi'(y/w)$, becomes identified with $\exists x\phi$. This has the consequence that whenever a finite sequence $\vec{x}$ of variables and a formula of the form $\exists x\phi$ is given, we can always assume that x is not among the $\vec{x}$; namely, we can pass to the form $\exists y\phi'$ with such a y. Briefly put, we do not distinguish between *alphabetic variants* of formulas which differ only in the exact identity of bound variables (but in which the same pairs of occurrences of bound variables are occupied by equal bound variables).

The usual way of writing $\phi \wedge \psi$ for $\bigwedge\{\phi, \psi\}$ and $\phi \vee \psi$ for $\bigvee\{\phi, \psi\}$ will be adopted. The empty conjunction $\bigwedge \varnothing$ will be denoted by $\top$ ("true"), the empty disjunction $\bigvee \varnothing$ by $\bot$ ("false").

The formula ϕ is a *subformula* of the formula ψ if "ϕ is constructed as an intermediate step in the construction of ψ". More precisely, we have the following inductive definition. The only subformula of an atomic formula is itself. The subformulas of $\bigwedge \Theta$ are: $\bigwedge \Theta$ itself as well as the subformulas of all the elements of Θ; similarly for $\bigvee \Theta$. The subformulas of $\neg\phi$ are $\neg\phi$ itself as well as all the subformulas of ϕ. There is a similar clause for each of $\phi \to \psi$, $\exists x\phi$ and $\forall x\phi$.

A subclass F of the class of all formulas of $L_{\infty\omega}$ is called a *fragment* of $L_{\infty\omega}$ if (a) with each formula $\phi \in F$ all the subformulas of ϕ also belong to F and (b) F is closed under substitution: if ϕ is in F, t is a term of L, x is a free variable in ϕ, then $\phi(x/t)$ is in F.

A *(Gentzen) sequent* of F is an object of the form $\Phi \Rightarrow \Psi$ (with $\Rightarrow$ a new symbol) where Φ and Ψ are finite (possible empty) sets of formulas belonging to F. A *theory* in F is a set of sequents of F.

We notice that the intersection of fragments is again a fragment and this allows us to speak of the fragment generated by a set of formulas.

The set of finitary formulas of L is denoted by $L_{\omega\omega}$ and it is the fragment F of $L_{\infty\omega}$ such that each of $\bigwedge \Theta \in F$ and $\bigvee \Theta \in F$ implies that Θ is a finite set; briefly, only finite conjunctions and disjunctions are allowed. $L_{\omega_1\omega}$ is the fragment where only *countable* (possibly finite) conjunctions and disjunctions are allowed.

Another kind of restriction leads to other fragments that are important for us. The *coherent logic* $L^g_{\infty\omega}$ is the fragment in which we have unrestricted use of $\bigvee$ and $\exists$, $\bigwedge$ can be applied only to finite sets and the rest of the logical operators cannot be used at all. In other words, the formulas of $L^g_{\infty\omega}$ form the least class X containing the atomic formulas such that if Θ is a subset of X, Σ is a finite subset of X, $\phi \in X$, and x is free in ϕ, then $\bigvee \Theta \in X$, $\bigwedge \Sigma \in X$ and $\exists x\phi \in X$. The fragments $L^g_{\omega\omega}$, $L^g_{\omega_1\omega}$ are defined naturally as $L^g_{\omega\omega} = L_{\omega\omega} \cap L^g_{\infty\omega}$, $L^g_{\omega_1\omega} = L_{\omega_1\omega} \cap L^g_{\infty\omega}$. A theory in $L^g_{\infty\omega}$ is called *coherent*, one in $L^g_{\omega\omega}$ *finitary coherent*.

The primary meaning of formulas is given by their *standard interpretation* in (ordinary) structures. A (many-sorted) *structure* M of type L is a function with domain L subject to the following conditions:

(1) for every sort s in L, $M(s)$ is a set;

(2) for every predicate symbol R in L, $R \subset s_1 \times \cdots \times s_n$, $M(R)$ is a subset of $M(s_1) \times \cdots \times M(s_n)$;

(3) for every operation symbol f in L such that $f : s_1 \times \cdots \times s_n \to s$, $M(f)$ is an operation $M(s_1) \times \cdots \times M(s_n) \to M(s)$. In particular, if f is an individual constant of sort s, $M(f) \in M(s)$.

An important point is that we allow the *(partial) domains* $M(s)$ of M to be empty. In model theory, usually the domains are stipulated to be non-empty. This difference slightly affects what sequents are considered logically valid; c.f. below.

The basic notion is that of the truth of a formula in a structure, once free variables have been interpreted by fixed but arbitrary elements in the structure. Let ϕ be a formula, with its free variables among $x_1, \ldots, x_n$, x_i of sort s_i, and let $a_i \in M(s_i)$. Then we write

$$\begin{aligned} M &\models \phi[a_1, \ldots, a_n] \quad \text{or} \\ M &\models \phi[x_1/a_1, \ldots, x_n/a_n] \quad \text{or} \\ M &\models \phi[\vec{a}] \end{aligned}$$

for: ϕ is true in M when x_i is interpreted as a_i, or: the a_i satisfy ϕ in M. The notion of truth has a straightforward inductive definition, suggested by the terminology introduced above relating to formulas. In particular, e.g. we have

$$\begin{aligned} M &\models (\textstyle\bigwedge\Theta)\,[\vec{a}] \;\Leftrightarrow\; \text{for every } \phi \in \Theta, M \models \phi[\vec{a}]. \\ M &\models (\textstyle\bigvee\Theta)\,[\vec{a}] \;\Leftrightarrow\; \text{for at least one } \phi \in \Theta, M \models \phi[\vec{a}]. \\ M &\models (\exists x\phi)[\vec{a}] \;\Leftrightarrow\; \text{for some } a \in M(x), M \models \phi[x/a, \vec{a}] \end{aligned}$$

(here $M(x) = M(s)$ where s is the sort of x). Also, it is important to keep in mind that equality ($\approx$) is always interpreted by real equality. Formally,

$$M \models (t_1 \approx t_2)[\vec{a}] \;\text{ iff }\; t_1^M[\vec{a}] = t_2^M[\vec{a}],$$

where $t_1^M[\vec{a}]$ is the value of the interpretation of the term t when the free variable x_i is assigned the value a_i ($i = 1, \ldots, n$).

§2 Some categorical notions

Here we briefly enumerate the handful of simple categorical notions that the interpretations of formulas in categories rests on.

All categories in this work are assumed to have finite left limits, i.e., the left limit, or inverse limit, of every finite diagram in the category should exist. Left limits are determined only up to a unique isomorphism *over* the given diagram, in the well-known sense. It is also well known that it is enough to assume the existence of certain finite left limits in order to have all finite left limits, viz. the existence of

(1) a final object 1 (= empty product);

(2) the product of any two objects;

(3) the equalizer of two morphisms with the same domain and with the same codomain.

We use the following standard notation concerning products. Let a product

$$\begin{tikzcd} & X_1 \times \cdots \times X_n \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_n"] & \\ X_1 & \cdots\cdots & X_n \end{tikzcd}$$

be given. If $A \xrightarrow{f_i} X$ is a morphism for each $i = 1, \ldots, n$, then $\langle f_1, \ldots, f_n\rangle$, or $\langle f_1, \ldots, f_n\rangle_{\pi_1,\ldots,\pi_n}$, or $\langle f_1, \ldots, f_n\rangle_{\pi}$, will denote the unique morphism $f : A \to X_1 \times \cdots \times X_n$ such that $\pi_i \circ f = f_i$, $i = 1, \ldots, n$.

We fix a category with finite left limits for the rest of this section.

For a given object X, a subobject of X is determined by a monomorphism $A \hookrightarrow X$ and two monomorphisms $A \hookrightarrow X$, $B \hookrightarrow X$ determine the same subobject of X if there are morphisms $A \rightleftarrows B$ such that both

$$\begin{tikzcd} A \arrow[r, hook] \arrow[dr] & X \\ & B \arrow[u, hook] \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} B \arrow[r, hook] \arrow[dr] & X \\ & A \arrow[u, hook] \end{tikzcd}$$

commute. We talk about *the* subobject $A \hookrightarrow X$, with a certain measure of abuse of language. We say that the subobject $A \hookrightarrow X$ is *smaller than* ($\leq$) $B \hookrightarrow X$ if there is a (necessarily unique) morphism $A \to B$ such that

$$\begin{tikzcd} A \arrow[r, hook] \arrow[dr] & X \\ & B \arrow[u, hook] \end{tikzcd} \quad \text{commutes.}$$

The $\leq$-relation partially orders the set of subobjects of X. As a consequence of the existence of left limits, we have that in this partial order the inf (or: meet) $A \wedge B$ of two elements (i.e., the greatest element C such that $C \leq A$, $C \leq B$) exists; in fact, it is given by $C \hookrightarrow X$ in the pullback

$$\begin{tikzcd} A \arrow[r, hook] & X \\ C \arrow[u, hook] \arrow[r, hook] \arrow[ur, phantom, "\text{p.b.}"] & B. \arrow[u, hook] \end{tikzcd}$$

Given an arbitrary set Θ of subobjects of a given object X, the inf of Θ, denoted by $\bigwedge \Theta$, is the greatest subobject of X that is $\leq$ than any element of Θ. For an infinite set Θ, $\bigwedge \Theta$ does not necessarily exists. For the empty set $\Theta = \varnothing$, $\bigwedge \varnothing$ equals the maximal subobject $X \xhookrightarrow{\text{id}} X$.

Given a set Θ of subobjects of X, $\bigvee \Theta$, the sup of Θ, is the smallest subobject (if it exists) among those that are $\geq$ than any subobject in Θ. $\bigvee\{A, B\}$ is denoted $A \vee B$. The phrase "$\mathcal{R}$ has finite sups" means that for any finite family (including the empty one) of subobjects of a given object in $\mathcal{R}$, the sup of the family exists.

A morphism $A \to Y$ is called *surjective* if *whenever* $B \to Y$ is a monomorphism such that $A \to Y$ factors through $B \to Y$:

$$\begin{tikzcd} A \arrow[r] \arrow[dr] & Y \arrow[d, phantom, ""] \\ & B \arrow[u, hook] \end{tikzcd}$$

then $B \to Y$ as a subobject of Y is the maximal subobject, i.e., $B \to Y$ is an isomorphism. A surjective morphism is always an epimorphism but not necessarily conversely. To show the first claim, assume that f is surjective and in

$$A \xrightarrow{f} B \underset{h_2}{\overset{h_1}{\rightrightarrows}} C,$$

we have $h_1 \circ f = h_2 \circ f$. Let $\mathrm{Equ}(h_1, h_2) \to B$ be the equalizer of h_1 and h_2. By the universal property of the equalizer, $A \xrightarrow{f} B$ factors through $\mathrm{Equ}(h_1, h_2) \to B$. By the definition of surjectivity, this implies that $\mathrm{Equ}(h_1, h_2) \to B$ is an isomorphism, $\mathrm{Equ}(h_1, h_2) \xrightarrow{\simeq} B$. But of course, this means that $h_1 = h_2$, as required. Under reasonable assumptions, a morphism is surjective iff it is an "effective epimorphism" in the sense of SGA4, I. 10.3; we return to this point in the next Chapter.

Given a diagram

$$\begin{tikzcd} A \arrow[r, hook] & X \arrow[d, "f"] \\ & Y \end{tikzcd}$$

the *image of* $A \hookrightarrow X$ *under* f is the subobject $\exists_f(A) \hookrightarrow Y$ such that there is a surjective g making the following commute:

$$\begin{tikzcd} A \arrow[r, hook] \arrow[d, "g"'] & X \arrow[d, "f"] \\ \exists_f(A) \arrow[r, hook] & Y \end{tikzcd}$$

The image $\exists_f(A) \hookrightarrow Y$, if it exists, is uniquely determined as a subobject of Y.

It is easy to see that this definition is equivalent to the following: the image $\exists_f(A) \hookrightarrow Y$ of $A \hookrightarrow X$ is the smallest subobject $B \hookrightarrow Y$ such that $A \hookrightarrow X \xrightarrow{f} Y$ factors through $B \hookrightarrow Y$. Namely, suppose that in

$$\begin{tikzcd} A \arrow[r, hook] \arrow[d, "g"'] & X \arrow[d, "f"] \\ B \arrow[r, hook] & Y \end{tikzcd}$$

g is surjective and let $C \hookrightarrow Y$ be another subobject such that $A \hookrightarrow X \xrightarrow{f} Y$ factors through $C \hookrightarrow Y$; we want to show that there is $B \to C$ such that $B \to C \hookrightarrow Y$ is $B \hookrightarrow Y$. Consider the "intersection":

$$\begin{tikzcd} B \wedge C = B \times_Y C \arrow[r] \arrow[d, hook] \arrow[dr, phantom, "\text{p.b.}"] & C \arrow[d, hook] \\ B \arrow[r, hook] & Y. \end{tikzcd}$$

By the universal property of the pullback, we will have $A \to B \wedge C$ such that the following is commutative

$$\begin{tikzcd} A \arrow[dr] \arrow[drr, bend left] \arrow[ddr, bend right, "g"'] & & \\ & B \wedge C \arrow[r] \arrow[d, hook] & C \arrow[d, hook] \\ & B \arrow[r, hook] & Y \end{tikzcd}$$

By the surjectivity of g, $B \wedge C \to B$ is an isomorphism. Denoting by j^{-1} its inverse, $B \xrightarrow{j^{-1}} B \wedge C \to C$ is the desired morphism.

This shows that if B is the image $\exists_f(A)$ according to the first definition, then it is that according to the second definition as well. The converse is easier.

The phrase "$\mathcal{R}$ has images" means that for every subobject $A \hookrightarrow X$ and every morphism $X \to Y$, $\exists_f A \to Y$ exists. Notice that this is equivalent to saying: every morphism $A \xrightarrow{f} Y$ in $\mathcal{R}$ is the product ip of a surjective morphism p and a monomorphism i.

The foregoing notions suffice to interpret formulas in $L^g_{\infty\omega}$, which is our primary interest. Next we mention the rest of the notions that are used in interpreting the full language $L_{\infty\omega}$.

The *Boolean complement* of a subobject A of X, if it exists, is the subobject B of X such that $A \vee B = X$ and $A \wedge B = 0$. Here X is the maximal subobject $X \xrightarrow{\mathrm{id}} X$ and 0 is the minimal subobject of X, the sup of the empty family. Again, B as a subobject is uniquely determined if it exists at all, at least in case the subobject lattice of X is distributive (which will mostly be the case).

The *Heyting complement* of a subobject A of X, if it exists, is the maximal subobject B of X such that $A \wedge B = 0$. B is again uniquely determined; if the Boolean complement of A exists, then the Heyting complement equals the Boolean complement.

Intuitionistic implication is formulated in the notion of *Heyting implication* $A \to B$ of the two subobjects A and B of X. $A \to B$, if exists, is the maximal subobject C of X such that $A \wedge C \leq B$. $A \to 0$ is the Heyting complement of A.

There is a Boolean formulation of the universal quantifier, based on the identity $\forall x A x = \neg\exists x \neg A x$. We formulate an "intuitionistic" notion. Given

$$\begin{tikzcd} A \arrow[r, hook] & X \arrow[d, "f"] \\ & Y \end{tikzcd}$$

the *dual image* (for lack of a better expression) of the subobject $A \hookrightarrow X$, denoted $\forall_f(A)$, is the largest subobject $B \to Y$ such that the pullback $f^{-1}(B) \hookrightarrow X$ factors through $A \hookrightarrow X$.

$$\begin{tikzcd} A \arrow[r, hook] & X \arrow[dd, "f"] \\ f^{-1}(B) \arrow[u, dotted] \arrow[ur, hook] \arrow[d] & \\ B \arrow[r, hook] & Y \end{tikzcd}$$

Finally, let us mention an expression that is a common generalization of the Heyting $\to$ and $\forall$. Suppose we are given

$$\begin{tikzcd} A_1 \arrow[dr, hook] & \\ & X \arrow[d, "f"] \\ A_2 \arrow[ur, hook] & Y \end{tikzcd}$$

By $\forall_f(A_1 \to A_2)$ we mean the largest subobject $B \hookrightarrow Y$ of Y such that $(f^{-1}(B) \wedge A_1) \hookrightarrow X$ factors through $A_2 \hookrightarrow X$. Putting $f = \mathrm{id}_X$, we get Heyting implication, and putting $A_1 = X$, we get the dual image $\forall_f(A_2)$.

Finally, we make a few remarks on the above notions in the category of sets, **Set**. First of all, the reader should be familiar with the meaning of left limits in **Set**.

Two monomorphisms $A \xrightarrow{f} X$, $B \xrightarrow{g} X$ determine the same subobject of X just in case the images $f(A) \subset X$, $f(B) \subset X$ coincide. Thus, subobjects in **Set** mean *subsets*. The lattice (actually: complete Boolean algebra) structure of the subobjects of X is that of the subsets X determined by inclusion. In particular, inf and sup are intersection and union, respectively.

The image $\exists_f(A) \hookrightarrow Y$ in

$$\begin{tikzcd} A \arrow[r, hook] \arrow[d] & X \arrow[d] \\ \exists_f(A) \arrow[r, hook] & Y \end{tikzcd}$$

is nothing but the usual image of the subset $A \subset X$ under f. The dual image $\forall_f(A) \to Y$ in

$$\begin{tikzcd} A \arrow[r, hook] & X \arrow[d] \\ \forall_f(A) \arrow[r, hook] & Y \end{tikzcd}$$

is determined by $y \in \forall_f(A) \Leftrightarrow \forall x \in X[f(x) = y \to x \in A]$, as it is easily seen.

§3 The categorical interpretation

Let $\mathcal{R}$ be a fixed category with finite left limits. Let L be a language as described in §1. The notion of an *$\mathcal{R}$-valued structure* (or: *$\mathcal{R}$-structure*) *of type* L is a natural generalization of that of an ordinary structure.

An $\mathcal{R}$-structure M of type L is a function with domain L such that

(1) for every sort s in L, $M(s)$ is an object of $\mathcal{R}$;

(2) for every predicate symbol R in L, $R \subset s_1 \times \cdots \times s_n$, $M(R)$ is a subobject $M(R) \to M(s_1) \times \cdots \times M(s_n)$ in $\mathcal{R}$;

(3) for every operation symbol f in L such that $f : s_1 \times \cdots \times s_n \to s$, $M(f)$ is a morphism $M(f) : M(s_1) \times \cdots \times M(s_n) \to M(s)$ in $\mathcal{R}$. If f is an individual constant of sort s, $M(f)$ is a morphism $1 \to M(s)$.

Remark *There is a certain amount of ambiguity in the notion, e.g. because products are determined only up to a (unique) isomorphism. A more precise version would be something like this:*

(2′) *for* $R \subset s_1 \times \cdots \times s_n$, M *specifies a particular product*

$$\begin{tikzcd} & M(s_1) \times \cdots \times M(s_n) \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_n"] & \\ M(s_1) & \cdots\cdots & M(s_n) \end{tikzcd}$$

as well as a particular monomorphism $M(R) \hookrightarrow M(s_1) \times \cdots \times M(s_n)$*;*

(3′) *a similar version for operation symbols.*

However, we will not find it necessary to insists on the more precise version.

Notice that if $\mathcal{R}$ is the category of sets, an $\mathcal{R}$-structure is essentially what we called an ordinary structure before.

Next we turn to the interpretation of terms and formulas. Let M be an $\mathcal{R}$-structure of type L.

For a sequence $\vec{x} = (x_1, \ldots, x_n)$ of *distinct* variables of respective sorts $s_1, \ldots, s_n$, we define $M(\vec{x}) = M(s_1) \times \cdots \times M(s_n)$. For a term t of sort s, having all its free variables among $\vec{x} = (x_1, \ldots, x_n)$, $M_{\vec{x}}(t)$ will be defined and it will be a morphism $M(\vec{x}) \to M(s)$.

Let $t := x_i$. $M_{\vec{x}}(x_i)$ is defined as the canonical projection $M(\vec{x}) \xrightarrow{\pi_i} M(x_i)$

Let $t = ft_1 \cdots t_n$, t_i of sort s_i, t of sort s. Then $M_{\vec{x}}(t)$ is the composite indicted by the following diagram

$$
\begin{tikzcd}
& \prod_{i=1}^{n} M(s_i) \arrow[dl, "\pi_i"'] \arrow[r, "M(f)"] & M(s) \\
\begin{array}{c} M(s_i) \\ \vdots \\ (i=1,\ldots,n) \end{array} & & \\
& M(\vec{x}) \arrow[ul, "M_{\vec{x}}(t_i)"] \arrow[uu, "\langle M_{\vec{x}}(t_1),\ldots,M_{\vec{x}}(t_n)\rangle"'] \arrow[uur, "M_{\vec{x}}(t)"'] &
\end{tikzcd}
$$

Next we give the interpretation of formulas. As a general remark, we note the following. Let ϕ be a formula with its free variables among $\vec{x} = (x_1, \ldots, x_n)$. Then $M_{\vec{x}}(\phi)$, the interpretation of ϕ in M, will be a subobject of $M(\vec{x})$, provided $M_{\vec{x}}(\phi)$ is defined at all. $M_{\vec{x}}(\phi)$ will be defined if and only if all the categorical operations called for by the various logical operators can actually be performed in $\mathcal{R}$.

The interpretation $M_{\vec{x}}(t_1 \approx t_2)$ of the atomic formula $t_1 \approx t_2$ is given as the following *equalizer* (more precisely, the corresponding subobject of $M(\vec{x})$) where s is the common sort of t_1 and t_2:

$$
\begin{tikzcd}
M_{\vec{x}}(t_1 \approx t_2) \arrow[r, hook] & M(\vec{x}) \arrow[r, shift left, "M_{\vec{x}}(t_1)"] \arrow[r, shift right, "M_{\vec{x}}(t_2)"'] & M(s)
\end{tikzcd}
$$

Let $Pt_1 \cdots t_n$ be an atomic formula and let t_i be of sort s_i. Then $M_{\vec{x}}(Pt_1 \cdots t_n)$ is given by the following *pull back* diagram

$$
\begin{tikzcd}
\prod_{i=1}^{n} M(s_i) & M(P) \arrow[l, hook'] \\
M(\vec{x}) \arrow[u, "\langle M_{\vec{x}}(t_i)\rangle_{i=1}^{n}"] & M_{\vec{x}}(Pt_1 \cdots t_n) \arrow[l, hook'] \arrow[u]
\end{tikzcd}
$$

In the next few clauses, we will deal with subobjects of $M(\vec{x})$. We define

$$M_{\vec{x}}(\bigwedge\Theta) \underset{\mathrm{df}}{\equiv} \bigwedge\{M_{\vec{x}}(\theta) : \theta \in \Theta\}$$

$$M_{\vec{x}}(\bigvee\Theta) \underset{\mathrm{df}}{\equiv} \bigvee\{M_{\vec{x}}(\theta) : \theta \in \Theta\}$$

On the right hand side, $\bigwedge$ and $\bigvee$ mean the inf and sup operations on subobjects of $M(\vec{x})$. The interpretations $M_{\vec{x}}(\bigwedge\Theta)$, $M_{\vec{x}}(\bigvee\Theta)$ exist if and only if each $M_{\vec{x}}(\theta)$ exists ($\theta \in \Theta$) and the inf (sup) on the right hand side exists.

To define $M_{\vec{x}}(\exists y\phi)$, first note that without loss of generality we can assume that y is not among the $\vec{x}$. Let π be the canonical projection $M(\vec{x}, y) \to M(\vec{x})$. We define

$$M_{\vec{x}}(\exists y\phi) = \exists_\pi(M_{\vec{x},y}(\phi)).$$

The above is sufficient for the definition of the interpretation of $L^g_{\infty\omega}$. The following last clause takes care of the full logic $L_{\infty\omega}$:

$$M_{\vec{x}}(\forall\vec{y}(\phi \to \psi)) = \forall_\pi(M_{\vec{x},\vec{y}}(\phi) \to M_{\vec{x},\vec{y}}(\psi))$$

where π is the canonical proyection

$$\pi : M(\vec{x}\vec{y}) \to M(\vec{x})$$

for disjoint sequences $\vec{x}$ and $\vec{y}$ of variables.

We note that, in order to take care of $\neg$, $\rightarrow$, $\forall$ (which are of secondary interest to us anyway) at the same time, we adopt the convention that the formula $\forall\vec{y}(\phi \rightarrow \psi)$ is considered as built up directly from ϕ and ψ. In other words, $\phi \rightarrow \psi$ is *not* a subformula of $\forall\vec{y}(\phi \rightarrow \psi)$; its subforumlas are itself, ϕ, ψ and the subformulas of the latter. If $\forall\vec{y}$ is the empty sequence, we are essentially reduced to $\phi \rightarrow \psi$, and if ϕ is $\top$, we have $\forall\vec{y}\psi$. By this device, e.g. we will have to state only two rules in our formal system in Chapter 5, instead of four (or six), relating to $\rightarrow$, $\forall$ (and $\neg$).

In view of the meaning of the relevant operations in **Set**, it is clear that the categorical interpretation in **Set** reduces to the standard interpretation (c.f. §1). More precisely, if M is a **Set**-structure, it is easy to see that the subobject $M_{\vec{x}}(\phi) \hookrightarrow M(\vec{x})$, as a subset of $M(\vec{x}) = M(x_1) \times \cdots \times M(x_n)$, coincides with the set $\{\langle a_1, \ldots, a_n\rangle \in M(x_1) \times \cdots \times M(x_n); M \models \phi[a_1, \ldots, a_n]\}$.

Returning to interpretations in a general category, we note that the same ambiguity as in the notion of an $\mathcal{R}$-structure appears in the notion of an interpretation. The interpretation $M_{\vec{x}}(t)$ of a term is only given relative to a specification of the product $M(\vec{x}) = M(x_1)\times\cdots\times M(x_n)$. Once this product is given, the morphism $M_{\vec{x}}(t) : M(\vec{x}) \rightarrow M(s)$ is also given. More particularly, if we compute the interpretation $M_{\vec{x}}(t)$ relative to two different copies of the product

$$\begin{array}{ccc} \cdots \; X \; \cdots & & \cdots \; X' \; \cdots \\ \downarrow{\scriptstyle \pi_i} & \text{and} & \downarrow{\scriptstyle \pi'_i} \\ M(x_i) & & M(x_i) \\ {\scriptstyle (i=1,\ldots n)} & & {\scriptstyle (i=1,\ldots n)} \end{array}$$

obtaining $f : X \rightarrow M(s)$ and $f' : X' \rightarrow M(s)$, then the unique isomorphism $j : X' \xrightarrow{\simeq} X$ such that $\pi'_i = \pi_i \circ j$ $(i = 1, \ldots, n)$ will carry f into f': $f' = f \circ j$. The reader is invited to check this by going through the definition of $M_{\vec{x}}(t)$. We can make an analogous detailed statement to what extent interpretations of formulas are determined.

Given a sequent $\Phi \Rightarrow \Psi$, we say that the $\mathcal{R}$-structure M *satisfies* it, in symbols: $M \models \Phi \Rightarrow \Psi$ if

$$\bigwedge\{M_{\vec{x}}(\phi) : \phi \in \Phi\} \leq \bigvee\{M_{\vec{x}}(\psi) : \psi \in \Psi\}.$$

Here $\vec{x}$ is the sequence of all free variables in the sequent $\Phi \Rightarrow \Psi$ and $\bigwedge$, $\bigvee$ denote inf and sup on the subobjects of $M(\vec{x})$. It is understood that $M \models \Phi \Rightarrow \Psi$ implies, among other things, that $M_{\vec{x}}(\phi)$ are all defined for $\phi \in \Phi \cup \Psi$, moreover, the sup on the right hand side is defined.

In particular, with a sequent

$$\phi \Rightarrow \psi$$

with single formulas on both sides, $M \models \phi \Rightarrow \psi$ means that

$$M_{\vec{x}}(\phi) \leq M_{\vec{x}}(\psi)$$

where $\vec{x}$ is the sequence of variables occurring free either in ϕ of in ψ.

§4. Expressing categorical notions by formulas: the first main fact

In this section, we will show that certain simple properties of diagrams (e.g., that it is a product diagram, etc.) can be expressed by the truth of certain (Gentzen) sequents. This fact has a rather tautologous nature; nevertheless one has to do some work to establish it in the form needed. At the end of the next chapter, there is a discussion how this so called 'first main fact' (namely, that such an expression is possible) combines with the 'second main fact' (c.f. the next chapter) to give a way of applying logic to categories.

Let $\mathcal{R}$ be a given category with finite left limits, fixed throughout this section. There is a canonical language associated with $\mathcal{R}$ as well as a canonical interpretation of this language, as follows.

Define the language $L = L_{\mathcal{R}}$ by declaring that the sorts in L are exactly the objects of $\mathcal{R}$, that each morphism $f : A \to B$ in $\mathcal{R}$ is a unary operation symbol, with sort A associated to its only place and sort B "associated to its value", and that there are no other symbols in L. In other words, L is obtained by *forgetting* the composition law of the category $\mathcal{R}$, but retaining the domain-codomain-relationships. Thus L is "$\mathcal{R}$ itself". Consequently, the "identity" map on L is an $\mathcal{R}$-structure of type L. If we call this 'identical interpretation' M, then we can e.g. talk about $M_{\vec{x}}(\phi)$, for a formula ϕ. The meaning of the interpretation $M_{\vec{x}}(\phi)$ will be the result of reading the formula ϕ with the symbols understood as objects and morphisms in $\mathcal{R}$, and the logical operators understood as operations in $\mathcal{R}$. We will write $[\phi]_{\vec{x}}$ for $M_{\vec{x}}(\phi)$, or simply $[\phi]$, if $\vec{x}$ are exactly the free variables in ϕ, with M the canonical interpretation. We also use the notation $[t]$, or $[t]_{\vec{x}}$, for the interpretation of terms in the canonical structure.

To give an example, consider the formula $f(g(x)) \approx h(x)$ in the canonical language L. First of all, this is syntactically well-formed, i.e. it is a formula to begin with, if and only if f, g and h are morphisms with domains and codomains as shown:

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

and also, x has to be a variable of sort X. Next, $[f(g(x))]$ turns out to be the composite $f \circ g : X \to Z$. The interpretation of the formula, $[f(g(x)) \approx h(x)]$, is the equalizer of the two morphisms $f \circ g$ and h:

$$[f(g(x)) \approx h(x)] \hookrightarrow X \overset{f\circ g}{\underset{h}{\rightrightarrows}} Z.$$

Finally, the sequent $\Rightarrow f(g(x)) \approx h(x)$ (with empty left hand side) is *true* in the canonical interpretation M, or: *true in* $\mathcal{R}$ as we might say it, if and only if the diagram above is commutative. This latter fact is equivalent to saying that $h = f \circ g$ iff the equalizer of $f \circ g$ and h is id_X. This last fact is well-known and easily seen.

This example shows how categorical facts are expressed, indeed quite naturally, by the truth of sequents of the canonical language in $\mathcal{R}$ (i.e., in the canonical interpretation). Below there will be more of this kind. Next, we make a couple of simple remarks.

The subobject $[f(x) \approx y]$ of $X \times Y$, for a morphism $X \xrightarrow{f} Y$, should naturally be called the *graph* of f. Although the literal definition of $[f(x) \approx y]$ coming from Section 3 is something a bit more complicated, it is easy to see that $[f(x) \approx y]$ is the same thing as the subobject

$$X \xhookrightarrow{\langle \mathrm{id}_X, f\rangle} X \times Y$$

where we use the notation $\langle \cdot, \cdot \rangle$ in the way described in Section 2. Hence, to say that a given subobject given by the monomorphism $R \hookrightarrow X \times Y$ is identical to the graph of

f, is equivalent to saying that there is an isomorphism $X \overset{\simeq}{\longleftrightarrow} R$, making the following commute:

$$\begin{tikzcd} X \arrow[r, hook, "\langle \mathrm{id}_X, f\rangle"] \arrow[d, leftrightarrow, "\simeq"'] & X \times Y \\ R \arrow[ur, hook] & \end{tikzcd}$$

Now, let $A \overset{f}{\hookrightarrow} X$ be a monomorphism. We claim that the *subobject* $A \overset{f}{\hookrightarrow} X$ is denoted by the formula $\exists a(f(a) \approx x)$. Consider the following diagram

$$\begin{tikzcd} A \arrow[r, hook, "\langle \mathrm{id}_A, f\rangle"] \arrow[d, "g"'] \arrow[dd, dotted, bend right] & A \times X \arrow[d, "\text{can.}"] \\ B \arrow[r, hook] & X \\ A \arrow[u, dotted] \arrow[ur, hook, "f"'] & \end{tikzcd}$$

The claim is equivalent to saying that (i) there is $A \dashrightarrow A$ making the outer quadrangle commute and (ii) whenever $B \hookrightarrow X$ is a monomorphism such that there is $A \xrightarrow{g} B$ with the inner quadrangle commuting, then there is $A \dashrightarrow B$ such that $\begin{tikzcd} B \arrow[r, hook] & X \\ A \arrow[u, dotted] \arrow[ur, hook, "f"'] & \end{tikzcd}$

commutes. Although the whole claim is trivial, let us see why it is true.

$$\begin{tikzcd} A \arrow[r, "\langle \mathrm{id}_A, f\rangle"] \arrow[dr, "f"'] & A \times X \arrow[d, "\text{can.}"] \\ & X \end{tikzcd}$$

By definition, the composition $A \xrightarrow{\langle \mathrm{id}_A, f\rangle} A \times X \xrightarrow{\text{can.}} X$ is $A \xrightarrow{f} X$. So clearly, in (i) we can take $A \xrightarrow{\mathrm{id}_A} A$. But for the same reason, if we have $A \xrightarrow{g} B$ as in (ii), then

$$\begin{tikzcd} A \arrow[d, "g"'] \arrow[dr, "f"] & \\ B \arrow[r, hook] & X \end{tikzcd}$$

is commutative, so we can take g for $A \dashrightarrow B$. Notice that we have shown that the image needed for $[\exists a(f(a) \approx x)]$ to be defined indeed always exists whenever f is a monomorphism.

The last remark gives a way of denoting subobjects in the canonical language. It is also possible to simply extend the canonical language to include a symbol for one or more subobjects in $\mathcal{R}$. E.g., given a subobject $R \hookrightarrow X \times Y$, we can introduce a new symbol $\utilde{R}$, or just R, declared to be a binary relation symbol with its places assigned the sorts X and Y. Then, in the canonical interpretation, $[\utilde{R}xy]$ is by definition the subobject R and we can use $\utilde{R}$ in building compound formulas.

Similarly, we can extend the canonical language to include an n-ary operation symbol, corresponding to a morphism

$$f : X_1 \times \cdots \times X_n \to Y.$$

Denoting the operation symbol also by f, f has the obvious sorting. The same morphism $f : X \to Y$ could correspond to more distinct operations symbols, depending on how X is considered to be, a product $X_1 \times \cdots \times X_n$.

The *extended canonical language* corresponding to a given category contains all possible predicate and operation symbols described above. The extended canonical language has the obvious canonical interpretation in the category itself.

When we talk about the canonical language, unless otherwise indicated, usually we understand it to be the narrower sense, i.e. having only unary operation symbols and no predicate symbols.

We mention another example of interpreting formulas that we will have occasion to use. Consider

$$\begin{tikzcd} A_1 \arrow[r, "f_1"] & B \\ & A_2 \arrow[u, "f_2"'] \end{tikzcd}$$

and the formula $f_1 a_1 \approx f_2 a_2$. As it is expected, and easy to see, the interpretation $R = [f_1 a_1 \approx f_2 a_2] \overset{i}{\hookrightarrow} A_1 \times A_2$ is the pullback (fibered product) $A_1 \times_B A_2$. More precisely

$$\begin{tikzcd} A_1 \arrow[r, "f_1"] & B \\ R \arrow[u, "\pi_1 i"] \arrow[r, "\pi_2 i"'] & A_2 \arrow[u, "f_2"'] \end{tikzcd}$$

is a pullback diagram, where π_1, π_2 are the canonical projections $A_1 \xleftarrow{\pi_1} A_1 \times A_2 \xrightarrow{\pi_2} A_2$.

Returning to graphs of morphisms, we claim

Proposition 2.4.1 *If $R \hookrightarrow X \times Y$ is the graph of* some *morphism $X \to Y$, then the following two sequents are true in $\mathcal{R}$:*

$$\begin{aligned} Rxy \wedge Rxy' \quad &\Rightarrow \quad y \approx y', \\ &\Rightarrow \quad \exists y\ Rxy. \end{aligned}$$

Then omit the easy proof. Note that these two sequents are natural expressions of the notion of a functional relation. Our first main aim is to establish a converse of 2.4.1, c.f. 2.4.4 below.

The following lemma is completely trivial on the basis of the definitions.

Lemma 2.4.2 *If f is a monomorphism and it is surjective, then it is an isomorphism.*

The following lemma takes more work.

Lemma 2.4.3 *Suppose $R \hookrightarrow X \times Y$ is univalent, i.e. the first of the two sequents in* 2.4.1 *is true in $\mathcal{R}$. The the composite*

$$p_X i : R \xrightarrow{i} X \times Y \xrightarrow{p_X} X,$$

with p_X the canonical projection, is a monomorphism.

Proof. We will first spell out the hypothesis in a diagrammatical way. Let us introduce the canonical projections in the following product diagrams:

$$\begin{tikzcd}[row sep=small] & X \\ X \times Y \arrow[ur, "p_X"] \arrow[dr, "p_Y"'] & \\ & Y \end{tikzcd} \qquad \begin{tikzcd}[row sep=small] & R \\ R \times Y \arrow[ur, "q_R"] \arrow[dr, "q_Y"'] & \\ & Y \end{tikzcd} \qquad \begin{tikzcd}[row sep=small] & X \\ X \times Y \times Y \arrow[ur, "\rho_1"] \arrow[r, "\rho_2"] \arrow[dr, "\rho_3"'] & Y \\ & Y \end{tikzcd}$$

$$p = (p_X, p_Y) \qquad q = (q_R, q_Y) \qquad \rho = (\rho_1, \rho_2, \rho_3).$$

We first note that the interpretations $[Rxy]_{xyy'}$, $[Rxy']_{xyy'}$ can be identified as follows:

$$[Rxy]_{xyy'} : R \times Y \xhookrightarrow{j_1 = \langle p_X i q_R, p_Y i q_R, q_Y \rangle_\rho} X \times Y \times Y,$$

$$[Rxy']_{xyy'} : R \times Y \xhookrightarrow{j_2 = \langle p_X i q_R, q_Y, p_Y i q_R \rangle_\rho} X \times Y \times Y.$$

Also, $[y \approx y']_{xyy'}$ is:

$$X \times Y \xhookrightarrow{j = \langle p_X, p_Y, p_Y \rangle_\rho} X \times Y \times Y.$$

Using the interpretation of $\wedge$ as a pullback, and using that $Rxy \wedge Rxy' \Rightarrow y \approx y'$ holds, we get the commutative diagram

$$\begin{tikzcd} & X \times Y \times Y & \\ R \times Y \arrow[ur, "j_1"] & X \times Y \arrow[u, "j"] & R \times Y \arrow[ul, "j_2"'] \\ & P = [Rxy \wedge Rxy']_{xyy'} \arrow[ul] \arrow[u, dotted] \arrow[ur] & \end{tikzcd}$$

such that the outer diagram is a pullback.

Let $S \overset{f_1}{\underset{f_2}{\rightrightarrows}} R$ be two morphisms such that $p_X i f_1 = p_X i f_2$; we have to show that $f_1 = f_2$. Consider the morphisms

$$S \xrightarrow{h_1 = \langle f_1, p_Y i f_2 \rangle_q} R \times Y,$$

$$S \xrightarrow{h_2 = \langle f_2, p_Y i f_1 \rangle_q} R \times Y.$$

(Assuming that we are in **Set**, i is an inclusion, and the products are standard Cartesian products, we will have that h_1 is the map $s \mapsto \langle x, y_2, y_2 \rangle$ and h_2 is $s \mapsto \langle x, y_2, y_1 \rangle$ where $f_1(s) = \langle x, y_1 \rangle$, and $f_2(s) = \langle x, y_2 \rangle$.)

We can check easily that

$$j_1 h_1 = \langle p_X i f_1, p_Y i f_1, p_Y i f_2 \rangle_\rho \tag{2}$$

$$j_2 h_2 = \langle p_X i f_2, p_Y i f_1, p_Y i f_2 \rangle_\rho \tag{3}$$

By the hypothesis on f_1 and f_2, we have $j_1 h_1 = j_2 h_2$, hence by the universality of the pullback we have the following commutative diagram:

$$\begin{tikzcd}
 & X \times Y \times Y & \\
R \times Y \arrow[ur, "j_1"] & X \times Y \arrow[u, "j"] & R \times Y \arrow[ul, "j_2"'] \\
 & P \arrow[ul] \arrow[u, dotted] \arrow[ur] & \\
 & S \arrow[uul, "h_1"] \arrow[u, dotted] \arrow[uur, "h_2"'] &
\end{tikzcd}$$

Let the composite $S \to P \to X \times Y$ be $g = \langle g_1, g_2 \rangle_p$. Then $jg = \langle p_X, p_Y, p_Y \rangle_\rho \cdot \langle g_1, g_2 \rangle_p = \langle g_1, g_2, g_2 \rangle_\rho$. Since $jg = j_1 h_1 = j_2 h_2$, from (2) and (3) we obtain

$$\begin{array}{rcl} p_Y i f_1 & = & p_Y i f_2, \\ p_X i f_1 & = & p_X i f_2. \end{array}$$

By the universal property of the product $X \xleftarrow{p_X} X \times Y \xrightarrow{p_Y} Y$, we obtain that $if_1 = if_2$, hence $f_1 = f_2$ (since i is a monomorphism). $\square$

Let us call $R \hookrightarrow X \times Y$ a *functional relation ("with domain X and codomain Y")* if the two sequents in 2.4.1 are true (in $\mathcal{R}$).

Theorem 2.4.4 *Every functional relation $R \overset{i}{\hookrightarrow} X \times Y$ is the graph of a unique morphism $X \xrightarrow{f} Y$.*

Proof. Consider

$$\begin{tikzcd}
R \arrow[r, "i"] \arrow[dr, "p_X i"'] & X \times Y \arrow[r, "p_Y"] \arrow[d, "p_X"] & Y \\
 & X \arrow[ur, "f = p_Y i (p_X i)^{-1}"'] &
\end{tikzcd}$$

By 2.4.3, $p_X i$ is a monomorphism. By the truth of $\Rightarrow \exists y \; Rxy$, we clearly have that $p_X i$ is surjective. Hence, by 2.4.2, $p_X i$ is an isomorphism. Define $f = p_Y i (p_X i)^{-1}$. Then $X \xrightarrow{i(p_X i)^{-1}} X \times Y$ is identical to $\langle \mathrm{id}_X, f \rangle$, and we have the commutative diagram

$$\begin{tikzcd}
X \arrow[r, hook, "\langle \mathrm{id}_X, f \rangle"] & X \times Y \\
R \arrow[u, "p_X i", "\simeq"'] \arrow[ur, hook, "i"'] &
\end{tikzcd}$$

Hence, R is the graph of f, We leave it to the reader to check the uniqueness of f. $\square$

We make a remark on a slight variant of the last theorem. Using the same notation as there, assume that X is a product $X_1 \times \cdots \times X_n$. Then a subobject $R \to X \times Y$ can be regarded as a subobject of $X_1 \times \cdots \times X_n \times Y$. The functionality axioms can be written

$$Rx_1 \cdots x_n y \wedge Rx_1 \cdots x_n y' \Rightarrow y \approx y',$$

$$\Rightarrow \exists y \; Rx_1 \cdots x_n y;$$

it is immediate that indeed these two axioms are equivalent to the original two when R was regarded binary. So, we have that under the assumption that the last two sequents hold in $\mathcal{R}$, there is a morphism $f : X_1 \times \cdots \times X_n \to Y$ whose graph is $R \to (X_1 \times \cdots \times X_n) \times Y$.

Next, we formulate and prove our second main aim in this section, which is "the first main fact" in relating logic and categories. We give a list of sets of one or more axioms whose meaning will be immediately obvious when understood in **Set**. Their expected roles in the case of an arbitrary category $\mathcal{R}$ with finite limits is stated in Theorem 2.4.5. Under each heading, first we exhibit a diagram, then we list a few axioms using objects and morphisms of the diagram and finally, we state the "intended meaning". Lower case letters $a, b, c, d, e, \ldots$ denote variables of the sort denoted by the corresponding upper case letter.

1. Axiom for identity

$$A \xrightarrow{f} A;$$
$$\Rightarrow fa \approx a;$$
$$f = \mathrm{id}_A.$$

2. Axiom for commutative diagram

$$\begin{tikzcd} & B \arrow[dr, "g"] & \\ A \arrow[ur, "f"] \arrow[rr, "h"'] & & C \end{tikzcd}$$
$$\Rightarrow gfa \approx ha;$$
$$h = g \circ f.$$

3. Axiom for monomorphism

$$A \xrightarrow{f} A;$$
$$fa \approx fa' \Rightarrow a \approx a';$$
$$f \text{ is a monomorphism.}$$

4. Axioms for terminal object

$$A;$$
$$\Rightarrow a \approx a',$$
$$\Rightarrow \exists a(a \approx a);$$
$$A \text{ is a terminal object.}$$

5. Axioms for equalizer

$$\begin{tikzcd} E \arrow[r, "\varepsilon"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B; \end{tikzcd}$$
$$\varepsilon e \approx \varepsilon e' \Rightarrow e \approx e',$$
$$\Rightarrow f\varepsilon e \approx g\varepsilon e,$$
$$fa \approx ga \Rightarrow \exists e(\varepsilon e \approx a);$$
$$\varepsilon \text{ is the equalizer of } f \text{ and } g.$$

6. Axioms for product

$$\begin{tikzcd} & C \arrow[dl, "f"'] \arrow[dr, "g"] & \\ A & & B; \end{tikzcd}$$

$$fc \approx fc' \wedge gc \approx gc' \Rightarrow c \approx c',$$

$$\Rightarrow \exists c(fc \approx a \wedge gc \approx b);$$

C is the product of A and B, with projections f and g.

7. Axiom for initial object

$$A;$$

$$a \approx a \Rightarrow ;$$

A is an initial object.

8. Axioms for sup

$$\begin{tikzcd} & A_i \arrow[d, hook, "f_i"] & \\ B \arrow[r, hook, "g"'] & X & (i \in I); \end{tikzcd}$$

$$\textstyle\bigvee_{i \in I} \underset{\sim}{A}_i(x) \Rightarrow \underset{\sim}{B}(x),$$

$$\underset{\sim}{B}(x) \Rightarrow \textstyle\bigvee_{i \in I} \underset{\sim}{A}_i(x);$$

(**Remark** *Here* $\underset{\sim}{B}(x)$ *denotes* $\exists b(gb \approx x)$ *as we introduced above; similarly for* $\underset{\sim}{A}_i(x)$.)

$B \overset{g}{\hookrightarrow} X$ is the sup of the $A_i \overset{f_i}{\hookrightarrow} X$.

9. Axioms for image

$$A \xrightarrow{f} B;$$

$$\Rightarrow \exists a(fa \approx b);$$

f is surjective.

10. Axioms for inf

$$\begin{tikzcd} & A_i \arrow[d, hook, "f_i"] & \\ B \arrow[r, hook, "g"'] & X & (i \in I); \end{tikzcd}$$

$$\underset{\sim}{B}(x) \Rightarrow \textstyle\bigwedge_{i \in I} \underset{\sim}{A}_i(x),$$

$$\textstyle\bigwedge_{i \in I} \underset{\sim}{A}_i(x) \Rightarrow \underset{\sim}{B}(x);$$

c.f. the remark under item no. 8

$B \overset{g}{\hookrightarrow} X$ is the inf of the $A_i \overset{f_i}{\hookrightarrow} X$.

11. Axioms for dual image

$$\begin{tikzcd}
A_1 \arrow[dr, hook, "f_1"] & \\
& X \arrow[d, "f"] \\
A_2 \arrow[ur, hook, "f_2"'] & \\
B \arrow[r, hook, "g"'] & Y
\end{tikzcd}$$

$$\underset{\sim}{B}(y) \Leftrightarrow \forall x[(fx \approx y \wedge \underset{\sim}{A}_1(x)) \rightarrow \underset{\sim}{A}_2(x)].$$

Remark $\Phi \Leftrightarrow \Psi$ *is an abbreviation for two sequents jointly:* $\Phi \Rightarrow \Psi$ *and* $\Psi \Rightarrow \Phi$.

Theorem 2.4.5 *The above axioms express their intended meaning. That is, given $\mathcal{R}$, an arbitrary category with finite left limits, if given a diagram in $\mathcal{R}$ as indicated under any one of the headings* 1-11*, the diagram satisfies the condition stated last under the heading* if and only if *all of the axioms under the heading are true in $\mathcal{R}$.*

Proofs. ad 6, products: (i) Assume first that the two axioms hold in the category. Consider the subobject $R = [fc \approx a \wedge gc \approx b] \rightarrow A \times B \times C$ and, via the canonical isomorphism $A \times B \times C \simeq (A \times B) \times C$, consider R as a subobject of $D \times C$, $D = A \times B$. The canonical projections $\pi_A : A \times B \rightarrow A$, $\pi_B : A \times B \rightarrow B$ are introduced. We first claim that R is a functional subobject of $D \times C$, "with domain D and codomain C". Making use of the remark after 2.4.4, the claim means that the sequents

$$\begin{aligned} Rabc \wedge Rabc' &\Rightarrow c \approx c' \\ &\Rightarrow \exists c\; Rabc \end{aligned}$$

are both true in $\mathcal{R}$. It turns out that this fact is equivalent to the assumption that the two "axioms for product" hold. The second axiom $\Rightarrow \exists c\; Rabc$ is actually identical to the second axiom $\Rightarrow \exists c(fc \approx a \wedge gc \approx b)$ "for products". The equivalence of the two forms of the first axiom could be verified by a straightforward computation right at this point but we prefer deferring it to the next chapter.

Granting the functoriality of R, we have a morphism $A \times B \xrightarrow{h} C$ whose graph is R.

Let h' be the morphism $C \xrightarrow{h'} A \times B$ resulting from the universal property of the product $A \times B$, such that $\pi_A \circ h' = f$, $\pi_B \circ h' = g$. The next thing to realize is that $R \rightarrow (A \times B) \times C$ is also the graph of the morphism h', *now with domain C and codomain* $A \times B$, i.e., in the opposite sense to h; and this holds without any assumption on f and g. This is an easy exercise in first definitions.

Next, we state a general fact. Let $R \hookrightarrow D \times C$ be a subobject of a product $D \times C$. If R is the graph of $D \xrightarrow{h} C$ and in the opposite sense, also of $C \xrightarrow{h'} D$, then h and h' are inverses of each other, hence h, h' are isomorphisms. This fact also is left to the reader as an exercise to check.

Returning to $D = A \times B$ and $A \times B \xrightarrow{h} C$, $C \xrightarrow{h'} A \times B$ as above, we now have that *the canonical morphism $C \xrightarrow{h'} A \times B$ such that $\pi_A \circ h' = f$, $\pi_B \circ h' = g$ is an isomorphism.* This fact is sufficient for having that $A \xleftarrow{f} C \xrightarrow{g} B$ is a product diagram (since it is an isomorphic copy of $A \xleftarrow{\pi_A} A \times B \xrightarrow{\pi_B} B$). This completes the proof in one direction.

(ii) To show the other direction, assume that $A \xleftarrow{f} C \xrightarrow{g} B$ is a product diagram. We can now essentially reverse our previous argument. Taking $A \xleftarrow{\pi_A} A \times B \xrightarrow{\pi_B} B$ to be another product, the subobject $R \hookrightarrow (A \times B) \times C$ as defined in part (i) is the graph of the canonical map $h' : C \to A \times B$. But now h' is an isomorphism. Denoting its inverse by h, R will be the graph of h as well, now with $A \times B$ as domain and C as codomain. By 2.4.1, R, in the sense: "with $A \times B$ as domain and C as codomain", will be functional. As we said in part (i), this fact is equivalent to having the two "axioms for product" hold. $\square$

AD 8, SUPS: This case is completely tautologous. By definition, the condition that the axioms hold is equivalent to saying that $[\bigvee_{i\in I} \underset{\sim}{A}_i(x)]$ is defined and is both $\leq$ and $\geq$ than $[\underset{\sim}{B}(x)]$, i.e., that $[\bigvee_{i\in I} \underset{\sim}{A}_i(x)] = [\underset{\sim}{B}(x)]$. But $[\underset{\sim}{B}(x)] = B \overset{g}{\hookrightarrow} X$ by earlier remarks; also $[\bigvee_{i\in I} \underset{\sim}{A}_i(x)] = \bigvee_{i\in I} A_i$, with one side defined iff the other is, and with A_i now abbreviating the subobject $A_i \overset{f_i}{\hookrightarrow} X$. In other words, the axioms are equivalent to saying that $B = \bigvee_{i\in I} A_i$, as required. $\square$

AD 9, IMAGES: Trivial. Given $f : A \to B$, $[fa \approx b]$, the graph of f, is the subobject

$$A \xhookrightarrow{\langle \mathrm{id}_A, f\rangle} A \times B.$$

To say that $\Rightarrow \exists a(fa \approx b)$ holds is, by definition, equivalent to saying that the composite $A \xhookrightarrow{\langle \mathrm{id}_A, f\rangle} A \times B \xrightarrow{\pi_B} B$ is surjective. But this composite is nothing but f. $\square$

FURTHER REMARKS ON THE PROOF OF 2.4.5. The case 2, "commutative diagram", was discussed as a first example in the meaning of formulas. The cases 1, 3 and 4 are very easy. The case of equalizers is similar to that of products worked out above. The rest are more or less tautologous as shown by the last two proofs.

Chapter 3

Axioms and rules of inference valid in categories

§1 Some simple rules

We begin by discussing a small fragment of $L_{\infty\omega}$, which we call Horn-logic, and denote it by L_H. The only logical operator in L_H is finite conjunction. Accordingly, the formulas in L_H are the ones that are built up from arbitrary atomic formulas using only $\bigwedge$ applied to finite sets. Without loss of generality, the formulas can be taken to be finite conjunctions

$$\pi_1 \wedge \cdots \wedge \pi_n$$

of atomic formulas, together with $\top$ (true). For the purposes of this discussion, a sequent of L_H will mean one of the form $\phi \Rightarrow \psi$ with single formulas ϕ, ψ of L_H.

Throughout this section, let $\mathcal{R}$ be a category with finite left limits and let L be its extended canonical language. Let T be a set of sequents of L_H, σ a single sequent. Let us write $T \models \sigma$ for: σ is a logical consequence of T, meaning that whenever M is an ordinary (**Set**-)structure that is a model of T, then also M is a model of σ. The following simple result is an example for a *completeness theorem* more instances of which we will see later.

Proposition 3.1.1 *Suppose $T \models \sigma$ and that every axiom in T is true in $\mathcal{R}$ (under the canonical interpretation). Then σ is true in $\mathcal{R}$.*

Proof. The only kind of structures of type L we need are the functors $\mathrm{Hom}(A, -)$, with objects A of $\mathcal{R}$. Given A,

$$\mathrm{Hom}(A, -) : \mathcal{R} \to \mathbf{Set}$$

is the functor F such that $F(B) = \mathrm{Hom}_{\mathcal{R}}(A, B) = \mathrm{Hom}(A, B)$ and $F(f) : \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, B')$ (for $f : B \to B'$) is the map such that $F(f)(g) = fg$, for any $g \in \mathrm{Hom}(A, B)$. Hence, $F = \mathrm{Hom}(A, -)$ is automatically a structure of type L, in the narrower sense of the canonical language L; the *sort* B is interpreted as $\mathrm{Hom}(A, B)$, the *operation symbol* $f : B \to B'$ as $F(f)$. Moreover, as is well-known and also easy to check, $\mathrm{Hom}(A, -)$ preserves all projective (left) limits in $\mathcal{R}$, and it preserves monomorphisms. Hence, denoting $\mathrm{Hom}(A, -)$ by M_A, we have that an n-ary predicate symbol $R \hookrightarrow X_1 \times \cdots \times X_n$, and an n-ary operation symbol $f : X_1 \times \cdots \times X_n \to Y$ are interpreted by the subset $M_A(R) \hookrightarrow M_A(X_1) \times \cdots \times M_A(X_n)$ and the map $M_A(f) : M_A(X_1) \times \cdots \times M_A(X_n) \to$

$M_A(Y)$, respectively; so we can regard M_A as an interpretation in **Set** of the extended canonical language L.

Let ϕ be a formula of L_H and let $R = [\phi]_{\vec{x}} \xhookrightarrow{i} X = X_1 \times \cdots \times X_n$ be the canonical interpretation of ϕ in $\mathcal{R}$. We claim that the interpretation of ϕ by M_A, $(M_A)_{\vec{x}}(\phi)$ is nothing but $M_A(R) \xhookrightarrow{M_A(i)} M_A(X)$. This is an easy consequence of the facts that M_A preserves left limits and that the interpretation of formulas in L_H uses finite left limits only.

An immediate consequence of this claim is that if a sequent of L_H is true in $\mathcal{R}$, then it will remain true in all models M_A, $A \in \mathrm{Ob}\,\mathcal{R}$.

As a final preliminary step, consider monomorphisms $R_1 \xhookrightarrow{i_1} X$ and $R_2 \xhookrightarrow{i_2} X$ and assume that for every $A \in \mathrm{Ob}\,\mathcal{R}$, the subsets $M_A(R_1)$, $M_A(R_2)$ of X satisfy $M_A(R_1) \subseteq M_A(R_2)$. Then we claim that the subobject $R_1 \xhookrightarrow{i_1} X$ of X is $\leq$ the subobject $R_2 \xhookrightarrow{i_2} X$. In fact, as it is easy to see, it is enough to consider $A = R_1$!

Now, assume the hypotheses of the proposition. Since every axiom in T is true in $\mathcal{R}$, every axiom in T is true in the structure M_A, for any $A \in \mathrm{Ob}\,\mathcal{R}$. By the assumption, M_A satisfies σ, for any $A \in \mathrm{Ob}\,\mathcal{R}$. If $\sigma = \phi \Rightarrow \psi$, $[\phi]_{\vec{x}} = R_1 \xhookrightarrow{i_1} X$ $[\psi]_{\vec{x}} = R_2 \xhookrightarrow{i_2} X$ then $M_A(\phi)$ is $M_A(R_1) \xhookrightarrow{M_A(i_1)} M_A(X)$ and $M_A(\psi)$ is $M_A(R_2) \xhookrightarrow{M_A(i_2)} M_A(X)$. Since $M_A(R_1) \leq M_A(R_2)$ for any $A \in \mathrm{Ob}\,\mathcal{R}$, we have $R_1 \leq R_2$, i.e. $\phi \Rightarrow \psi$ is true in $\mathcal{R}$, as claimed. □

Remark *For a more general context for the last proposition, c.f. Kock and Reyes* [1977] *and the references there.*

As an application of the last proposition, we complete the proof of 2.4.5, for the case of products. Defining the subobject $R \hookrightarrow A \times B \times C$ as we did there by

$$R = [fc \approx a \wedge gc \approx b]$$

we have that the two sequents σ_1, σ_2

$$\begin{aligned} Rabc \quad &\Rightarrow \quad fc \approx a \wedge gc \approx b \\ fc \approx a \wedge gc \approx b \quad &\Rightarrow \quad Rabc \end{aligned}$$

are true in $\mathcal{R}$.

Let σ_3 be: $\quad fc \approx fc' \wedge gc \approx gc' \Rightarrow c \approx c'$,

and σ_4: $\quad Rabc \wedge Rabc' \Rightarrow c \approx c'$.

Then clearly $\quad \{\sigma_1, \sigma_2, \sigma_3\} \models \sigma_4$,

and $\quad \{\sigma_1, \sigma_2, \sigma_4\} \models \sigma_3$,

in the sense of ordinary **Set**-models. Hence, by 3.1.1 σ_3 is true in $\mathcal{R}$ iff σ_4 is, as required in the appropriate place the proof of 2.4.5.

We now turn to discussing the rest of the logical operators. We start with $\exists_f$.

Proposition 3.1.2 *Given $X \xrightarrow{f} Y$, then the following are equivalent for arbitrary subobjects A of X and B of Y, provided that $\exists_f A$ exists:*

$$\begin{aligned} &\text{(i)} \quad A \leq f^{-1}(B), \\ &\text{(ii)} \quad \exists_f A \leq B. \end{aligned}$$

The proposition is essentially equivalent to the definition of $\exists_f A$. The direction (i)$\rightarrow$(ii) uses the universal property of the pullback $f^{-1}(B)$. We note that it is essentially this form in which Lawvere [1965] first introduced the categorical notion of existential quantifier. Similarly, he used that 'adjoint' formulation of $\forall$ as formulated in

Proposition 3.1.3 *With the notation of 3.1.2, if $\forall_f(A)$ exists, then the following are equivalent*

$$\begin{array}{rrcl} \text{(i)} & f^{-1}(B) & \leq & A, \\ \text{(ii)} & B & \leq & \forall_f(A). \end{array}$$

A generalization of 3.1.3 is

Proposition 3.1.3' *Supposing that $\forall_f(A_1 \rightarrow A_2)$ exists, the following are equivalent for any $B \rightarrow Y$.*

$$\begin{array}{rrcl} \text{(i)} & f^{-1}(B) \wedge A_1 & \leq & A_2, \\ \text{(ii)} & B & \leq & \forall_f(A_1 \rightarrow A_2). \end{array}$$

Given subobjects $A \hookrightarrow X \times Y$, and $B \hookrightarrow Y$, the previous propositions applied to the projection $\pi_Y : X \times Y \rightarrow Y$, we obtain that

Corollary 3.1.4

$$\underset{\sim}{A}xy \Rightarrow \underset{\sim}{B}y$$

and

$$\exists x \underset{\sim}{A}xy \Rightarrow \underset{\sim}{B}y$$

are equivalent, and similarly

$$\underset{\sim}{B}y \Rightarrow \underset{\sim}{A}xy$$

and

$$\underset{\sim}{B}y \Rightarrow \forall x \underset{\sim}{A}xy$$

are equivalent, provided $[\exists x \underset{\sim}{A}xy]$ ($[\forall x \underset{\sim}{A}xy]$) is defined. (Here $\underset{\sim}{A}xy$ and $\underset{\sim}{B}y$ are formulas such that $[\underset{\sim}{A}xy] = A \hookrightarrow X \times Y$, $[\underset{\sim}{B}y] = B \hookrightarrow Y$.)

Similar facts hold with an arbitrary formula $\phi(x, \vec{y})$ in place of $\underset{\sim}{A}xy$, etc., provided $[\phi(x, \vec{y})]$, etc., are defined.

Proposition 3.1.6 *The composition gf of two surjective morphisms is surjective.*

Proof. Suppose $A \xrightarrow{f} B$ and $B \xrightarrow{g} C$ are surjective. Suppose $D \hookrightarrow C$ is a monomorphism such that gf factors through it:

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

By the universal property of the pullback $g^{-1}(D)$, we have $A \rightarrow g^{-1}(D)$ such that

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[dr] \arrow[drr, "h"] & B \arrow[r, "g"] & C \\
& g^{-1}(D) \arrow[u, hook] \arrow[r] & D \arrow[u, hook]
\end{tikzcd}$$

is commutative. Since f is surjective, $g^{-1}(D) \rightarrow B$ is an isomorphism. Letting $B \rightarrow g^{-1}(D)$ be its inverse, the composite

$$B \rightarrow g^{-1}(D) \rightarrow D$$

shows that g factors through $D \hookrightarrow C$. Since g is surjective, $D \to C$ is an isomorphism. $\square$

As an immediate consequence, we have the following

Corollary 3.1.6 *Given* $A \hookrightarrow X \xrightarrow{f} Y \xrightarrow{g} Z$,

$$\exists_{gf}(A) = \exists_g(\exists_f(A)).$$

Consider the projection

$$\begin{tikzcd} X \times Y \times Z \arrow[r, "p"] \arrow[dr, "p'"'] & X \times Y \arrow[r, "q"] & X \\ & X \times Z \arrow[ur, "q'"'] & \end{tikzcd}$$

and a subobject $A \to X \times Y \times Z$. We obtain

Corollary 3.1.7 $\exists y \exists z \underset{\sim}{A} xyz \iff \exists z \exists y \underset{\sim}{A} xyz$
holds in $\mathcal{R}$, provided the required interpretations exist in $\mathcal{R}$.

The reason is that both sides, when interpreted, become $\exists_f(A)$ where $f = qp = q'p'$. An easy fact is

Proposition 3.1.8 *For $A \to X$, $B \to X$, $X \xrightarrow{f} Y$ if $A \leq B$ and both $\exists_f(A)$ and $\exists_f(B)$ exist, then $\exists_f(A) \leq \exists_f(B)$.*

Next, we derive a formula connecting sups and $\exists$.

Proposition 3.1.9 *Suppose $R_i \hookrightarrow X$ $(i \in I)$ and $X \xrightarrow{f} Y$ are given.*

(i) *Assume that $\bigvee_{i \in I} R_i$, $\exists_f(\bigvee_{i \in I} R_i)$ and $\exists_f(R_i)$ $(i \in I)$ all exist. Then $\bigvee_{i \in I} \exists_f(R_i) = \exists_f(\bigvee_{i \in I} R_i)$*

(ii) *The same equality holds if we assume that $\bigvee_{i \in I} R_i$, $\exists_f(R_i)$ $(i \in I)$, and $\bigvee_{i \in I} \exists_f(R_i)$ all exist.*

Proof. (ad i). $\exists_f(R_i) \leq \exists_f(\bigvee_{i \in I} R_i)$ follows from 3.1.8. Assume that each $\exists_f(R_i) \hookrightarrow Y$ factors through $B \hookrightarrow Y$. Consider $f^{-1}(B) \hookrightarrow X$. We obtain that $R_i \leq f^{-1}(B)$, $i \in I$. Hence $\bigvee_{i \in I} R_i \leq f^{-1}(B)$ and thus

$$\exists_f(\textstyle\bigvee_{i \in I} R_i) \leq B.$$

$\square$

The proof of (ii) is similar.

Corollary 3.1.10 *With subobjects $A_i \hookrightarrow X \times Y$, we have that $\exists y \bigvee_{i \in I} \underset{\sim}{A}_i xy \iff \bigvee_{i \in I} \exists y \underset{\sim}{A}_i xy$ holds in $\mathcal{R}$, provided the interpretations*

$$[\textstyle\bigvee_{i \in I} \underset{\sim}{A}_i xy], \ [\exists y \bigvee_{i \in I} \underset{\sim}{A}_i xy], \ [\exists y \underset{\sim}{A}_i xy]$$

or *the ones* $[\bigvee_{i \in I} \underset{\sim}{A}_i xy]$, $[\exists y \underset{\sim}{A}_i xy]$, $[\bigvee_{i \in I} \exists y \underset{\sim}{A}_i xy]$ *exist.*

§2. Stability and distributivity

In order that sups and images behave sufficiently well, we have to require that they be *stable*, i.e. stable under pullbacks. "Stability" is used here in the same sense as "universality" is in SGA 4.

Let L be the extended canonical language of $\mathcal{R}$ as before. All formulas and terms are in $L_{\infty\omega}$ and $[\phi]$, $[t]$ refer to the canonical interpretation.

Definition 3.2.1 (i) *A morphism* $A \xrightarrow{f} B$ *is called* stable surjective *if it is surjective and for every* $B' \xrightarrow{g} B$, $f' : A \times_B B' \to B'$ *in the pullback diagram*

$$\begin{tikzcd} A \arrow[r, "f"] & B \\ A \times_B B' \arrow[u] \arrow[r, "f'"'] & B' \arrow[u, "g"'] \end{tikzcd}$$

is surjective too.

(ii) *Given* $A \hookrightarrow X \xrightarrow{f} Y$, *the image* $\exists_f(A) \hookrightarrow Y$ *is called* stable *if the surjective map* h *in*

$$\begin{tikzcd} A \arrow[r, hook] \arrow[d, "h"'] & X \arrow[d, "f"] \\ \exists_f(A) \arrow[r, hook] & B' \end{tikzcd}$$

is stable surjective.

Given $A \hookrightarrow X \xrightarrow{f} Y$, the stability of the image $\exists_f(A)$ is equivalent to the following. Let $Y' \xrightarrow{g} Y$ be any morphism. Form the pullback

$$\begin{tikzcd} X' \arrow[r, "h"] \arrow[d, "f'"'] & X \arrow[d, "f"] \\ Y' \arrow[r, "g"'] & Y. \end{tikzcd}$$

Then we have that $\exists_{f'}(h^{-1}(A))$ exists and equals $g^{-1}(\exists_f(A))$.

Remark (ii) *is the so-called Beck-Chevalley condition, c.f. e.g. Kock-Reyes* [1977].

(iii) *Given* $A_i \hookrightarrow X$ $(i \in I)$ *such that* $\bigvee_{i\in I} A_i \hookrightarrow X$ *exists, we say that the sup* $\bigvee A_i$ *is* stable *if for any* $X' \xrightarrow{g} X$, $\bigvee_{i\in I} g^{-1}(A_i)$ *exists and equals to* $g^{-1}(\bigvee_{i\in I} A_i)$.

(iv) *Given* $A_1 \hookrightarrow X$, $A_2 \hookrightarrow X$, *and* $X \xrightarrow{f} Y$, *the generalized dual image*

$$\forall_f(A_1 \to A_2) \to Y$$

is called stable *if for any*

$$Y' \xrightarrow{g} Y,$$

for the pullback

$$\begin{tikzcd} X' \arrow[r, "h"] \arrow[d, "f'"'] & X \arrow[d, "f"] \\ Y' \arrow[r, "g"'] & Y \end{tikzcd}$$

we have that $\forall_{f'}(h^{-1}(A_1) \to h^{-1}(A_2))$ *exists and equals to* $g^{-1}(\forall_f(A_1 \to A_2))$.

(v) *Let* ϕ *be a formula of* L *such that* $[\phi]_{\vec{x}}$ *exists, with* $\vec{x}$ *the sequence of free variables in*

ϕ. We say that ϕ is stable *if every image, sup and dual image evaluated in the course of computing $[\phi]_{\vec{x}}$ is stable.*

Given a sequence $\vec{x}$ of distinct variables $\vec{x} = \langle x_1, \ldots, x_n \rangle$, x_i of sort A_i (an object in $\mathcal{R}$), let $[\vec{x}]$ denote the product $A_1 \times \cdots \times A_n$. Let $\vec{x}$, $\vec{y}$ be two sequences of distinct variables, $\vec{x}$ contained in $\vec{y}$. Let p be the canonical projection $[\vec{y}] \to [\vec{x}]$.

A first consequence of the definition is

Proposition 3.2.2 *With the previous notation, if ϕ is a stable formula with free variables among $\vec{x}$, we have that*

$$M_{\vec{y}}(\phi) = p^{-1}(M_{\vec{x}}(\phi)).$$

The proof is an easy induction on the complexity of ϕ.

Recall that $\phi(t/x)$ denotes the result of formally substituting the term t for x at each occurrence of x in ϕ. Here t has the same sort as x. In elementary logic, one has the substitution lemma saying that, roughly, one can evaluate $\phi(t/x)$ by first evaluating at t and then using the value thus obtained as the value for x in evaluating ϕ. We are going to state the substitution lemma for the categorical interpretation.

Let ϕ, t, x and $\phi(t/x)$ as above.

Substitution lemma 3.2.3 *Suppose ϕ is a stable formula. Suppose t does not contain the variable x. Let $\vec{y}$ be a sequence of variables not containing x but containing all free variables in $\phi(t/x)$. Then we have a pullback diagram as follows*

$$\begin{tikzcd}
{[\phi]_{x\vec{y}}} \arrow[r, hook] & {[x] \times [\vec{y}]} \\
{[\phi(t/x)]_{\vec{y}}} \arrow[r, hook] \arrow[u] & {[\vec{y}]} \arrow[u, "\langle [t]_{\vec{y}}, \mathrm{id}_{[\vec{y}]} \rangle"']
\end{tikzcd}$$

Remarks *This uses the morphism*

$$[t]_{\vec{y}} : [\vec{y}] \to [x]$$

obtained by interpreting t. Also, if we denote $\langle [t]_{\vec{y}}, \mathrm{id}_{[\vec{y}]} \rangle$ by g, the assertion is equivalent to the equality $[\phi(t/x)]_{\vec{y}} = g^{-1}([\phi]_{x\vec{y}})$.

As another remark, we note that, of course, we could substitute t for x even if t does contain x free. However, in this case we can choose a free variable x' such that x' is new, *i.e. it does not occur in ϕ or t, form $\phi' = \phi(x'/x)$ and have $\phi(t/x) = \phi'(t/x')$. Now, the last substitution is of the kind in the lemma. Applying the lemma and eliminating ϕ' we can show the existence of a pullback*

$$\begin{tikzcd}
{[\phi]_{x\vec{y}}} \arrow[r, hook] & {[x] \times [\vec{y}]} \\
{[\phi(t/x)]_{x\vec{y}}} \arrow[r, hook] \arrow[u] & {[x] \times [\vec{y}]} \arrow[u, "\langle [t]_{x\vec{y}}, p \rangle"']
\end{tikzcd}$$

where p is the canonical projection $[x] \times [\vec{y}] \xrightarrow{p} [\vec{y}]$.

Proof of 3.2.2. The proof is by induction on the complexity of the formula ϕ. For atomic formulas, the assertion follows from familiar commutation properties of left limit diagrams. For conjunctions $\bigwedge \Sigma$, one has an automatic stability of infs that gives the desired result. For the rest of the logical operators, one has to use stability as postulated for logical operators in ϕ. We will show the case $\phi = \exists z\psi$.

Without loss of generality, we assume that z does not occur in t, and $z \neq x$. Then we have that the formula $\phi(t/x)$ is identical to $\exists z(\psi(t/x))$. Let A be the subobject $[\psi]_{x\vec{y}z}$ of $X = [x] \times [\vec{y}] \times [z]$ and consider the following pullback diagram, taking the role of the pullback diagram in Definition 3.2.1(ii):

$$\begin{tikzcd}
{[x] \times [\vec{y}] \times [z]} \arrow[r, "{\langle [t]_{x\vec{y}z}, \mathrm{id}_{[\vec{y}z]}\rangle}"] \arrow[d, "\text{can.}"'] & {[\vec{y}] \times [z]} \arrow[d, "\text{can.}"] \\
{[x] \times [\vec{y}]} \arrow[r, "{\langle [t]_{x\vec{y}}, \mathrm{id}_{[\vec{y}]}\rangle}"'] & {[\vec{y}]}
\end{tikzcd}$$

Denote the last pullback diagram more briefly by

$$\begin{tikzcd}
X' \arrow[r, "h"] \arrow[d, "f'"'] & X \arrow[d, "f"] \\
Y' \arrow[r, "g"'] & Y
\end{tikzcd}$$

as in 3.2.1(ii). Then
$$\begin{aligned}[\phi(t/x)]_{\vec{y}} &= [\exists z(\psi(t/x))]_{\vec{y}} \\ &= \exists_{f'}([\psi(t/x)]_{\vec{y}z}).\end{aligned}$$
By the induction hypothesis, the substitution lemma is true for ψ. Apply it to get

$$[\psi(t/x)]_{\vec{y}z} = h^{-1}([\psi]_{x\vec{y}z}) = h^{-1}(A).$$

So $[\phi(t/x)]_{\vec{y}} = \exists_{f'} h^{-1}(A)$. By the stability of ϕ, and 3.2.1(ii), the latter equals
$$g^{-1}(\exists_f(A)) = g^{-1}([\phi]_{x\vec{y}}).$$
□

The main consequence of the substitution lemma is

Proposition 3.2.4 *For a stable formula $\exists x\phi$, the sequent*

$$\phi(t/x) \Rightarrow \exists x\phi$$

is valid.

Proof. Without loss of generality, we can assume that x does not occur in t. Putting together the diagrams of the substitution lemma and the definition of $[\exists x\phi]_{\vec{y}}$ we get the commutative diagram

$$\begin{tikzcd}
{[\phi(t/x)]_{\vec{y}}} \arrow[r, hook] \arrow[d] & {[\vec{y}]} \arrow[d, "{\langle [t]_{\vec{y}}, \mathrm{id}_{[\vec{y}]}\rangle = g}"] \\
{[\phi]_{x\vec{y}}} \arrow[d] & {[x] \times [\vec{y}]} \arrow[d] \\
{[\exists x\phi]_{\vec{y}}} \arrow[r, hook] & {[\vec{y}]}
\end{tikzcd}$$

Since the composition $[\vec{y}] \xrightarrow{g} [x] \times [\vec{y}] \to [\vec{y}]$ is the identity, $[\phi(t/x)]_{\vec{y}} \leq [\exists\phi]_{\vec{y}}$. □

Another formulation of the substitution lemma is 3.2.5 below. Let ϕ, or $\phi(x_1, \ldots, x_n)$, be a formula whose free variables are exactly the distinct free variables $x_1, \ldots, x_n = \vec{x}$. Let $t_1, \ldots, t_n$ be arbitrary terms such that t_i is of the same sort as x_i $(i = 1, \ldots, n)$ and let $\vec{y}$ be a sequence of free variables containing all the variables in any of $t_1, \ldots, t_n$. Let $\phi(t_1, \ldots, t_n)$ denote the result of substituting t_i for x_i, $i = 1, \ldots, n$.

Corollary 3.2.5 *With the above notation, if ϕ is stable, we have a pullback diagram*

$$\begin{tikzcd}
{[\phi(x_1, \ldots, x_n)]_{\vec{x}}} \arrow[r, hook] & {[\vec{x}] = [x_1] \times \cdots \times [x_n]} \\
{[\phi(t_1, \ldots, t_n)]_{\vec{y}}} \arrow[r, hook] \arrow[u] & {[\vec{y}]} \arrow[u, "{\langle [t_1]_{\vec{y}}, \ldots, [t_n]_{\vec{y}} \rangle}"']
\end{tikzcd}$$

The proof can be given by repeated application of 3.2.3.

A direct consequence of stability is

Proposition 3.2.6 *If the sup $\bigvee_{i \in I} A_i$ of subobjects of X is stable, and B is another subobject of X, then*

$$B \wedge \bigvee_{i \in I} A_i = \bigvee_{i \in I} (B \wedge A_i).$$

As a consequence, if all finite sups exist and are stable, the subobjects of X form a distributive lattice*.*

Similarly we have

Proposition 3.2.7 (i) *If, for $A \hookrightarrow X \xrightarrow{f} Y$, the image $\exists_f(A)$ is stable, and $B \hookrightarrow Y$ is another subobject, then*

$$\exists_f(f^{-1}(B) \wedge A) = B \wedge \exists_f(A).$$

(ii) *The sequent*

$$\exists y(\underset{\sim}{B}x \wedge \underset{\sim}{A}xy) \iff \underset{\sim}{B}x \wedge \exists y \underset{\sim}{A}xy$$

is valid if the formula $\exists y \underset{\sim}{A}xy$ is stable.

Next we formulate some special properties of infs and dual images. Allowing that our remark might be obscure, we note that while the above principles were all intuitionisitically valid, distributivity as we introduce it below is only classically valid.

In the rest of this section, we assume that $\mathcal{R}$ *has finite sups*, i.e., sups of finite families of subobjects always exist.

Suppose the inf $\bigwedge_{i \in I} A_i \hookrightarrow X$ exists. We say that it is *distributive* if for any $B \hookrightarrow X$, $B \vee \bigwedge_{i \in I} A_i = \bigwedge_{i \in I}(B \vee A_i)$. We say that $\bigwedge_{i \in I} A_i$ is *distributive in a stable way* (or *stably distributive*) if for every $X' \xrightarrow{f} X$, $\bigwedge_{i \in I} f^{-1}(A_i)$ is distributive. Note that stability of finite sups implies that finite infs are distributive.

Given $\forall_f(A_1 \to A_2) \hookrightarrow Y$, with

$$\begin{tikzcd}
A_1 \arrow[dr, hook] & \\
& X \arrow[d, "f"] \\
A_2 \arrow[ur, hook] & Y
\end{tikzcd}$$

we say $\forall_f(A_1 \to A_2)$ is *distributive* if for every $B \hookrightarrow Y$, $\forall_f(A_1 \to (A_2 \vee f^{-1}(B))$ exists and is equal to $\forall_f(A_1 \to A_2) \vee B$. We talk about distributivity *in a stable way* if (in addition) $\forall_f(A_1 \to A_2)$ is stable, and for every pullback diagram

$$\begin{tikzcd}
X' \arrow[r, "h"] \arrow[d, "f'"'] & X \arrow[d, "f"] \\
Y' \arrow[r, "g"'] & Y
\end{tikzcd}$$

$\forall_{f'}(h^{-1}(A_1) \to h^{-1}(A_2))$ is distributive. We call a formula (of the canonical language) *distributive* if it is stable, and each inf and $\forall_f$ needed to evaluate it in $\mathcal{R}$ is distributive in a stable way.

Finally, in this section, we formulate the second main fact about the categorical interpretation. We ask the reader to look up the two *formal systems* we introduce in Chapter 5. In the formal systems, axioms and rules of inferences are given and a corresponding notion of derivability is defined. An instance of a rule of inference is of the form

$$\frac{\{\sigma_i : i \in I\}}{\sigma}$$

with one or more *hypotheses* σ_i and a *conclusion* σ; each of σ_i, σ is a sequent. Let $\mathcal{R}$ be a category with finite left limits, L its canonical language, and let F be a fragment of $L_{\infty\omega}$. We say that F is *stable (distributive) (in $\mathcal{R}$)* if every formula in F can be interpreted in $\mathcal{R}$, and actually, every formula in F is stable (distributive). We say that an instance of a rule is *valid (in $\mathcal{R}$)* in case the truth of all the hypotheses of the instance implies the truth of the conclusion in $\mathcal{R}$.

We denote by $T \overset{1}{\vdash} \sigma$ ($T \overset{2}{\vdash} \sigma$) the fact that "$\sigma$ is a formal consequence of T", according to the formal system of Section 1 (Section 2) of Chapter 5. Here we call these two formal systems G_T^1 and G_T^2, respectively. In other words, $T \overset{1}{\vdash} \sigma$ iff σ es obtained from the axioms of G_T^1 by repeated applications of the rules of inference.

Soundness Theorem 3.2.8 (i) *Assume $\mathcal{R}$ has stable finite sups. Assume the fragment F is distributive in $\mathcal{R}$, and T is a theory (set of sequents) in F all of whose elements are true in $\mathcal{R}$. Then all axioms of G_T^1 belonging to F, and all instances of rules of G_T^1 with conclusions belonging to F, are valid in $\mathcal{R}$. Hence, if $T \overset{1}{\vdash} \sigma$ and $\mathcal{R}$ satisfies all sequents in T, then $\mathcal{R}$ satisfies σ (in the canonical interpretation).*

(ii) *Assume F is a* coherent *fragment ($F \subset L^g_{\infty\omega}$), F is stable in $\mathcal{R}$, and T is as before. Then the same conclusion holds with respect to the "one sided" system G_T^2.*

Proof. Above we collected all the necessary facts needed for the proof.

(ad (i)) The axioms (A1) and the rules ($\bigwedge \Rightarrow$), ($\Rightarrow \bigvee$) are direct consequences of the definitions and the don't use stability of distributivity.

(ad $(\mathrm{A2})_{T^=}$) The equality axioms that are added to T to form $T^=$ are all true under the ordinary interpretation, and the are in L_H, so by 3.1.1 they are true in $\mathcal{R}$. Starting with this fact, we have to use the substitution lemma to prove $(\mathrm{A2})_{T^=}$. If $\Theta \Rightarrow \Gamma$ belongs to $T^=$, then it is true in $\mathcal{R}$. Using the shorthand $\vec{t}$ for $t_1, \ldots, t_n$,

$$\begin{aligned}
[\textstyle\bigwedge\Theta(\vec{t})]_{\vec{y}} &= \textstyle\bigwedge_{\theta\in\Theta}[\theta(\vec{t})]_{\vec{y}}, \\
[\textstyle\bigvee\Gamma(\vec{t})]_{\vec{y}} &= \textstyle\bigvee_{\gamma\in\Gamma}[\gamma(\vec{t})]_{\vec{y}}.
\end{aligned}$$

Denoting $\langle [t_1]_{\vec{y}}, \ldots, [t_n]_{\vec{y}} \rangle : [\vec{y}] \to [\vec{x}]$ by g, by 3.2.5 we have

$$[\phi(\vec{t})]_{\vec{y}} = g^{-1}([\phi]_{\vec{x}})$$

for $\phi \in \Theta$, or $\phi \in \Gamma$. Hence

$$[\textstyle\bigwedge\Theta(\vec{t})]_{\vec{y}} = g^{-1}([\bigwedge\Theta]_{\vec{x}}) \qquad \text{and} \qquad [\bigvee\Gamma(\vec{t})]_{\vec{y}} = g^{-1}([\bigvee\Gamma]_{\vec{x}})$$

by the stability of finite sups in $\mathcal{R}$. Since $A \leq B$ obviously implies $g^{-1}(A) \leq g^{-1}(B)$, we have

$$[\textstyle\bigwedge\Theta(\vec{t})]_{\vec{y}} \leq [\bigvee\Gamma(\vec{t})]_{\vec{y}},$$

hence clearly $\Phi,\, \Theta(\vec{t}) \Rightarrow \Phi,\, \Gamma(\vec{t})$ is true as desired.

(AD $\Rightarrow \bigwedge$) By distributivity of the inf $\bigwedge_{\theta\in\Theta}[\theta]_{\vec{y}}$, we have that

$$\bigwedge_{\theta\in\Theta}[\bigvee\Psi \vee \bigwedge\Theta \vee \theta]_{\vec{y}} = [\bigvee\Psi \vee \bigwedge\Theta]_{\vec{y}} \vee [\bigwedge\Theta]_{\vec{y}} = [\bigvee\Psi \vee \bigwedge\Theta]_{\vec{y}}.$$

Suppose $[\bigwedge\Phi]_{\vec{y}} \leq [\bigvee\Psi \vee \bigwedge\Theta \vee \theta]_{\vec{y}}$ for all $\theta \in \Theta$. By the definition of $\bigwedge$, we can take the inf $\bigwedge_{\theta\in\Theta}$ of the right hand side sub objects and still have a valid inequality. By the equality deduced first, we obtain exactly the desired conclusion.

(AD $\bigvee \Rightarrow$) This is very similar to the previous proof, using 3.2.6 in the appropriate place.

(AD $\exists \Rightarrow$) Since y does not occur in $\bigwedge\Psi$, by 3.1.2 (or 3.1.4) the hypothesis implies that

$$\exists y(\bigwedge\Phi \wedge \exists x\theta(x) \wedge \theta(y)) \Rightarrow \Psi$$

is true. Since y does not occur in $\bigwedge\Phi$ and $\exists x\theta(x)$, and by the stability of the sup in $[\exists y\theta(y)]$, 3.2.7 yields that the left-hand-side is equivalent to

$$\bigwedge\Phi \wedge \exists x\theta(x) \wedge \exists y\theta(y)$$

hence to $\bigwedge\Phi \wedge \exists x\theta(x)$. Therefore,

$$\bigwedge\Phi \wedge \exists x\theta(x) \Rightarrow \Psi$$

is true as required.

(AD $\Rightarrow \exists$) Let $\vec{y}$ be the set of all free variables in the hypothesis. We assume that

$$[\bigwedge\Phi]_{\vec{y}} \leq [\bigvee\Psi]_{\vec{y}} \vee [\exists x\theta(x)]_{\vec{y}} \vee [\theta(t)]_{\vec{y}}.$$

By 3.2.4, $[\theta(t)]_{\vec{y}} \leq [\exists x\theta(x)]_{\vec{y}}$, hence $[\bigwedge\Phi]_{\vec{y}} \leq [\bigvee\Psi]_{\vec{y}} \vee [\exists x\theta(x)]_{\vec{y}}$. By the free-variable-proviso of the rule ($\Rightarrow \exists$), $\vec{y}$ is exactly the set of free variables in the conclusion. Hence, the last equality means exactly that the conclusion is true in $\mathcal{R}$.

(AD $\forall \Rightarrow$) This is similar to ($\Rightarrow \exists$); the details are omitted.

(AD $\Rightarrow \forall$) We use 3.1.3' and the distributivity of the $\forall$-formula involved. Let $\vec{x}$ be the set of free variables in the conclusion

$$\begin{array}{rl}
 & X = [\vec{x}] \times [\vec{y}] \\
 & Y = [\vec{x}] \\
 & f : X \to Y \text{ the canonical projection} \\
 & B = [\bigwedge\Phi]_{\vec{x}} \hookrightarrow Y \\
\text{(then} & [\bigwedge\Phi]_{\vec{x}\vec{y}} = f^{-1}(B)\ ) \\
 & A_1 = [\phi(y)]_{\vec{x}\vec{y}} \hookrightarrow X \\
 & A_2 = [\psi(y)]_{\vec{x}\vec{y}} \hookrightarrow X \\
 & C = [\bigwedge\Psi \vee \forall\vec{z}(\phi \to \psi)]_{\vec{x}} \hookrightarrow Y \\
\text{so} & [\bigwedge\Psi \vee \forall\vec{z}(\phi \to \psi)]_{\vec{x}\vec{y}} = f^{-1}(C).
\end{array}$$

With this notation, the hypothesis can be written

$$f^{-1}(B) \wedge A_1 \leq A_2 \vee f^{-1}(c).$$

Hence, by 3.1.3', we have

$$B \leq \forall_f (A_1 \to (A_2 \vee f^{-1}(c)))$$

provided the right hand side is defined. But of course $[\forall \vec{z}(\phi \to \psi)]_{\vec{x}} = [\forall \vec{y}(\phi(\vec{y}) \to \psi(\vec{y}))]_{\vec{x}} = \forall_f(A_1 \to A_2)$ and the latter is distributive. Hence $\forall_f(A_1 \to (A_2 \vee f^{-1}(c)))$ exists and is equal to $[\forall \vec{z}(\phi \to \psi)]_{\vec{x}} \wedge C$. Thus the last inequality is equivalent to the validity of the conclusion of the rule.

The rule $(\mathrm{CUT})_T =$ is left to the reader. The free-variable-proviso is used as in $(\Rightarrow \exists)$.

(AD (ii)) The details are similar, and sometimes simpler, than those for (i). Since we don't have infinite infs or $\forall$'s, we don't need distributivity. Also, since the sequence have only one formula on the right hand side, we don't need the blanket assumption on the stability of finite sups (although of course, we need the stability of the sups that are built into *formulas* of F).

§3. Further categorical notions and their expression by formulas

In this section we will relate some notions in SGA4 Exposé I, Section 10 with our present framework. Notice that by our blanket assumption that all our categories have finite left limits, some of the distinctions in SGA4 automatically disappear. Let $\mathcal{R}$ be a category fixed throughout this section and L be the canonical language of $\mathcal{R}$.

In SGA4, we have the definitions: an object 0 is an *initial object* if for every object A, there is exactly one morphism $0 \to A$. 0 is a *strict initial object* if, in addition, every morphism $B \to 0$ is an isomorphism.

Proposition 3.3.1 *Let* 0 *be the minimal subobject of* 1*,* 1 *the empty product. Assume that* $0 \to 1$ *as the empty sup is stable. Then* 0 *is a strict initial object.*

PROOF. The stability means that the minimal subobject $0_A \to A$ is $0 \times A \xhookrightarrow{\text{can. project.}} A$. Let f, g be two morphisms $B \rightrightarrows 0_A$; we claim $f = g$. The reason is that both graph(f) and graph(g) are subobjects of $B \times 0_A = 0_{A \times B}$; by the minimality of $0_{A\times B}$, graph$(f) =$ graph$(g) = 0_{A \times B}$, thus $f = g$. The claim implies that, a fortiori, the canonical projection $0_A \to 0$ is a monomorphism. By the minimality of $0 \hookrightarrow 1$, it follows that the canonical projection $0_A \to 0$ is an isomorphism. This clearly gives us a morphism $0 \to A$. By referring to graphs again, it is easy to see that there can be only one $0 \to A$. Finally, let $B \xrightarrow{f} 0$ be a morphism. With $R = \mathrm{graph}(f) \hookrightarrow B \times 0$, we have the commutative diagram

$$\begin{tikzcd}
R \arrow[r, "\sim"] \arrow[d, hook, "i"'] & B \arrow[dl, hook] \arrow[ddl, hook, bend left, "\langle \mathrm{id}_B, f\rangle"] \\
B \times 0 \arrow[d, "\sim"'] & \\
0. &
\end{tikzcd}$$

Since by minimality, i is an isomorphism too, it follows that f is an isomorphism. $\square$

We have the following definition in loc. cit.:

Definition 3.3.2 (i) *A family* $A_i \xrightarrow{f_i} B$ $(i \in I)$ *of morphisms is called an* effective epimorphic *family if the following is satisfied: for any* C *and any family*

$$A_i \xrightarrow{g_i} C \quad (i \in I)$$

such that for the pullback diagram

$$\begin{tikzcd} & A_i \arrow[dr, "f_i"] & \\ A_i \times_B A_j \arrow[ur, "e_i^j"] \arrow[dr, "e_j^i"'] & & B \\ & A_j \arrow[ur, "f_j"'] & \end{tikzcd}$$

we have $g_i e_i^j = g_j e_j^i$*, for any choice of* $i, j \in I$*, there is a unique morphism* $B \xrightarrow{g} C$ *such that*

$$g_i = g f_i$$

for all $i \in I$*.*

(ii) $\langle f_i : i \in I \rangle$ *is a* universal *(or* stable*) effective epimorphic family if, in addition, it remains effective epimorphic after pulling back along any* $B' \xrightarrow{h} B$*, i.e., for the pullback diagrams*

$$\begin{tikzcd} A_i \arrow[r, "f_i"] & B \\ A_i \times_B B' \arrow[u] \arrow[r, "f_i'"'] & B' \arrow[u, "h"'] \end{tikzcd}$$

$\langle f_i' : i \in I \rangle$ *is effective epimorphic.*

Remark This condition is equivalent to saying that the family $(A_i \xrightarrow{f_i} A)_{i \in I}$ belongs to the finest Grothendieck topology in which the representable presheaves of $\mathcal{R}$ are sheaves: c.f. loc. cit. or Chapter 1. Also notice that an effective epimorphic family is *epimorphic*, meaning that whenever g_1 and g_2 are such that $g_1 f_i = g_2 f_i$ for all $i \in I$, then $g_1 = g_2$. This is a consequence of the uniqueness part of the definition.

Proposition 3.3.3 *Assume that* $\mathcal{R}$ *has stable images, i.e. for every* $A \hookrightarrow X \xrightarrow{f} Y$*,* $\exists_f(A) \hookrightarrow Y$ *exists, and it is stable. Also assume that, for a given set* I*, sups of families* $\{B_i \hookrightarrow X; i \in I\}$ *of subobjects indexed by* I *exist and they are stable. Then the following are equivalent:*

(i) $\{A_i \xrightarrow{f_i} B : i \in I\}$ *is an effective epimorphic family.*

(ii) $B = \bigvee_{i \in I} \exists_{f_i}(A_i)$*.*

Moreover, every effective epimorphic family indexed by I *is stable.*

Proof. (i) $\Rightarrow$ (ii): Assume (i) and in addition, let $D \xrightarrow{\ell} B$ be a monomorphism such that each f_i factors through ℓ:

$$\begin{tikzcd} A_i \times_B A_j \arrow[r, "e_i^j"] & A_i \arrow[rr, "f_i"] \arrow[dr, "h_i"'] & & B, \\ & & D \arrow[ur, "\ell"'] & \end{tikzcd} \qquad f_i = \ell h_i.$$

From $f_i e_i^j = f_j e_j^i$ and the fact that ℓ is a monomorphism, we deduce that

$$h_i e_i^j = h_j e_j^i.$$

Hence, by (i), there is $g : B \to D$ such that

$$\begin{aligned} h_i &= g f_i. \qquad \text{Then} \\ \ell h_i &= \ell g f_i \\ \text{i.e.} \qquad f_i &= \ell g f_i \qquad \text{for all } i \in I. \end{aligned}$$

Since the family is *epimorphic*, it follows that $\ell g = \mathrm{id}_B$. Considering now $g\ell : D \to D$, we have

$$\ell g \ell = \ell\, \mathrm{id}_D$$

since both sides equal ℓ. Since ℓ is a monomorphism, $g\ell = \mathrm{id}_D$. We see that our assumptions on $D \xrightarrow{\ell} B$ leads to the conclusion that ℓ is an isomorphism.

Now assume that each $\exists_{f_i}(A) \hookrightarrow B$ exists. To show that $B = \bigvee_{i\in I} \exists_{f_i}(A_i)$, take $D \overset{\ell}{\hookrightarrow} B$, such that each $\exists_{f_i}(A_i) \leq D$. Then $D \overset{\ell}{\hookrightarrow} B$ will have the property we assumed above, hence D, as a subobject, equals to B by what we proved above, completing the proof.

(ii) $\Rightarrow$ (i): This is the direction where we have to use the existence of stable images and sups in an essential way. Assume (ii) and assume that the g_i are given as in 3.3.2(i). Consider the formula $\phi(b,c)$:

$$\bigvee\nolimits_{i\in I} \exists a_i (f_i a_i \approx b \wedge g_i a_i \approx c).$$

By assumption, the subobject $R = [\phi(b,c)] \hookrightarrow B \times C$ exists. We claim that it is functional, "with domain B and codomain C".

First notice that the assumption $g_i e_i^j = g_j e_j^i$ can be equivalently written as the truth of the sequent

$$f_i a_i \approx f_j a_j \Rightarrow g_i a_i \approx g_j a_j.$$

This can be easily seen, on the basis of our remarks in the previous chapter relating fiber products and formulas. Consider

$$f_i a_i \approx b \wedge g_i a_i \approx c \wedge f_j a_j \approx b \wedge g_j a_j \approx c' \Rightarrow c \approx c'.$$

Since this is a consequence of the previous sequent in ordinary models, by 3.1.1, it is true in $\mathcal{R}$. Using the stability of images, by 3.1.4 and 3.2.7 (or, by the validity of the rule ($\exists \Rightarrow$), c.f. 3.2.8), we infer

$$\exists a_i (f_i a_i \approx b \wedge g_i a_i \approx c) \wedge f_j a_j \approx b \wedge g_j a_j \approx c' \Rightarrow c \approx c'$$

and then again

$$\exists a_i (f_i a_i \approx b \wedge g_i a_i \approx c) \wedge \exists a_j (f_j a_j \approx b \wedge g_j a_j \approx c') \Rightarrow c \approx c'.$$

By the definition of sups, prefixing $\bigvee_{i\in I}$ to the whole left side, what we get is still valid. By stability and 3.2.6, and doing the same with $\bigvee_{j\in I}$, we obtain

$$\bigvee\nolimits_{i\in I} \exists a_i (f_i a_i \approx b \wedge g_i a_i \approx c) \wedge \bigvee\nolimits_{j\in I} \exists a_j (f_j a_j \approx b \wedge g_j a_j \approx c') \Rightarrow c \approx c'.$$

i.e., $\underset{\sim}{R}bc \wedge \underset{\sim}{R}bc' \Rightarrow c \approx c'$, showing the first part of the claim that R is functional. For later reference, we note that the truth of the sequent

$$\bigvee\nolimits_{i\in I} \exists a_i (f_i a_i \approx b \wedge g_i a_i \approx c) \wedge f_j a_j \approx b \Rightarrow g_j a_j \approx c \tag{1}$$

is contained in our previous arguments.

Turning to the second part, notice first of all

$$\Rightarrow \exists c (g_i a_i \approx c).$$

Hence by 3.2.7, 3.1.7 and 3.1.10, we successively have that

$$\begin{array}{rcl} f_i a_i \approx b & \Rightarrow & \exists c[f_i a_i \approx b \wedge g_i a_i \approx c] \\ \exists a_i\ f_i a_i \approx b & \Rightarrow & \exists a_i\ \exists c[f_i a_i \approx b \wedge g_i a_i \approx c] \\ \exists a_i\ f_i a_i \approx b & \Rightarrow & \exists c\ \exists a_i[f_i a_i \approx b \wedge g_i a_i \approx c] \\ \bigvee_{i\in I} \exists a_i\ f_i a_i \approx b & \Rightarrow & \exists c \bigvee_{i\in I} \exists a_i[f_i a_i \approx b \wedge g_i a_i \approx c] \end{array}$$

are all true. The last left hand side, when interpreted, equals to the subobject

$$\bigvee_{i\in I} \exists_{f_i}(A_i),$$

which equals to B by assumption. Thus we have

$$\Rightarrow \exists c\ Rbc$$

as desired.

Having that R is functional, by 2.4.4 we have $g : B \to C$ such that

$$gv \approx c \ \Leftrightarrow\ \bigvee_{i\in I} \exists a_i[f_i a_i \approx b \wedge g_i a_i \approx c]$$

is true (in $\mathcal{R}$). The required equality $g_j = gf_j$ is equivalent to saying

$$gb \approx c \wedge f_j a_j \approx b \Rightarrow g_j a_j \approx c$$

which is equivalent to (1) above.

Finally, we have to show the uniqueness of g. Suppose g' has the property that $g_i = g' f_i$ for $i \in I$. Then we successively have

$$\begin{array}{rcl} f_i a_i \approx b \wedge g' b \approx c & \Rightarrow & g_i a_i \approx c, \\ f_i a_i \approx b \wedge g' b \approx c & \Rightarrow & f_i a_i \approx b \wedge g_i a_i \approx c, \\ \exists a_i(f_i a_i \approx b) \wedge g' b \approx c & \Rightarrow & \exists a_i(f_i a_i \approx b \wedge g_i a_i \approx c) \\ (\bigvee_{i\in I} \exists a_i(f_i a_i \approx b)) \wedge g' b \approx c & \Rightarrow & \bigvee_{i\in I} \exists a_i(f_i a_i \approx b \wedge g_i a_i \approx c) \end{array}$$

with justifications that should by now be familiar. Since the first conjunct 'is true' (it is the full subobject B when interpreted), we obtain

$$g' b \approx c \Rightarrow gb \approx c$$

i.e. $g' = g$, as promised.

The 'moreover' part is clear on the basis of the equivalence of (i) and (ii) and the assumed stabilities. □

Corollary 3.3.3' *If $\mathcal{R}$ has stable images, then any surjective morphism is an effective epimorphism. Conversely, and effective epimorphism is always surjective (without any hypothesis).*

Proof. The first statement is contained in 3.3.3, with I a singleton. The second part is established by our proof of the implication (i) $\Rightarrow$ (ii) of 3.3.3.

As another application, we discuss disjoint sums.

Definition 3.3.4 *We say that B is the* disjoint sum *of the objects A_i ($i \in I$), $B = \coprod_{i\in I} A_i$, with canonical injections $f_i : A_i \to B$ if the following axioms hold:*

(i) $f_i a_i \approx f_i a_i' \Rightarrow a_i \approx a_i'$ *(f_i is a monomorphism)*

(ii) $\underset{\sim}{A}_i(b) \wedge \underset{\sim}{A}_j(b) \Rightarrow \perp$ *for* $i \neq j, i, j \in I$.

(iii) $b \approx b \Rightarrow \bigvee_{i\in I} \underset{\sim}{A}_i(b)$;

here $\underset{\sim}{A}_i$ *stands for* $\exists a_i(f_i a_i \approx b)$.

The definition of disjoint sum in SGA4 is as follows: the diagram

$$\begin{array}{c} \vdots \\ A_i \xrightarrow{\ f_i\ } B \\ \vdots \end{array}$$

is a coproduct (i.e., and inductive limit), the f_i are monomorphisms, and $A_i \times_B A_j$ is an initial object, for $i \neq j$, $i, j \in I$. Also, the disjoint sum $B = \coprod_{i\in I} A_i$ is *stable* if for any $C \xrightarrow{g} B$, $C = \coprod_{i\in I} g_i^{-1}(A_i)$.

Proposition 3.3.5 *Assume that $\mathcal{R}$ has stable images, stable sups indexed by the fixed set I, and also, a stable empty sup. Then $B = \coprod_{i\in I} A_i$, with canonical injections $f_i : A_i \to B$, according to Definition 3.3.4, iff it is the disjoint sum according to the definition in SGA4. Moreover, the disjoint sum $\coprod_i A_i$ is stable.*

Proof. Suppose the conditions of 3.3.4 are satisfied. By 3.3.4(ii) and 3.3.1, we have that $A_i \times_B A_j$ is an (actually strict) initial object for $i \neq j$. We have to show that

$$\begin{array}{c} \vdots \\ A_i \xrightarrow{\ f_i\ } B \\ \vdots \end{array}$$

is a coproduct. By 3.3.4(iii) and 3.3.3, the family $\langle f_i : i \in I\rangle$ is effective epimorphic. Referring to the definition of 'effective epimorphic', we see that *any* system $\langle g_i : i \in I\rangle$ of morphisms $g_i : A_i \to C$ satisfies the 'compatibility condition' in 3.3.2; this is because each f_i is a monomorphism (by (i)) and $A_i \times_B A_j$ is an initial object for $i \neq j$. Hence, for any such $\langle g_i : i \in I\rangle$ there is a unique $g : B \to C$ such that $g_i = gf_i$ as required.

The converse is left to the reader. $\square$

Definition 3.3.6 *A subobject $R \hookrightarrow X \times X$ is called and* equivalence relation (on X) *if the following are true in $\mathcal{R}$*

$$\begin{aligned} &\Rightarrow \quad Rxx \\ Rxx' &\Rightarrow \quad Rx'x \\ Rxx' \wedge Rx'x'' &\Rightarrow \quad Rxx'' \end{aligned}$$

Remark Using the arguments given for 3.1.1, it is easy to see that this definition is equivalent to the one given in SGA4, which we don't repeat here (c.f. Chapter 1).

Definition 3.3.7 *Let $R \hookrightarrow X \times X$ be an equivalence relation. We call a morphism $X \xrightarrow{p} Y$ a quotient of R if the following are satisfied*

$$\begin{array}{lrcl} \text{(i)} & & \Rightarrow & \exists x(px \approx y)(p \text{ is surjective}) \\ \text{(ii)} & Rxx' & \Rightarrow & px \approx px'. \end{array}$$

In SGA4, we find the following definition: Given an equivalence relation $R \overset{i}{\hookrightarrow} X \times X$, or $R \underset{\pi_2}{\overset{\pi_1}{\rightrightarrows}} X$ with $\pi_j = p_j \circ i$, $p_1, p_2 : X \times X \to X$ the canonical projections, R is called *effective* if there is a commutative diagram

$$\begin{tikzcd} R \arrow[r, "\pi_1"] \arrow[d, "\pi_2"'] & X \arrow[d, "p"] \\ X \arrow[r, "p"'] & Y \end{tikzcd}$$

which is cartesian (a pullback), and in which p is an effective epimorphism. Also, R is called *stable effective* if p is a stable effective epimorphism.

Proposition 3.3.8 *Every effective equivalence relation has a quotient. Conversely, if $\mathcal{R}$ has stable images, then every equivalence relation having a quotient is stable effective.*

PROOF. Immediate by 3.3.3'. □

Proposition 3.3.9 *Assume the equivalence relations in $\mathcal{R}$ are stable effective, in the sense of* SGA4 *(c.f. after* 3.3.7*). Then $\mathcal{R}$ has stable images.*

PROOF. Let $A \xrightarrow{f} B$ be an arbitrary morphism. Let $R \hookrightarrow A \times A$ be the subobject $[fa \approx fa']$. In other words, R is $A \times_B A \overset{\pi_1}{\underset{\pi_2}{\rightrightarrows}} A$ with π_1, π_2 the canonical projections. Clearly, R is an equivalence relation (e.g. by 3.1.1). Let $A \xrightarrow{p} C$ be a quotient of R, by 3.3.8 p is surjective and since R is *stable* effective, p is stable surjective. It follows that

$$\Rightarrow \exists a\ pa \approx c \tag{2}$$

holds in $\mathcal{R}$ and $\exists a(pa \approx c)$ is a stable formula. By the definitions involved, we also have

$$pa_1 \approx pa_2 \ \Leftrightarrow\ fa_1 \approx fa_2. \tag{3}$$

Since p is an effective epimorphism, and

$$\begin{tikzcd} A \times_B A \arrow[r, "\pi_1"] \arrow[d, "\pi_2"'] & A \arrow[d, "p"] \\ A \arrow[r, "p"'] & C \end{tikzcd}$$

is a pullback, by the definition of 'effective epimorphism' it follows that there is $i : C \to B$ such that $f = ip$, we have

$$pa_1 \approx c_1 \wedge pa_2 \approx c_2 \wedge ic_1 \approx ic_2 \Rightarrow c_1 \approx c_2$$

(c.f. 3.1.1). By the stability of the formula $\exists a(pa \approx c)$ and (1), we successively have that the following hold in $\mathcal{R}$:

$$\begin{aligned} \exists a_1(pa_1 \approx c) \wedge pa_2 \approx c_2 \wedge ic_1 \approx ic_2 \quad &\Rightarrow \quad c_1 \approx c_2 \\ pa_2 \approx c_2 \wedge ic_1 \approx ic_2 \quad &\Rightarrow \quad c_1 \approx c_2 \\ \exists a_2(pa_2 \approx c_2) \wedge ic_1 \approx ic_2 \quad &\Rightarrow \quad c_1 \approx c_2 \\ ic_1 \approx ic_2 \quad &\Rightarrow \quad c_1 \approx c_2 \end{aligned}$$

(c.f. 3.2.7). The last fact means i is a monomorphism. □

Proposition 3.3.10 *Assume that the equivalence relations in $\mathcal{R}$ are stable effective. Assume that $\mathcal{R}$ has coproducts $\coprod_{i \in I} A_i$ indexed by the fixed set I. Then $\mathcal{R}$ has sups*

$\bigvee_{i\in I} A_i$ of subobjects A_i, indexed by I, of a given object. If the coproducts indexed by I are stable, then the sups indexed by I are stable as well.

PROOF. Let $A_i \overset{\alpha_i}{\hookrightarrow} X$ be subobjects, $i \in I$. Let $E = \coprod_{i\in I} A_i$ be the disjoint sum of the objects A_i, with canonical injections $j_i : A_i \to E$. Since $\coprod_{i\in I} A_i$ is a coproduct, there is a unique $f : E \to X$ such that $\alpha_i = f j_i$, for $i \in I$.

$$\begin{tikzcd}
 & & A_i \arrow[dll, bend right, "j_i"'] \arrow[d, hook, "\alpha_i"'] \arrow[dr, "\alpha_i'"] & \\
E = \coprod_{i\in I} A_i \arrow[r, two heads, "g"] \arrow[rr, bend left, "f"] \arrow[rrr, bend right, "h"'] & C \arrow[r, hook, "\gamma"'] & X & D \arrow[l, hook', "\delta"']
\end{tikzcd}$$

Using 3.3.9, let $g : E \to C$ be surjective and $C \xrightarrow{\gamma} X$ a monomorphism such that $f = \gamma g$. We claim that the subobject $C \overset{\gamma}{\hookrightarrow} X$ is the sup of the subobjects $A_i \overset{\alpha_i}{\hookrightarrow} X$, $C = \bigvee_{i\in I} A_i$. To verify this, let the monomorphism $D \overset{\delta}{\hookrightarrow} X$ be such that each α_i factors through δ, $\alpha = \delta\alpha_i'$ $(i \in I)$. By the coproduct property of E, there is $h : E \to D$ such that $\alpha_i' = h j_i$ for $i \in I$. Then, for $f' = \delta h$, we have $f' j_i = \alpha_i$ for $i \in I$, hence $f = f' = \delta h$. Since $\exists_f(E) = C \overset{\gamma}{\hookrightarrow} X$ and $E \xrightarrow{f} X$ factors through δ, it follows that the subobject $C \overset{\gamma}{\hookrightarrow} X$ is $\leq$ the subobject $D \overset{\delta}{\hookrightarrow} X$. This shows that, indeed, $C = \bigvee_{i\in I} A_i$. Verifying stability is left to the reader.

Proposition 3.3.11 *Assume that $\mathcal{R}$ has stable images, stable finite sups, quotients of all its equivalence relations and disjoint sums. Then every epimorphism is surjective.*

PROOF. Suppose that $A \xrightarrow{f} B$ is an epimorphism. Let $C = B \amalg B$ be the disjoint sum, with canonical injections $i_1, i_2 : B \to C$. Let $\underset{\sim}{A}_1(c)$, $\underset{\sim}{A}_2(c)$ denote the formulas $\exists a(i_1 f a \approx c)$, $\exists a(i_2 f a \approx c)$, respectively. Define the subobject $R \hookrightarrow C \times C$ by

$$\underset{\sim}{R}(c, c') \;\Leftrightarrow\; c' \vee ((\underset{\sim}{A}_1(c) \vee \underset{\sim}{A}_2(c)) \wedge (\underset{\sim}{A}_1(c') \vee \underset{\sim}{A}_2(c'))).$$

Using the distributivity of the subobject-lattices in $\mathcal{R}$, it is quite easy to check that R is an equivalence relation. Let $C \xrightarrow{p} D$ be a quotient of R; hence

$$\underset{\sim}{R}(c, c') \;\Leftrightarrow\; pc \approx pc'. \tag{4}$$

Notice that by the definition of R, $\Rightarrow \underset{\sim}{R}(i_1 f a, i_2 f a)$ is true, hence $p i_1 f = p i_2 f$. Since f is an epimorphism, $p i_1 = p i_2$.

Using (4) again, $p i_1 = p i_2$ implies that

$$\Rightarrow \underset{\sim}{R}(i_1 b, i_2 b)$$

holds. Since the sum $C = B \amalg B$ is disjoint, $i_1 b \approx i_2 b$ is false, i.e., its interpretation is the zero subobject. It follows that

$$\Rightarrow (\underset{\sim}{A}_1(i_1 b) \vee \underset{\sim}{A}_2(i_1 b))$$

holds. Again by disjointness, $[A_2(i_1 b)] = 0_B$, i.e., we have $\Rightarrow \underset{\sim}{A}_1(i_1 b)$, i.e. $\Rightarrow \exists a(i_1 f a \approx i_1 b)$. Since i_1 is a monomorphism, $\Rightarrow \exists a(f a \approx b)$ follows which is equivalent to saying that f is surjective. $\square$

§4 Logical categories

In this section we introduce those kinds of categories that best represent first order logic, both finitary and infinitary.

Definition 3.4.1 *The category $\mathcal{R}$ is called* logical *if the following are satisfied:*

(i) *$\mathcal{R}$ has finite left limits,*

(ii) *$\mathcal{R}$ has stable finite sups,*

(iii) *$\mathcal{R}$ has stable images.*

In the course of this work, it should become clear that the notion of a logical category can be regarded as the basic notion in a categorical formulation of logic. In Chapter 8 we will show that, in a sense made precise there, logical categories are *the same* as theories in a finitary coherent $L^g_{\omega\omega}$. There are good reasons why it is better to take $L^g_{\omega\omega}$ as basic rather than $L_{\omega\omega}$; we will discuss them below. Notice that **Set**, the category of sets, is logical.

Definition 3.4.2 *A functor $F\colon\mathcal{R}\to\mathcal{S}$ between logical categories is called* logical *if it preserves finite left limits, finite sups and images.*

Remarks E.g. the phrase "F preserves finite sups" means that whenever A_i $(i\in I)$ are subobjects of X in $\mathcal{R}$, I is a finite set, then the subobject $F(\bigvee_{i\in I}A_i)$ of $F(X)$ in $\mathcal{S}$ is the sup of the subobjects $F(A_i)$, $i\in I$:

$$F(\textstyle\bigvee^{(\mathcal{R})}_{i\in I}A_i)=\bigvee^{(\mathcal{S})}_{i\in I}F(A_i).$$

The notion of a logical functor is the categorical rendering of the notion of a model, and of an interpretation of a theory in another theory, at the same time.

The notion of *pretopos* was introduced by Grothendieck, SGA4, Exposé VI, in connection with coherent toposes. We will return to this connection in Chapter 9. One of the main results of this work, Theorem 7.1.8 in Chapter 7, shows that pretopoi are characterized by an interesting abstract property. On the other hand, in Chapter 8 we show that every logical category can be 'completed' to a pretopos. First we state the definition of a pretopos in our preferred terminology and the proceed to show that it is equivalent to Grothendieck's definition.

Definition 3.4.3 *A* pretopos *$\mathcal{P}$ is a logical category having the additional properties:*

(iv) *$\mathcal{P}$ has quotients of equivalence relations,*

(v) *$\mathcal{P}$ has finite disjoint sums.*

Remark In more detail, (iv) means that every equivalence relation in $\mathcal{P}$ has a quotient in $\mathcal{P}$, (v) means that every finite family of objects in $\mathcal{P}$ has a disjoint sum in $\mathcal{P}$.

Grothendieck gives the following definition (SGA4, Exposé VI, 3.11) $\mathcal{P}$ is a pretopos if (a) it has finite left limits, (b) $\mathcal{P}$ has stable finite disjoint sums (in the sense of SGA4, c.f. above), (c) the equivalence relations of $\mathcal{P}$ are effective, (d) every epimorphism in $\mathcal{P}$ is stable effective, (e) $\mathcal{P}$ is small. We will ignore the smallness condition and show that Grothendieck's definition (minus (e)) is equivalent to ours.

First assume that $\mathcal{P}$ is a pretopos according to 3.4.3. Then (b), (c) and (d) follow, respectively, by 3.3.5, 3.3.8, 3.3.11 and 3.3.3'. Conversely, if $\mathcal{P}$ satisfies (a) to (d), then $\mathcal{P}$ has *stable* effective equivalence relations (by (c) and (d) jointly). Hence, by 3.3.9 and 3.3.10 $\mathcal{P}$ is logical. Finally, as stated in 3.3.5 and 3.3.8, the requirements regarding disjoint sums and equivalence relations are the same if we formulate them as in 3.4.3, or as in (b) and (c).

Notice that according to these arguments, (d) can be relaxed to read: "every *effective* epimorphism is stable effective" without changing the notion of pretopos. Actually, another inessential change is to drop (d) entirely but strengthening (c) requiring *stable* effectiveness of equivalence relations. This last version should be considered the "SGA4 definition" in our opinion.

Notice that, as an immediate consequence of our definition of the notions, a logical functor $F : \mathcal{P} \to \mathcal{S}$ automatically preserves finite disjoint sums, equivalence relations and quotients of equivalence relations.

Next we give "infinitary generalizations" of our previous notion, finally arriving at (Grothendieck) topoi and geometric morphisms. Let κ be an infinite regular cardinal number.

Definition 3.4.4 *A category $\mathcal{R}$ is called* κ-logical *if it satisfies:*

(i) *$\mathcal{R}$ has finite left limits*

(ii)$_\kappa$ *$\mathcal{R}$ has stable $< \kappa$-sups, i.e. every family of power $< \kappa$ of subobjects of a fixed object has a stable sup.*

(iii) *$\mathcal{R}$ has stable images.*

Definition 3.4.5 *A functor $F : \mathcal{R} \to \mathcal{S}$ between κ-logical categories is called* κ-logical *if it preserves finite left limits, $< \kappa$-sups and images.*

Definition 3.4.6 *A* κ-pretopos *is a κ-logical category having the additional properties:*

(iv) *$\mathcal{P}$ has quotients of equivalence relations,*

(v)$_\kappa$ *$\mathcal{P}$ has disjoint $< \kappa$-sums.*

Putting $\kappa = \aleph_0$, we obtain the notion of (ordinary) pretopos.

If we remove the restrictions on the size of sups and sums we arrive at

Definition 3.4.7 *A* ∞-pretopos *is a category $\mathcal{P}$ satisfying* (i), (ii)$_\infty$, (iii), (iv), (v)$_\infty$ *with* (i), (ii) *and* (iv) *from* 3.4.5 *and* 3.4.6*; and* (iii)$_\infty$, (v)$_\infty$ *as follows:*

(ii)$_\infty$ *$\mathcal{P}$ has stable sups of arbitrary sets of subobjects of a given object,*

(v)$_\infty$ *$\mathcal{P}$ has arbitrary disjoint sums (according to Definition* 3.3.4*).*

This notion is practically the same as the notion of topos.

Proposition 3.4.8 *A category $\mathcal{P}$ is a (Giraud) topos (c.f.* 1.4.3 *if and only if it is an ∞-pretopos and has a set of generators.*

The proof is immediate on the basis of 3.3.5, 3.3.8 (for the 'if' direction) and 3.3.9, 3.3.10 (for the 'only if' direction).

Definition 3.4.9 *A functor $M : \mathcal{E}_1 \to \mathcal{E}_2$ between ∞-pretopoi is* ∞-logical *if it preserves finite left limits, arbitrary sups and images.*

Proposition 3.4.10 *A functor $M : \mathcal{E}_1 \to \mathcal{E}_2$ between (Giraud) topoi (or ∞-pretopoi) is continuous (or M is an $\mathcal{E}_2$-model of $\mathcal{E}_1$, c.f. Chapter* 1*, Section* 3*) if and only if M is ∞-logical.*

Proof. Suppose M is continuous. Then M preserves effective epimorphic families. So it preserves images. If $\bigvee_{i \in I} A_i = B \to X$, then $(A_i \to B)_{i \in I}$ is an effective epimorphic family (c.f. e.g. 3.3.3). Hence M preserves sups. It follows that M is ∞-logical. Conversely, if M is ∞-logical then by 3.3.3 it immediately follows that M preserves effective epimorphic families. $\square$

Finally, we fill in the holes that were left in Chapter 1.

Proposition 3.4.11 (= 1.4.6) *In an ∞-pretopos (in a Giraud topos), every epimorphic family is stable effective.*

PROOF. Suppose that $(A_i \xrightarrow{f_i} A)_{i\in I}$ is an epimorphic family. Consider the subobject $B = \bigvee_{i\in I} \exists_{f_i}(A_i) \overset{j}{\hookrightarrow} A$. With $g_i : A_i \to B$ such that $ig_i = f_i$ (for $i \in I$), it is easy to see that $B = \bigvee_{i\in I} \exists_{g_i}(A_i)$. Hence by 3.3.3, $(A_i \xrightarrow{g_i} B)_{i\in I}$ is an effective epimorphic family. Using only the epimorphic property of $(A_i \xrightarrow{f_i} A)_{i\in I}$, we can directly verify that $B \xrightarrow{j} A$ is an epimorphism.

$$\begin{tikzcd} A_i \arrow[r, "g_i"'] \arrow[rr, bend left, "f_i"] & B \arrow[r, "j"'] & A \arrow[r, shift left] \arrow[r, shift right] & C \end{tikzcd}$$

By 3.3.11, j is surjective and since it is a monomorphism, j is an isomorphism. Hence $(A_i \xrightarrow{f_i} A)_{i\in I}$ is an effective epimorphic family, since $(A_i \xrightarrow{g_i} B)_{i\in I}$ is such.

The stability of $(A_i \xrightarrow{f_i} A)_{i\in I}$ is asserted, in fact, in 3.3.3. □

Proposition 3.4.12 (= 1.3.13) *If a functor $M : \mathcal{E}_1 \to \mathcal{E}_2$ between topoi preserves finite left limits and all small inductive limits, then M preserves (effective epimorphic) families.*

PROOF. From the hypothesis is follows that M preserves disjoint sums (defined in the sense of SGA4). Moreover, if $R \underset{p_2}{\overset{p_1}{\rightrightarrows}} A$ is an equivalence relation in $\mathcal{E}_1$ and p is a quotient of R, i.e.,

$$\begin{tikzcd} & A \arrow[dr, "p"] & \\ R \arrow[ur, "p_1"] \arrow[dr, "p_2"'] & & B \\ & A \arrow[ur, "p"'] & \end{tikzcd}$$

is both a pullback and a push out, then the same is true of

$$\begin{tikzcd} & MA \arrow[dr, "Mp"] & \\ MR \arrow[ur, "Mp_1"] \arrow[dr, "Mp_2"'] & & MB \\ & MA \arrow[ur, "Mp"'] & \end{tikzcd}$$

en $\mathcal{E}_2$. From these two facts it follows, through the ways images and sups are constructed from quotients of equivalence relations and disjoint sums (coproducts) in 3.3.9 and 3.3.10, that M preserves images and sups. Finally, by 3.3.3, it follows that M preserves effective epimorphic families. □

On Sublemma 1.4.10, we should say this. Let $I : \mathcal{R} \to \mathcal{S}$ be conservative and left exact. If A and B are subobjects of X in $\mathcal{R}$, then from $IA \leq IB$ (as subobjects of IX) in $\mathcal{S}$ it follows that $A \leq B$; this is immediately seen by considering the monomorphisms $A \times_X B \to A$. Next, look at the sequents defining equivalence relations in 3.3.6. Using the left exactness of I as well as the last statement, we have that if these sequents are true for the image $IR \hookrightarrow IX \times IX$, then they are true for $R \hookrightarrow X \times X$, what is the claim in 1.4.10 about equivalence relations. Faithfulness follows by considering a sequent expressing the equality of two morphisms.

Lemma 3.4.13 *Suppose F is a κ-logical morphism between κ-pretopoi (κ is an infinite regular cardinal or $\kappa = \infty$). Then F preserves disjoint $< \kappa$ sums.*

This is immediate on the basis of the "logical" definition 3.3.4 of disjoint sum.

§5 Summary of the two main facts

Recall that we called the two main facts in applying logic to categories: the first, summarized mainly in 2.4.5, is the fact that properties of diagrams could be expressed by using formulas (note that additional such facts were proved in this Chapter); the second, 3.2.8, is the soundness of certain ordinary 'complete' formal system with respect to the categorical interpretation. Both these facts were formulated in the context of the canonical language of a given category and the canonical interpretation of this language. First we show how they generalize to a much more general situation.

Let us start with an arbitrary category $\mathcal{S}$ (with finite left limits), an 'arbitrary' language L and an interpretation of L in $\mathcal{S}$, i.e. an $\mathcal{S}$ structure of type L

$$M : L \to \mathcal{S}.$$

Let $L_{\mathcal{S}}$ be the canonical language of $\mathcal{S}$. We can translate the logic $L_{\infty\omega}$ into $(L_{\mathcal{S}})_{\infty\omega}$ such that M 'becomes' the canonical interpretation of $L_{\mathcal{S}}$ as follows. First of all, regard M as a map between languages

$$M : L \to L_{\mathcal{S}}.$$

In more detail, a sort s of L is mapped to the sort $M(s)$ of $L_{\mathcal{S}}$ (an object of $\mathcal{S}$), and operation symbol $f : s_1 \times \cdots \times s_n \to s$ of L is mapped to the operation symbol $M(f) : M(s_1) \times \cdots \times M(s_n) \to M(s)$ of $L_{\mathcal{S}}$ (which is the same as the morphism $M(f)$, together with the specification of the product $M(s_1) \times \cdots \times M(s_n)$) and similarly for predicate symbols. In connection with the *translation* $M : L \to L_{\mathcal{S}}$ we'll write s^M, f^M etc. for $M(s)$, $M(f)$, respectively.

Now, this 'translation' of L into $L_{\mathcal{S}}$ carries over naturally to the corresponding logics. First of all, the variables of L should be mapped in a 1-1 way into the variables of $L_{\mathcal{S}}$ such that the free (bound) variables of sort s are mapped to free (bound) variables of sort s^M. To simplify notation, let us *identify* variables of sort s with those of s^M; though, strictly speaking, one has to be careful since for distinct s and s' s^M might be equal to $(s')^M$. Once this identification is made, we can define the *translate* ϕ^M of a formula of $L_{\infty\omega}$ to be a formula in $(L_{\mathcal{S}})_{\infty\omega}$ by simply replacing each symbol of L by its M-image in $L_{\mathcal{S}}$. Of course, this involves a translate t^M of terms t of L as well. E.g. we will have

$$\begin{aligned} (ft_1 \cdots t_n)^M &\underset{\mathrm{df}}{=} (f^M)t_1^M \cdots t_n^M, \\ (Pt_1 \cdots t_n)^M &= (P^M)t_1^M \cdots t_n^M \\ (\bigvee \Sigma)^M &= \bigvee \{\phi^M : \phi \in \Sigma\}, \end{aligned}$$

etc. We also have the natural notion of the M-translate σ^M of a sequent σ. Now notice the following *trivial fact:*

Proposition 3.5.1 *The interpretation of a formula ϕ in $L_{\infty\omega}$ by M in $\mathcal{S}$ coincides with the canonical interpretation of its translate ϕ^M by the canonical interpretation of $L_{\mathcal{S}}$. I.e.,*

$$M_{\vec{x}}(\phi) = [\phi^M]_{\vec{x}}$$

in the sense that either side is defined once the other is and they are equal (as subobjects of the object $M(\vec{x})$ in $\mathcal{S}$). As a consequence, for a sequent σ in $L_{\infty\omega}$, $M \models \sigma$ iff σ^M is true in the canonical interpretation of $L_{\mathcal{S}}$.

We want to apply this device to the following situation. Let $M : \mathcal{R} \to \mathcal{S}$ be a functor between the categories $\mathcal{R}$ and $\mathcal{S}$. We are interested in whether M preserves e.g.

finite left limits, or images, etc.; in general, whether, for a given diagram in $\mathcal{R}$ with a certain property, the image of the diagram by M in $\mathcal{S}$ still has the same property (now understood in $\mathcal{S}$). We say that a property P of a diagram D is *described by axioms* if a result of the sort of the ones listed in 2.4.5 is true, in an arbitrary category. This means that there is a set of axioms, call it A, such that the diagram D in $\mathcal{R}$ has P iff (each axiom in) A is true in the canonical interpretation of $L_{\mathcal{R}}$. Actually, we should talk about $A_{\mathcal{R}}$, the specialization of the general form of the axioms in the category $\mathcal{R}$. Notice the following fact: (*) Given the functor $M : \mathcal{R} \to \mathcal{S}$, the set $A_{\mathcal{R}}$ of axioms formulated for D in $\mathcal{R}$, when *translated* vie the above translation $(\cdot)^M$ of $L_{\mathcal{R}}$ into $L_{\mathcal{S}}$ becomes the set $A_{\mathcal{S}}$, the 'same' axioms formulated for the image $M(D)$. This can be seen by inspecting each item referred to in 2.4.5.

Next, let us emphasize that any functor $M : \mathcal{R} \to \mathcal{S}$ *is* an $\mathcal{S}$-structure of type $L_{\mathcal{R}}$.

Conversely, let M be an $\mathcal{S}$-structure of type $L_{\mathcal{R}}$,

$$M : L_{\mathcal{R}} \to \mathcal{S}$$

M is *almost* a functor $\mathcal{R} \to \mathcal{S}$; namely it maps objects of $\mathcal{R}$ to objects of $\mathcal{S}$ and morphisms of $\mathcal{R}$ into morphemes of $\mathcal{S}$, with the appropriate domains and codomains. If we require that M also *satisfies* the axioms in groups 1 and 2 before 2.4.5 (axioms for identity and commutative diagram), then (and only then) M becomes a functor $\mathcal{R} \to \mathcal{S}$. This means that the property of $M : L_{\mathcal{R}} \to \mathcal{S}$ of being a functor is described by axioms.

Metatheorem 3.5.2 *Suppose D is a diagram in $\mathcal{R}$ with property P where P is a property described by a set of axioms $A_{\mathcal{R}}$, formulated in $L_{\mathcal{R}}$. Let M be an $\mathcal{S}$-structure $M : L_{\mathcal{R}} \to \mathcal{S}$. Then M satisfies all axioms in $A_{\mathcal{R}}$ if and only if M* preserves *the property P for D, i.e. the image $M(D)$ of D by M in $\mathcal{S}$ has property P understood in $\mathcal{S}$.*

PROOF. $M(D)$ has property P off (**) the axioms in $A_{\mathcal{S}}$ are true in the canonical interpretation of $L_{\mathcal{S}}$. Now, by fact (*) the axioms in $A_{\mathcal{S}}$ are exactly the M-translates of the axioms in $A_{\mathcal{R}}$. Hence, by 3.5.1, (**) is equivalent to saying that M, an $\mathcal{S}$-structure of type $L_{\mathcal{R}}$, satisfies $A_{\mathcal{R}}$. □

As an example, we mention the following. Suppose we have the product diagram $A \xleftarrow{f} C \xrightarrow{g} B$ in $\mathcal{R}$ and $M : \mathcal{R} \to \mathcal{S}$ is a functor. Then M preserves the given product off M (as an $\mathcal{S}$-structure of type $L_{\mathcal{R}}$) satisfies the "axioms for product" stated for the given diagram.

We will now apply our discussion to a particular case. Let $\mathcal{R}$ be a logical category. We define the *internal theory* $T_{\mathcal{R}}$ of $\mathcal{R}$ (as a logical category) as given by the following axioms in $(L_{\mathcal{R}})_{\omega\omega}$:

(i) the 'axioms of category', i.e., (a) each axiom for identity (Group 1 before 2.4.5) for every identity map $A \xrightarrow{f} A$, $f = \mathrm{id}_A$ in $\mathcal{R}$, (b) each axiom for commutative diagram (Group 2), for each commutative diagram

$$\begin{tikzcd} & B \arrow[dr, "g"] & \\ A \arrow[ur, "f"] \arrow[rr, "h"'] & & C \end{tikzcd}$$

in $\mathcal{R}$;

(ii) the 'axioms for left limits', i.e. those in groups 3, 4, 5 and 6 (3 is superfluous), corresponding to diagrams in $\mathcal{R}$ that do indeed have the respective property;

(iii) axioms for finite sups, i.e. all those acceding to Group 8, with finite sets I, for true sups in $\mathcal{R}$;

(iv) axioms for images, i.e. all those according to Group 9, for true images in $\mathcal{R}$.

Notice that each axiom in $T_{\mathcal{R}}$ is true in $\mathcal{R}$, by 2.4.5. Recall that every functor $M:\mathcal{R}\to\mathcal{S}$ is a structure of type $L_{\mathcal{R}}$ and conversely, every $\mathcal{S}$-structure M of type $L_{\mathcal{R}}$ is a functor $M:\mathcal{R}\to\mathcal{S}$ once it satisfies axioms (i).

The next theorem is the full formulation of the 'first main fact' for the 'logical' case:

Theorem 3.5.3 *Given the logical categories $\mathcal{R}$ and $\mathcal{S}$, the logical functors $M:\mathcal{R}\to\mathcal{S}$ are exactly the $\mathcal{S}$-models of the internal theory $T_{\mathcal{R}}$ of $\mathcal{R}$.*

The theorem is immediate on the basis of our lengthy discussion above.

We have a similar result for the "κ-logical" case. Now the axioms include all axioms for $<\kappa$-sups, call the (ii)$_\kappa$.

Notice that the internal theory of a logical category is a *finitary coherent* theory. The internal theory of a κ-logical category is formulated in $L^g_{\infty\omega}$, actually $L^g_{\infty\omega}$ (the last notion refers to the fact that all disjunctions are cover sets of power $<\kappa$).

Next we turn to a discussion of the use of the second main fact, Theorem 3.2.8. Let $M:L\to\mathcal{S}$ be an $\mathcal{S}$-structure of type L, F a fragment of $L_{\infty\omega}$. We say that F is *stable (distributive) with respect to* M if the translate ϕ^M in $L_{\mathcal{S}}$ of every formula ϕ of F is stable (distributive) as defined above. (E.g., if $\mathcal{S}$ is a logical category, F is $L^g_{\omega\omega}$, then F is stable with respect to any $M:L\to\mathcal{S}$.)

Theorem 3.5.4 (i) *Assume $\mathcal{S}$ has stable finite sups. Assume the fragment F is distributive with respect to $M:L\to\mathcal{S}$ and T is a theory in F all of whose elements are true in $\mathcal{S}$. Then if $T\overset{1}{\vdash}\sigma$, then $M\models\sigma$.*

(ii) *Assume F is a coherent fragment, F is stable with respect to $M:L\to\mathcal{S}$ and T is as before. Then, if $T\overset{2}{\vdash}\sigma$, then $M\models\sigma$.*

The proof is immediate, using 3.5.1 and 3.2.8. □

Finally, we show a typical application of the two main facts, together with a theorem, the Gödel completeness theorem, borrowed from logic. In Chapter 5, we will prove

Completeness Theorem *Let T be a finitary coherent theory (in $L^g_{\omega\omega}$), σ a sequent of $L^g_{\omega\omega}$. Suppose that σ is* not *a consequence of T, $T\overset{2}{\nvdash}\sigma$. Then there is a model M of T that is not a model of σ.*

We use Göde's completeness theorem to prove

Theorem 3.5.5 Completeness Theorem for logical categories *(Deligne-Joyal) Let $\mathcal{R}$ be a small logical category.*

(i) *If A and B are two subobjects of X in $\mathcal{R}$ such that $A\not\leq B$ then there is a logical functor $M:\mathcal{R}\to$* **Set** *such that $M(A)\not\leq M(B)$ (the latter being a subset of $M(X)$).*

(ii) *There is a set I (of the power* $\max(\aleph_0,\mathrm{card}(\mathrm{Ob}(\mathcal{R})),\hom_{\mathcal{R}}(A,B)\ (A,B\in\mathrm{Ob}(\mathcal{R})))$ *and a logical functor*

$$M:\mathcal{R}\to\mathbf{Set}^I$$

that is faithful.

Proof. (AD (i)): Let L be $L_{\mathcal{R}}$, the canonical language of $\mathcal{R}$. Let $\underset{\sim}{A}(x)$, $\underset{\sim}{B}$ be formulas in $L^g_{\omega\omega}$ such that $[\underset{\sim}{A}(x)]=A$, $[\underset{\sim}{B}(x)]=B$. Notice that the assumption means that $\underset{\sim}{A}(x)\Rightarrow\underset{\sim}{B}(x)$ does not hold in the canonical interpretation, i.e., in $\mathcal{R}$. Let $T=T_{\mathcal{R}}$, the internal theory of $\mathcal{R}$. We have that the axioms of T are true in $\mathcal{R}$. We also have that $T\overset{2}{\nvdash}\underset{\sim}{A}(x)\Rightarrow\underset{\sim}{B}(x)$; otherwise by the soundness theorem (either 3.2.8, or 3.5.4) we would have that $\underset{\sim}{A}(x) \Rightarrow \underset{\sim}{B}(x)$ would be true in $\mathcal{R}$, contrary to the assumption. Hence, by the completeness theorem, there is a model M of T, $M : T \to \mathbf{Set}$ such that $M \not\models \underset{\sim}{A}(x) \Rightarrow \underset{\sim}{B}(x)$. Hence, by the completeness theorem, there is a model M of T, $M : T \to \mathbf{Set}$ such that $M \not\models \underset{\sim}{A}(x) \Rightarrow \underset{\sim}{B}(x)$. Using 3.5.3, M *is* a logical functor $M : \mathcal{R} \to \mathbf{Set}$. Clearly $M(A) \not\leq M(B)$ as required.

(AD (ii)): For any object X, and any pair A, B of subobjects of X such that $A \not\leq B$, form $i(X, A, B)$, and let I be the set of all such indices i. For $i = (X, A, B)$, let $M_i : \mathcal{R} \to \mathbf{Set}$ be a logical functor such that $M_i(A) \not\leq M_i(B)$. The functor $M : \mathcal{R} \to \mathbf{Set}^I$ is defined so that its i^{th} coordinate is M_i. We leave it to the reader to check that M satisfies (ii).

Remarks 1. There is a "purely categorical" proof of this theorem, c.f. a sketch in Reyes [1974] or Kock and Reyes [1977] and the proof of the related theorem of Deligne in SGA4, Exposé 6, pp. 63-72. Inspection shows however that the categorical proof is no essentially different from the usual proofs of Gödel completeness theorem (c.f. the remark in Kock and Reyes [1977] on the connection to Henkin's method).

2. The categorical completeness theorem, 3.5.5, is *equivalent* to Göde's completeness theorem. First of all, we have shown that 3.5.5 can be derived from Gödel's completeness theorem formulated for the formal system " $\models^2$ " (described in Chapter 5). Conversely, Göde's completeness for " $\models^2$ " can be inferred from 3.5.5(i) via the construction of the "associated logical category" of a theory, described in Chapter 8. Finally, we emphasize that although negation and $\to$, $\forall$ are not present in $\models^2$ or in the notion of a logical category, completeness for logical categories, i.e. for coherent logic, already contains completeness for full first order logic. The reason is that a logical functor $M : \mathcal{R} \to \mathcal{S}$ between logical categories automatically preserves every Boolean complement that exists in $\mathcal{R}$. Therefore, if we modify logical categories and logical functors to Boolean categories and Boolean functors by requiring the existence and preservation of Boolean complements, the the Boolean version of 3.5.5 is simply a special case of 3.5.5 itself. The Gödel type completeness for an appropriate formal system can now be inferred via the construction of a Boolean category associated to a theory in full finitary first order logic.

Chapter 4

Boolean and Heyting valued models

Introduction

In this chapter we discuss the interpretation of formulas of $L_{\infty\omega}$ in a certain kind of Grothendieck topos, namely the category of sheaves over a partial ordering, and as the main special case of this the category of sheaves $\mathrm{Sh}_{\mathcal{B}}$ over a complete Boolean algebra $\mathcal{B}$. D. Higgs [1973] has given an alternative description of such toposes (for the case of a complete Heyting algebra). This description will be useful for us because it identifies models in $\mathrm{Sh}_{\mathcal{B}}$ in the sense of Chapters 1 and 2 with what logicians have called a (general) $\mathcal{B}$-valued model, c.f. e.g. Mansfield [1972]. In Chapter 5, working with the formulation familiar to logicians, we present detailed proofs of completeness theorems for Boolean valued models. Using the Higgs identifications, in Chapter 6 we will be able to derive "purely categorical" formulations of "completeness theorems" from those in Chapter 5.

In Section 1, we give definitions related to Heyting valued (and in particular, Boolean-valued) models in the ordinary "logical" sense. Some refinements of he ordinary formulation are needed because of the possibility of "empty domains".

In Section 2, we describe the identifications mentioned above. Finally, in Section 3, we state some facts concerning Boolean valued models that are needed later.

§1 Heyting and Boolean valued models

Let $\mathcal{H}$ be a complete Heyting algebra (c.f. e.g. Rasiowa and Sikorski [1963], in particular, $\mathcal{H}$ may be a complete Boolean algebra). An $\mathcal{H}$-valued structure M, appropriate for a given many-sorted language L, consists of

(i) The *(partial) domains* $|M|_s$ for each sort s in L, which are arbitrary sets. The sets $|M|_s$ can be stipulated to be non-empty, and also disjoint from each other, without loss of generality.

(ii) Interpretations P^M, f^M, c^M and $\approx^M$ of all the symbols P (relation symbol), f (operation symbol), c (individual constant) in L, and $\approx$, as follows. Let P be e.g. ternary with places having the respective sorts s_1, s_2, s_3. Then $P^M : |M|_{s_1} \times |M|_{s_2} \times |M|_{s_3} \to \mathcal{H}$. f^M and c^M are defined without reference to $\mathcal{H}$, exactly as in the two-valued case (c.f. Chapter 2, Section 1). (**Remark** *there is a natural "Heyting" ("Boolean") interpretation of operations that is more general than ours, but it turns out that our more restricted notion is sufficient for our purposes.) The interpretation of equality is actually given* *separately on each partial domain* $|M|_s$. *We have* $\approx^{M,s}\ : |M|_s \times |M|_s \to \mathcal{H}$ *for each sort* s *in* L. *Usually, we omit the second superscript and write* $\approx^M$.

(iii) "Membership functions" $\|\cdot\| = \|\cdot\|_s : |M|_s \to \mathcal{H}$, for each s. $\|a\|$ serves to tell us "with what value in $\mathcal{H}$ the existence of $a \in |M_s|$ as an element of M is admitted." E.g., if $\|a\| = 0$, then we can discard a from $|M|_s$ without any essential change in M (c.f. below). It is essential that elements $a \in |M|_s$ can have all "degrees" $\in \mathcal{H}$ of existence $\|a\|$ in M. We will write $\|\vec{a}\|$ for $\|a_1\| \cdot \|a_2\| \cdots \|a_n\|$ if $\vec{a} = \langle a_1, a_2, \ldots, a_n\rangle$, (for $\beta, \gamma \in \mathcal{H}$, $\beta \cdot \gamma$ is synonymous with $\beta \wedge \gamma$). (**Remark** *strictly speaking, the membership functions are not necessary since it will turn out that they could be defined as* $\|a\| = \|a \approx a\|$, *c.f. below. We have found, however, that their separate mention is more natural.)*

(iv) There are some conditions on the above items (basically: equality axioms) that can be formulated later more conveniently.

We will use the notation $|M|_x$, or even $|M|_t$, for $|M|_s$ if x is a variable of sort s, t is a term of sort s.

Given M, an $\mathcal{H}$-valued structure for the language L, the terms of L have the obvious interpretation in M just as in the case of ordinary models. In particular, if t is a term, $x_1, \ldots, x_n$ are free variables including all the variables in t, $a_i \in |M|_{x_i}$ $(i = 1, \ldots, n)$ then

$$t^M[a_1/x_1, \ldots, a_n/x_n]$$

or simply

$$t^M[a_1, \ldots, a_n]$$

is defined and it is an element of $|M|_t$.

Next we describe the interpretation of $L_{\infty\omega}$ in M. Given a formula ϕ in $L_{\infty\omega}$, distinct free variables $x_1, \ldots, x_n$ such that each free variable of ϕ is among the x_i (but some x_i might not actually occur in ϕ) and elements $a_i \in |M|_{x_i}$, we are going to define the *$\mathcal{H}$-value of ϕ when x_i is interpreted by a_i*, in notation

$$\begin{aligned} &\|\phi[a_1/x_1, \ldots, a_n/x_n]\|_M, \quad \text{or simply} \\ &\|\phi[a_1, \ldots, a_n]\|, \quad \text{or} \\ &\|\phi[\vec{a}]\| \qquad \text{if } \ \vec{a} = \langle a_1, \ldots, a_n\rangle. \end{aligned}$$

A slight change in the definition of $\|\phi[\vec{a}]\|$ with respect to the familiar definition is effected by the presence of the "membership-values" $\|a\|$. In particular, the definition is so designed that we always have $\|\phi[\vec{a}]\| \le \|a\|$, i.e., the truth-value of a formula is no greater that the degree of existence of the interpreting elements. Accordingly, we define for an atomic formula $\phi := Pt_1 \cdots t_m$

$$\|\phi[\vec{a}]\| = \|\vec{a}\| \cdot P^M(b_1, \ldots, b_m)$$

where $b_i = t_i^M[\vec{a}]$. Similarly,

$$\|t_1 \approx t_2\ [\vec{a}]\| = \|\vec{a}\| \cdot \|b_1 \approx b_2\|$$

with b_i as before. The following inductive rules complete the definition: (on the right hand sides $\neg$, $\to$, $\wedge$ (or: $\cdot$) denote the usual operations in $\mathcal{H}$, $\bigwedge\mathcal{X}$, $\bigvee\mathcal{X}$ denote the order-theoretic infimum and supremum, respectively, of the elements in the set $\mathcal{X} \subset \mathcal{H}$)

(1) $\|\neg\phi[\vec{a}]\| = \|\vec{a}\| \cdot \neg\|\phi[\vec{a}]\|$

(2) $\|(\phi \to \psi)[\vec{a}]\| = \|\vec{a}\| \cdot (\|\phi[\vec{a}]\| \to \|\psi[\vec{a}]\|)$

(3) $\|(\bigwedge\Theta)[\vec{a}]\| = \|\vec{a}\| \cdot \bigwedge\{\|\theta[\vec{a}]\| : \theta \in \Theta\}$

(4) $\|(\bigvee\Theta)[\vec{a}]\| = \|\vec{a}\| \cdot \bigvee\{\|\theta[\vec{a}]\| : \theta \in \Theta\}$

(5) $\|(\forall x\phi)[\vec{a}]\| = \|\vec{a}\| \cdot \bigwedge\{\|\vec{a}\| \to \|\phi[a/x, \vec{a}]\| : a \in |M|_x\}$

(6) $\|(\exists x\phi)[\vec{a}]\| = \|\vec{a}\| \cdot \bigvee\{\|\vec{a}\| \cdot \|\phi[a/x, \vec{a}]\| : a \in |M|_x\}$.

Remarks *The interpretation of the formulas of the form $\forall\vec{z}(\phi \to \psi)$ is obtained by reading them as if they were built up using the primitives $\forall$, $\to$. The factor $\|\vec{a}\|$ on the right in line* (3) *can be omitted whenever Θ is non empty without changing the value (since then $\|\vec{a}\| \geq \|\theta[\vec{a}]\|$ for some $\theta \in \Theta$). The factor $\|\vec{a}\|$ in lines* (4) *and* (6) *can always be omitted without changing values. If $\vec{a}$ is the empty sequence (i.e., there are no free variables to be interpreted), then $\|\vec{a}\|$ is understood (as an "empty meet") to be $1 = 1_{\mathcal{H}}$. In a natural way, we may use abbreviated notation, e.g. as follows. For*

$$\|(x \approx y)[a/x, b/y]\|$$

we write $\|a \approx b\|$. So we have

$$\|a \approx b\| = \|a\| \cdot \|b\| \cdot (\approx^M (a, b)).$$

Also, $\|P(a, b, c)\| = \|P(x, y, z)[a/x, b/y, c/z]\| = \|a\| \cdot \|b\| \cdot \|c\| \cdot P^M(a, b, c)$, etc.

Turning to the interpretation of Gentzen sequents, we say M *satisfies* $\Phi \Rightarrow \Psi$, $M \models \Phi \Rightarrow \Psi$ if

$$\|(\bigwedge\Phi)[\vec{a}]\| \leq \|\bigvee\Psi[\vec{a}]\|$$

for any $\vec{a} = \langle a_1, \ldots, a_n\rangle$ in $|M|$. Here a_i interprets x_i, and $x_1, \ldots, x_n$ are exactly the distinct free variables occurring in $\Phi \cup \Psi$. This definition coincides with saying that $M \models \Phi \Rightarrow \Psi$ if $\|\forall\vec{x}(\bigwedge\Phi \to \bigvee\Psi)\|_M = 1$.

Finally, we can return to point (iv) above concerning the requirements on the $\mathcal{H}$-valued structure M. These are that M should satisfy the following *axioms of equality*:

$$\begin{array}{rcl} & \Rightarrow & x \approx x \\ x \approx y & \Rightarrow & y \approx x \\ x \approx y, \theta(x) & \Rightarrow & \theta(y) \end{array}$$

where x, y are variables of the same sort, $\theta(x)$ is any atomic formula (and $\theta(y)$ is obtained by substituting y for x in $\theta(x)$). Spelling out some consequences of these requirements, we obtain that

$$\begin{array}{rcl} \|a \approx a\| & = & \|a\| \\ \|a \approx b\| & = & \|b \approx a\| \\ \|a \approx b\| \cdot \|b \approx c\| & \leq & \|a \approx c\| \\ \|a \approx b\| & \leq & \|a\| \cdot \|b\| \end{array}$$

and also

$$\begin{array}{rcl} \|a \approx b\| & \leq & \|f^M(a) \approx f^M(b)\| \\ \|\vec{a}\| & \leq & \|f^M(\vec{a})\|. \end{array}$$

We note that every $\mathcal{H}$-valued structure will satisfy "the axiom for equality with respect to an arbitrary formula", in other words

$$\|a_1 \approx a_1'\| \cdots \|a_n \approx a_n'\| \cdot \|\phi[\vec{a} \text{ for } \vec{x}]\| \leq \phi[\vec{a} \text{ for } \vec{x}]\|.$$

Finally, let us point out that the ordinary (standard) interpretation of $L_{\infty\omega}$ is essentially a special case of the above, namely, when $\mathcal{H}$ is the two-element Boolean algebra 2. Given M, an $\mathcal{H}$-valued structure with $\mathcal{H} = 2$ according to the codification in this section, define M^* by $|M^*|_s = \{a \in |M|_s : \|a\|_M = 1\}$, $f^{M^*} = f^M$ restricted to the $|M|_s$, etc. Since $\|\vec{a}\|_M \leq \|f^M(\vec{a})\|_M$, the operations f^{M^*} are well-defined on the sets $|M^*|_s$. M^* will still differ from an ordinary structure since we have that $\approx^{M^*}$ is not necessarily the true identity. But the relations $\approx^{M,s}$ will be sufficiently well-behaved so that we can perform the familiar construction of the quotient, $M' =$ "$M^*/\approx^{M^*}$". M' is the ordinary structure that M can be identified with for all practical purposes.

§2 Sheaves over Heyting algebras

Any partial ordering $\mathcal{P}$ can be considered a site (c.f. Chapter 1) in the following well-known way. First of all, $\mathcal{P}$ is regarded as the category whose objects are the elements of $\mathcal{P}$ and whose morphisms are: exactly one arrow $a \to b$ for those a, b such that $a \leq b$, no arrow between a, b otherwise (the definitions of composition and identity morphisms are the uniquely given). Secondly, using the order theoretic supremum, we declare that the family $(a_i)_{i\in I}$ (where $a_i \leq a$ for all $i \in I$) *covers* a iff $\bigvee_{i\in I} a_i = a$ where $\bigvee_{i\in I} a_i$ stands for the sup of the a_i. ($\bigvee_{i\in I} a_i = a$ hence means that a is the smallest element such that $a_i \leq a$ for all $i \in I$.) If, in particular, $\mathcal{P}$ is a Heyting algebra, then $\mathcal{P}$ as a category will have left limits. The topology on $\mathcal{P}$ just described is identical to the canonical topology in the sense of Chapter 1, Section 1.

$\mathrm{Sh}_{\mathcal{P}}$ denotes the category of sheaves over $\mathcal{P}$ as a site. For $\mathcal{P} = \mathcal{H}$, a complete Heyting algebra, D. Higgs has given an alternative description of $\mathrm{Sh}_{\mathcal{H}}$. We are going to define the *category of $\mathcal{H}$-valued sets*, denoted $\mathbf{Set}_{\mathcal{H}}$. An object of $\mathbf{Set}_{\mathcal{H}}$, i.e., and $\mathcal{H}$-valued set, is a pair (X, δ) with X an arbitrary set and

$$\delta : X \times X \to \mathcal{H}$$

satisfying the following conditions:

$$\begin{aligned} \delta(x, x') &= \delta(x', x) \\ \delta(x, x') \cdot \delta(x', x'') &\leq \delta(x, x''). \end{aligned}$$

(Intuitively, $\delta(x, x)$ is the degree of existence of x in (X, δ), and is denoted sometimes $\|x\|$. Also, $\delta(x, y)$ is the degree of equality of x and y, and it is denoted $\|x = y\|$ as well. Notice that $\|x = y\| \leq \|x\| \cdot \|y\|$, in particular, $\|x = x\| = 1$ may fail.) A morphism in $\mathbf{Set}_{\mathcal{H}}$,

$$f : (X, \delta) \to (Y, \varepsilon)$$

is a map

$$f : X \times Y \to \mathcal{H}$$

such that the following are true:

$$\begin{aligned} \varepsilon(y, y') \cdot f(x, y) &\leq f(x, y') \\ \delta(x, x') \cdot f(x, y) &\leq f(x', y) \\ f(x, y) \cdot f(x, y') &\leq \varepsilon(y, y') \\ \delta(x, x) &= \textstyle\bigvee_{y\in Y} f(x, y). \end{aligned}$$

This last condition implies that

$$f(x, y) \leq \delta(x, x)$$

and the third one that

$$f(x, y) \leq \varepsilon(y, y).$$

(Intuitively, we have in mind an "$\mathcal{H}$-valued map" $F : X \to Y$ for which it makes sense to ask for the value of $y = F(x)$, for $x \in X$, $y \in Y$, denoted $\|y = F(x)\|$. What we actually define is the function $\|\cdot = F(\cdot)\|$, i.e., the intention is that $f(x, y) = \|y = F(x)\|$. Under this interpretation, the above conditions express the "$\mathcal{H}$-valued functionality" of F. E.g., the last condition says that "$F(x)$ is always defined" with value $\|x\|$.)

The composition of two morphisms $(X, \delta) \xrightarrow{f} (Y, \varepsilon)$ and $(Y, \varepsilon) \xrightarrow{g} (Z, \eta)$ is the function

$$h : X \times Z \to \mathcal{H}$$

defined by

$$h(x, z) = \bigvee_{y \in Y} f(x, y) \cdot g(y, z).$$

It is easy to check that h is actually a morphism and the law of associativity holds. The identity morphism

$$i : (X, \delta) \to (X, \delta)$$

is defined by $i(x, y) = \delta(x, y)$.

Theorem 4.2.1 (D. Higgs [1973]). *The categories* $\mathrm{Sh}_{\mathcal{H}}$ *and* $\mathbf{Set}_{\mathcal{H}}$ *are equivalent.*

It is the category $\mathbf{Set}_{\mathcal{H}}$ that we can relate ordinary logical construction to more easily. The rest of this section is devoted to explaining in what precise sense an $\mathcal{H}$-valued structure for the language L, and a (categorical) interpretation $L \to \mathbf{Set}_{\mathcal{H}}$ of L is $\mathbf{Set}_{\mathcal{H}}$ are essentially the same thing.

Given M, an $\mathcal{H}$-valued structure for L as defined in Section 1, this is how we define $M : L \to \mathbf{Set}_{\mathcal{H}}$ (there will not be any confusion by our use of M to denote this second item too).

(i) For a sort s in L, we put $M(s)$ to be the $\mathcal{H}$-valued set $(X, \delta) = (|M|_s, \|\cdot \approx \cdot\|)$, i.e.

$$\delta(a, b) = \|a \approx b\|.$$

Since M satisfies the equality axioms, $M(s)$ is indeed and $\mathcal{H}$-valued set.

(ii) Notice that the product in $\mathbf{Set}_{\mathcal{H}}$

$$(X_1, \delta_1) \times \cdots \times (X_n, \delta_n)$$

is (X, δ) where $X = X_1 \times \cdots \times X_n$ and $\delta(\vec{x}, \vec{x}') = \delta_1(x_1, x_1') \cdots \delta_n(x_n, x_n')$, with the projections

$$\pi_i : (X_i, \delta_i) \to (X, \delta)$$

i.e.,

$$\pi_i : X_i \times X \to \mathcal{H},$$

defined by $\pi_i(x_i', \langle x_1, \ldots, x_n \rangle) = \delta_i(x_i', x_i)$.

Now, for an operation symbol $f : s_1 \times \cdots \times s_n \to s$ we define $M(f) : M(s_1) \times \cdots \times M(s_n) \to M(s)$ by

$$M(f)(\langle x_1, \ldots, x_n \rangle, x) = \|x_1\| \cdots \|x_n\| \cdot \|x \approx f^M(x_1, \ldots, x_n)\|.$$

The latter value is the value of the formula $v \approx f(u_1, \ldots, u_n)$ with u_i interpreted by x_i, v by x. It is easy to check that $M(f)$ thus defined is indeed a morphism

$$M(f) : M(s_1) \times \cdots \times M(s_n) \to M(s)$$

in $\mathbf{Set}_{\mathcal{H}}$.

(iii) Let P be an n-place relation symbol $P \hookrightarrow s_1 \times \cdots \times s_n$. We define the subobject $M(P) \xhookrightarrow{i} M(s_1) \times \cdots \times M(s_n) \underset{\mathrm{df}}{=} X$ as follows. As an object, we put $M(P) = (X, \delta)$ where $\delta(\vec{x}, \vec{x}') = \|x_1 \approx x_1'\| \cdots \|x_n \approx x_n'\| \cdot P^M(x_1, \ldots, x_n)$ where $\vec{x} = \langle x_1, \ldots, x_n \rangle$, $\vec{x}' = \langle x_1', \ldots, x_n' \rangle$. For the morphism $i : M(P) \to X$ we put $i(\vec{x}, \vec{x}') = \delta(\vec{x}, \vec{x}')$ ("inclusion map"). It is easy to check that $i : M(P) \to X$ is a monomorphism in $\mathbf{Set}_{\mathcal{H}}$.

Having legitimately defined M as an interpretation $M : L \to \mathbf{Set}_{\mathcal{H}}$, we can evaluate formulas of $L_{\infty\omega}$ in M, i.e., can consider $M_{\vec{x}}(\phi) \hookrightarrow M(\vec{x})$. We now establish a link between this interpretation and the $\mathcal{H}$-valued interpretation introduced in the preceding section.

Given any formula ϕ of $L_{\infty\omega}$, define the object $\widehat{M}_{\vec{x}}(\phi) = (X, \delta)$ in $\mathbf{Set}_{\mathcal{H}}$ as follows:

$$X = |M|_{x_1} \times \cdots \times |M|_{x_n},$$

$$\delta(\vec{a}, \vec{a}') = \|a_1 \approx a_1'\| \cdots \|a_n \approx a_n'\| \cdot \|\phi[\vec{a} \text{ for } \vec{x}]\|_M$$

It can be easily seen that (X, δ) is an $\mathcal{H}$-valued set. Define $i : \widehat{M}_{\vec{x}}(\phi) \to M(\vec{x})$ by $i(\vec{a}, \vec{a}') = \delta(\vec{a}, \vec{a}')$ to obtain the subobject

$$\widehat{M}_{\vec{x}}(\phi) \xhookrightarrow{i} M(\vec{x})$$

Proposition 4.2.2 (i) *With the above notations, $\widehat{M}_{\vec{x}}(\phi) \hookrightarrow M(\vec{x})$ is isomorphic, as a subobject, to $M_{\vec{x}}(\phi)$.*

(ii) *$M \models \sigma$ for a sequent σ of $L_{\infty\omega}$ in the categorical sense iff $M \models \sigma$ in the sense of Section 1.*

The proof of (i) is a straightforward induction on the complexity of ϕ, and (ii) is an easy consequence of (i) and the definitions involved.

4.2.2 is a generalization of the basic fact stated in Chapter 2, Section 3 that relates Tarski type truth definition for ordinary models and the categorical interpretation in **Set**. The reader will find that there is a word-for-word translation of any detailed verification of the latter fact into one of 4.2.2.

Finally, let us mention that conversely, every categorical interpretation $M : L \to \mathbf{Set}_{\mathcal{H}}$ can be derived from an $\mathcal{H}$-valued structure M as above, but for this a slightly more general definition of $\mathcal{H}$-valued structures is needed, namely, the definition of operations f^M should be made appropriately "$\mathcal{H}$-valued".

§3 Boolean homomorphisms

In certain important cases, Boolean-valued models give rise to two-valued (or ordinary) ones, by way of a two-valued homomorphism on the value-algebra. Placing our discussion into a slightly more general context, let $\mathcal{B}$, $\mathcal{B}_0$ be two complete Boolean algebras, and let $h : \mathcal{B} \to \mathcal{B}_0$ be a homomorphism (preserving the finitary Boolean operations $\neg$, $\wedge$, $\vee$). Since in infinitary logic, we have to deal with infinitary $\bigwedge$, $\bigvee$, we will have to consider approximately stronger notions of homomorphism. We say that h *preserves* the sup $\bigvee \mathcal{X}$, where $\mathcal{X}$ is a subset of $\mathcal{B}$, or the inf $\bigwedge \mathcal{X}$, if

$$\begin{aligned} h(\textstyle\bigvee^{(\mathcal{B})} \mathcal{X}) &= \textstyle\bigvee^{(\mathcal{B}_0)} \{h(\beta) : \beta \in \mathcal{X}\} \\ h(\textstyle\bigwedge^{(\mathcal{B})} \mathcal{X}) &= \textstyle\bigwedge^{(\mathcal{B}_0)} \{h(\beta) : \beta \in \mathcal{X}\} \end{aligned}$$

respectively.

Let M be a $\mathcal{B}$-valued structure and let F be a fragment of $L_{\infty\omega}$; assume F is a *set* (as oppose to a *proper class*). The disjunctions, conjunctions and quantified formulas in F coupled with elements of $|M|$ induce sups and infs of certain subsets of $\mathcal{B}$ that we call the logical sups and infs induced by M and F. E.g., a formula $\bigwedge\Sigma(\vec{x})$ in F and appropriate elements $\vec{a}$ in $|M|$ induce the "logical" inf $\bigwedge^{(\mathcal{B})}\{\|\phi[a,\vec{a}]\| : a \in |M|_x\}$. Now, assume that $h : \mathcal{B} \to \mathcal{B}_0$ is a homomorphism preserving all the logical sups and infs in $\mathcal{B}$ induced by M and F. Then, we claim, there is an essentially unique $\mathcal{B}_0$-valued structure denoted M/h such that $|M/h|_s = |M!_s$ for any sort s of L and we have

$$\|\phi[\vec{a}]\|_{M/h} = h\|\phi[\vec{a}]\|_M \tag{1}$$

for any formula ϕ in F. Our claim is practically obvious. First of all, the required equality suggests the definition of M/h as follows

$$\begin{array}{rcl} \|a\|_{M/h} & = & h(\|a\|_M) \\ f^{M/h}(\vec{a}) & = & f^M(\vec{a}) \\ P^{M/h}(\vec{a}) & = & h(P^M\vec{a}) \\ (or, & = & h\|P(\vec{a})\|_M, \text{which does not make any difference for (1))} \\ \approx^{M/h}(a,b) & = & h(\approx^M(a,b)). \end{array}$$

Then a straightforward induction on ϕ in F shows that (1) is true for all ϕ in F; we will use the homomorphism property as well as the preservation properties of h.

What is left is to see that homomorphisms with the appropriate additional preservation properties exist in sufficient numbers.

The following is a classical result. C.f. e.g. Rasiowa and Sikorski [1963].

Theorem 4.3.1 (Rasiowa-Sikorski Lemma). *Let $\mathcal{B}$ be a Boolean algebra. Given an element $b \neq 1$ in $\mathcal{B}$, and given countably many sups and infs: $\bigvee\mathcal{X}_i$, $\bigwedge\mathcal{Y}_i$* existing *in $\mathcal{B}$, There is a two-valued homomorphism $h : \mathcal{B} \to 2$ preserving the given sups and infs and such that $h(b) = 0$.*

Corollary 4.3.2 *Let L be a countable language, F a* countable *fragment of $L_{\infty\omega}$. Let Σ be a set of sequents of F, σ a sequent of F. Assume that there is a (non-trivial) complete Boolean algebra $\mathcal{B}$ and a countable $\mathcal{B}$-valued model M of Σ (i.e., $M \models \sigma'$ for $\sigma' \in \Sigma$) such that $M \not\models \sigma$. Then there is an ordinary model M' of Σ such that $M \not\models \sigma$.*

Remark *"M is countable" means that each set $|M|_s$ is countable. Actually, the countability assumption on M is not necessary (by the $\mathcal{B}$-valued downward Löwenheim-Skolem theorem), but the present version is sufficient for our purposes.*

Proof of 4.3.2 By assumption $b \underset{\text{df}}{=} \|(\forall)\sigma\| \underset{\text{df}}{=} \|\forall\vec{x}(\bigwedge\Phi \Rightarrow \bigvee\Psi)\|$ (where $\sigma := \Phi \Rightarrow \Psi$, $\vec{x}$ is the sequence of free variables in σ) is $\neq 1$. Let, by the Rasiowa-Sirkoski lemma, $h : \mathcal{B} \to 2$ be a homomorphism preserving all logical sups and infs induced by M and F, and such that $h(b) = 0$ and let $M' = M/h$. By the identity (1) $M \models \Sigma$ implies $M' \models \Sigma$. Also, since $h(b) = 0$, (1) implies that $M' \not\models \sigma$. M' is essentially identical to an ordinary model, c.f. our discussion of this point in Section 1. □

Another situation when we can guarantee the existence of two-valued models is when the fragment F is *finitary*, i.e., when for every conjunction $\bigwedge\Sigma \in F$, or disjunction $\bigvee\Sigma \in F$, Σ is a finite set. We briefly describe the situation as follows.

Definition 4.3.3 *A $\mathcal{B}$-valued structure M is called* full *if for any formula $\exists x\phi(x,\vec{x})$ in $L_{\infty\omega}$, and any $\vec{a}$ in $|M|$, $b = \|\exists x\phi(x,\vec{x})[\vec{a}/\vec{x}]\| = \|\phi(x,\vec{x})[a/x,\vec{a}/\vec{x}]\|$ for some $a \in |M|_x$ i.e., the sup defining b is actually a* maximum.

Proposition 4.3.4 *For any $\mathcal{B}$-valued structure M ($\mathcal{B}$ is a complete Boolean algebra), there is a full $\mathcal{B}$-valued structure M' such that M' is $L_{\infty\omega}$-equivalent to M, i.e.,*

$$M \models \sigma \;\Leftrightarrow\; M' \models \sigma$$

for any sequent σ of $L_{\infty\omega}$.

The proof (and further information on Boolean-valued models) can be found in Rosser [1969]. Incidentally, the construction of M' is closely related to the construction of the sheaf over $\mathcal{B}$ that corresponds to a given $\mathcal{B}$-valued set, in the sense of Higgs' theorem 4.2.1.

Proposition 4.3.5 *Suppose F is a finitary fragment, Σ is a set of sequents in F, σ is a sequent of F, and M is a $\mathcal{B}$-valued model of Σ for $\mathcal{B}$ a non-trivial Boolean algebra such that $M \not\models \sigma$. The there is an ordinary two-valued model $M' \models \Sigma$ such that $M' \not\models \sigma$.*

Proof. First of all, by 4.3.4, we can assume that M is full. Now, by the Stone ultrafilter existence theorem, there is a homomorphism $h : \mathcal{B} \to 2$ such that $h(b) = 0$ where $b \underset{\text{df}}{=} \|(\forall)\sigma\| \neq 1$. Putting $M' = M/h$, we can prove the identity (1) above for finitary formulas ϕ in F. All conjunction and disjunctions involved are finitary, so h will automatically preserve them. The sups (and infs) corresponding to quantifiers will also be automatically preserved because M is full. It follows that M' satisfies the assertions. □

There is one more matter we have to deal with in this section, viz. the construction of the complete Boolean algebra of *regular open subsets* of a partially ordered (p.o.) set $\mathcal{P}$. Let $\mathcal{P} = (P, \leq)$ be a p.o. set. A subset U of P is called *open* if $p \leq q$, $q \in U$ imply $p \in U$. The *regularization* of a set $U \subset P$, U^*, or $\neg\neg U$, is defined as

$$U^* = \{p \in P : \forall q \leq p\; \exists r \leq q\; r \in U\}.$$

If U is open, so is U^*. An open set U is called *regular* if $U^* = U$; U^* is always regular. It is well-known (and it is easy to show) that the regular open subsets of $\mathcal{P}$ endowed with inclusion as a partial orienting will form a complete Boolean algebra, denoted by $\mathcal{P}^*$. For future reference, we note the following easily proved additional facts:

(i) the 0 and 1 of $\mathcal{P}^*$ are $\varnothing$ and P, respectively,

(ii) if $U \subseteq V$ then $U^* \subseteq V^*$, $\bigvee_{i\in I} U_i^* = (\bigcup_{i \in I} U_i)^*$ where "$\bigvee_{i\in I}$" denotes the Boolean sup in the sense of $\mathcal{P}^*$.

(iii) $U_1^* \wedge U_2^* = (U_1 \cap U_2)^* = U_1^* \cap U_2^*$.

Here U, V, etc., denote arbitrary open sets.

Chapter 5

Completeness

Introduction

In this chapter we will give two completeness theorems, one for the full logic $L_{\infty\omega}$, and the other for the coherent sub logic $L^g_{\infty\omega}$ of $L_{\infty\omega}$. Here we briefly discuss some of the features of the proof systems and the completeness proofs.

The first system (call it $\mathcal{G}$) refers to arbitrary fragments of the full language $L_{\infty\omega}$. The proof system $\mathcal{G}$ we formulate is a Gentzen-type system for deriving Gentzen sequents. The completeness of $\mathcal{G}$ is proved with respect to interpretations in models with truth values in non-trivial complete Boolean algebras. Our proof is an inessential modification of Mansfield's proof in Mansfield [1972] for a related system. However, the following features of the proof system as compared to traditional formulations should be emphasized:

(i) $\mathcal{G}$ refers to many-sorted logic, and in this respect is related to Feferman's formulation in Feferman [1968].

(ii) $\mathcal{G}$ is heard to be sound with respect to interpretations when one or more of the partial universes are allowed to be empty. This feature is essential for our purposes. The effect of "possibly empty domains" can be summarized in the following simple way. With the notable exception of $\exists \Rightarrow$ and $\Rightarrow \forall$, all rules have to have the feature that each free variable occurring in at leas one premise should actually occur free in the conclusion. It turns out that this takes care of excluding the introduction of the "existential assumption" of the non-emptiness of a domain. Some rules in the ordinary Gentzen system already have this feature, some others have to be restricted by requiring the above as a proviso.

(iii) The system $\mathcal{G}$ is for deriving sequence from (non-logical) assumptions (axioms) hence some form of the cut-rule is necessary. We have formulated a cut-rule in which the cut formulas are allowed to be exactly the substitution instances of the formulas ϕ_i, ψ_j, in any of the assumption sequents $\phi_1, \ldots, \phi_n \Rightarrow \psi_1, \ldots, \psi_m$. Hence, with the empty set of non-logical axioms, our system reduces to the basic Gentzen type cut-free system. Although the resulting generalization of Gentzen's Hauptsatz can be derived from the original version with not too much (but non-zero amount of) effort, the generalization seems to be interesting even for finitary logic. We have learned that Barwise [1967] contains this type of restricted cut rule.

We will give full proofs mainly to emphasize the naturalness of $\mathcal{G}$ from the point of view of the completeness proof. We will indicate that $\mathcal{G}$ can be recovered from a certain plan of the completeness proof. This feature is one aspect of the *naturality* of Gentzen type systems – and is well-known for logicians. We include this discussion for the benefit of the non-logicians.

Our second system is one that can work only with the sub logic $L^g_{\infty\omega}$. It is a "one sided" system in which the right-hand-sides of the sequents do not 'change' at all during the derivations. We have found good implicit use of this system in the proofs in Chapter 7. This system seems to be new. It is interesting to note that this system too is the result of a certain outline of a completeness proof.

§1 A Boolean-complete formalization of $L_{\infty\omega}$

We remind the reader of some of the features of the logic $L_{\infty\omega}$ as considered in this book; for more detailed information, c.f. also Chapter 2. It is based on a language L appropriate for many sorted logic.

There is a distinction, in the usual manner, between free and bound variables although in notation this is mostly neglected. E.g. for a formula $\phi(x)$ with a free variable x, $\exists x\phi(x)$ actually denotes the formula $\exists w\phi(w)$ where w is a bound variable not occurring in $\phi(x)$ and $\phi(w)$ is obtained by substituting w for x. There is a similar convention applied to the use of the universal quantifier. It has the logical operators

$\bigwedge$	(conjunction; applicable to any *set* of formulas with altogether finitely many free variables)
$\bigvee$	(disjunction; applicable similarly as $\bigwedge$)
$\exists$	(existential qualifier; applied to a single variable at a time)
$\forall - (\cdot \to \cdot\cdot)$	(compound universal quantifier).

If $\phi(x_1, \ldots, x_n)$, $\psi(x_1, \ldots, x_n)$ are formulas with the indicated free variables and $\vec{z}$ is a sequent of distinct variables, then $\forall\vec{z}(\phi \to \psi)$ is a formula. We consider the formulas $\forall x\phi$, $\phi \to \psi$ and $\neg\phi$ as special cases obtained as $\forall x(\uparrow\to \phi)$, $\forall\varnothing(\phi \to \psi)$ and $\forall\varnothing(\phi \to\downarrow)$.

A *fragment* F of $L_{\infty\omega}$ is any class of formulas that is closed under (i) subformulas, (ii) substitution of terms of L for free variables. Note that $\forall\vec{z}(\phi \to \psi)$ has ϕ and ψ as subformulas but not $\phi \to \psi$.

A *sequent* of F is a formal expression of the form $\Phi \to \Psi$ with finite sets Φ, Ψ of formulas of F. A *theory* in F is an arbitrary set of sequents of F.

For an *interpretation* M, in the ordinary (set-valued) sense, we allow M to have empty partial domains. For a theory T and a sequent σ, both in a given fragment F, we write $T \models \sigma$ to mean that σ is true in any ordinary (set-valued) model of T. We will sometimes write $T \overset{\mathrm{b}}{\models} \sigma$ to mean that σ is true in every Boolean-valued model of T. (For a discussion of Boolean-valued models, see Chapter 4.)

It is very instructive to consider how one can arrive at a proof system, by starting out with the outline of a proposed completeness proof. We will indicate this process under the assumption that the fragment F is a *countable set*; in this case we can aim at set-valued models. Accordingly, assume that we have some notion

$$T \vdash \sigma \quad (\text{“}\sigma \text{ is derivable from } T\text{”})$$

with a given fixed theory T in F, and with a variable sequent σ of F; we now show requirements arising from a proposed proof of the fact that

$$T \models \sigma \;\Rightarrow\; T \vdash \sigma.$$

Assume $T \not\vdash \sigma$, i.e., σ is not derivable. We seek to construct a model M of T such that $M \not\models \sigma$. We wish to construct M out of a pair $\langle H_1, H_2 \rangle$ of sets of formulas in F in a way such that we will have:

(i) the domain of M consists of the terms whose (free) variables occur free in at least on formula in $H_1 \cup H_2$, and

(ii) for any formula $\phi(\vec{x})$ in F, with the indicated free variables which are all in $|M|$, the domain of M, we have that

(ii)$_\mathrm{a}$ $\phi(\vec{x}) \in H_1$ if and only if $\phi(\vec{x})$ is true in M (with $\vec{x}$ interpreting themselves in $|M|$) and

(ii)$_\mathrm{b}$ $\phi(\vec{x}) \in H_2$ if and only if $\phi(\vec{x})$ is false in M.

Intuitively, what will happen is that by a gradual building of the sets H_1, H_2 we approximate the total description of which formulas of F are true in M.

Next we observe that these conditions put some requirements on H_1 and H_2, some examples of which are:

(iii) $H_1 \cap H_2 = \varnothing$.

(iv) For every $\phi \in F$ with free variables in $|M|$, either $\phi \in H_1$, or $\phi \in H_2$.

(v) If $\bigwedge\Theta \in H_1$, then every $\phi \in \Theta$ belongs to H_1.

(vi) If $\bigwedge\Theta \in H_2$, then some $\phi \in \Theta$ belongs to H_2.

Spelling out all these conditions, we obtain what we call a *Hintikka pair* $\langle H_1, H_2 \rangle$. (Since our proof below will be "Boolean valued" and hence formally quite different, there is no point in spelling out the full definition of a Hintikka pair.) Then we verify that in fact, every Hintikka pair defines a model satisfying (ii). We notice that in the notation of Hintikka pair one of H_1, H_2 is redundant since e.g., $H_2 = F - H_1$. However, from the point of view of the proof, the given one is the natural formulation of the notion.

The next step is to translate the required properties of M into properties of $\langle H_1, H_2 \rangle$. E.g., we force M *not* to satisfy $\sigma = \Phi_0 \Rightarrow \Psi_0$ by requiring $\Phi_0 \subset H_1$ and $\Psi_0 \subset H_2$. (Notice that these relationships indeed mean that σ fails to hold in M for the *particular* elements denoted by the free variables in Φ_0, Ψ_0.) To say that an arbitrary sequent $\Phi(x_1, \ldots, x_n) \Rightarrow \Psi(x_1, \ldots, x_n)$ holds in M is equivalent to saying that it is *not the case* that $\Phi(t_1, \ldots, t_n) \subset H_1$ and $\Psi(t_1, \ldots, t_n) \subset H_2$, where

$$\Phi(t_1, \ldots, t_n) = \{\phi(t_1/x_1, \ldots, t_n/x_n) : \phi \in \Phi\}$$
$$\Psi(t_1, \ldots, t_n) = \{\psi(t_1/x_1, \ldots, t_n/x_n) : \psi \in \Psi\}$$

for any set $t_1, \ldots, t_n$ of terms having the respective sorts of $x_1, \ldots, x_n$.

The next task is to construct $\langle H_1, H_2 \rangle$. We seek to construct H_i $(i = 1, 2)$ as an increasing union $\bigcup_{n<\omega} H_n^i$ of finite sets H_n^i. Naturally, we set $H_0^1 = \Phi_0$ and $H_0^2 = \Psi_0$. It is very natural now to impose the following *induction hypothesis* on $\langle H_n^1, H_n^2 \rangle$: $T \not\vdash H_n^1 \Rightarrow H_n^2$, referring explicitly to $\vdash$ for the first time. Next we see that all required closure properties of $\vdash$ (i.e. the rules of inference) can be recovered from the assumption that the construction of the H_n^i can indeed be performed. To start with, we will ensure (iii), i.e. the condition $H_1 \cap H_2 = \varnothing$ by making sure that $H_n^1 \cap H_n^2 = \varnothing$ for all n. This in turn will be made to hold as a consequence of the main property $T \not\vdash H_n^1 \Rightarrow H_n^2$. To this end, we define '$\vdash$' is such a way that whenever $H_n^1 \cap H_n^2 \neq \varnothing$, then $T \vdash H_n^1 \mod H_n^2$. This effect will be achieved by postulating each sequent of the form $\Phi, \phi \Rightarrow \Psi, \phi$ (with common member ϕ of the sets on the left and the right) an *axiom*, and we will have $T \vdash \Phi, \phi \Rightarrow \Psi, \phi$.

To make e.g., (iv) true we should be able to do the following. Suppose H_n^1, H_n^2 have been constructed such that

$$T \nvdash H_n^1 \mod H_n^2.$$

We take some ϕ whose free variables all occur in $H_n^1 \Rightarrow H_n^2$. We want to put *either* $H_{n+1}^1 = H_n^1 \cup \{\phi\}$ and $H_{n+1}^2 = H_n^2$ *or* $H_{n+1}^1 = H_n^1$ and $H_{n+1}^2 = H_n^2 \cup \{\phi\}$, to contribute to making (iv) true.

Hence we want that either

$$T \nvdash H_n^1, \phi \Rightarrow H_n^2 \quad \text{or} \quad T \nvdash H_n^1 \Rightarrow H_n^2, \phi.$$

All this is equivalent to saying that the rule

$$\frac{\Phi, \phi \Rightarrow \Psi \qquad \Phi \Rightarrow \Psi, \phi}{\Phi \Rightarrow \Psi}$$

(with $\Phi = H_n^1$, $\Psi = H_n^2$ and with ϕ satisfying the conditions regarding the free variables) should at least be a *derived rule* for our proof-system, i.e., if the two sequents above the line are derivable from T, then so is the one under the line. The exhibited rule is the well-known *cut-rule* (with a restriction on variables necessitated by the appearance of possible empty domains).

As a final example, let us try to make (v) true. Assume that H_n^1, H_n^2 have been constructed so that $T \nvdash H_n^1 \Rightarrow H_n^2$, and let $\bigwedge\Theta \in H_n^1$. We pick an arbitrary $\theta \in \Theta$ and we want to put $H_{n+1}^1 = H_n^1 \cup \{\theta\}$ and $H_{n+1}^2 = H_n^2$. Hence we want that the fact

$$T \nvdash H_n^1, \theta \Rightarrow H_n^2.$$

is a consequence of $T \nvdash H_n^1 \Rightarrow H_n^2$. In other words, taking the contra-positive, we want that the rule

$$\frac{\Phi, \bigwedge\Theta, \theta \Rightarrow \Psi}{\Phi, \bigwedge\Theta \Rightarrow \Psi} \quad \text{subject to } \theta \in \Theta$$

should be at least a *derived rule*.

By a similar argument, we can show that condition (vi) above entails the rule

$$\frac{\{\Phi \Rightarrow \Psi, \bigwedge\Theta, \theta : \theta \in \Theta\}}{\Phi \Rightarrow \Psi, \bigwedge\Theta}$$

with a possibly infinite set of premises.

The actual system we give below will be a bit weaker that the one recovered by the above procedure. Essentially, what we will make sure is only that the left-to-right implications of the above conditions $(\mathrm{ii})_{\mathrm{a}}$, $(\mathrm{ii})_{\mathrm{b}}$ will hold. Then a special measure is needed to make sure that M will be a model of T; this is our *restricted* cut rule.

We warn the reader that our proof below will bear no formal resemblance to the argument given above; this is because we will give a Boolean completeness proof. The set-valued on will then be obtained by an application of the Rasiowa-Sikorski theorem. Besides being more general, the Boolean construction has the advantage that it is a *canonical* construction.

This should suffice to indicate how to find our rules of inference and we may turn to the definition of our formal system. We take a fixed but arbitrary set of sequents T. Some of the axioms and rules will depend on T. We denote the formal system associated to T by $\mathcal{G}_T^1$, (and the "one-sided" system introduced later by $\mathcal{G}_T^2$). Below, capital Greek letters denote finite sets of formulas, lower case Greek letters denote formulas in F. $T^=$ will denote the set $T \cup \{\text{all axioms of equality}\}$, c.f. Chapter 4, Section 1.

Axioms

$$(A1) \qquad \Phi, \phi \Rightarrow \Psi, \phi$$
$$(\text{abbreviating } \ \Phi \cup \{\phi\} \Rightarrow \Psi \cup \{\phi\})$$

for any *atomic* formula ϕ.

$$(A2)_{T^=} \qquad \Phi, \Theta(t_1, \ldots, t_n) \Rightarrow \Psi, \Gamma(t_1, \ldots, t_n)$$

provided $\Theta(x_1, \ldots, x_n) \Rightarrow \Gamma(x_1, \ldots, x_n)$ belongs to $T^=$, the arbitrary terms $t_1, \ldots, t_n$ are substituted for the free variables $x_1, \ldots, x_n$ with due regard for syntactic correctness.

Rules of Inference

$$(\bigwedge \Rightarrow) \qquad \frac{\Phi, \bigwedge\Theta, \theta \Rightarrow \Psi}{\Phi, \bigwedge\Theta \Rightarrow \Psi} \qquad \text{if } \theta \in \Theta$$

(that is, the rule is admitted if $\theta \in \Theta$).

$$(\Rightarrow \bigwedge) \qquad \frac{\{\Phi \Rightarrow \Psi, \bigwedge\Theta, \theta : \theta \in \Theta\}}{\Phi \Rightarrow \Psi, \bigwedge\Theta}$$

(This is a rule with possibly infinitely many premises; it becomes an *axiom* (with zero premises) if $\Theta = \varnothing$: $\quad \Phi \Rightarrow \Psi, \bigwedge\varnothing$.)

$$(\bigvee \Rightarrow) \qquad \frac{\{\Phi, \bigvee\Theta, \theta \Rightarrow \Psi : \theta \in \Theta\}}{\Phi, \bigvee\Theta \Rightarrow \Psi}$$

$$(\Rightarrow \bigvee) \qquad \frac{\Phi \Rightarrow \Psi, \bigvee\Theta, \theta}{\Phi, \Rightarrow \Psi, \bigvee\Theta} \qquad \text{if } \theta \in \Theta$$

$$(\forall \Rightarrow) \qquad \frac{\Phi, \forall\vec{z}(\phi \to \psi), \psi(\vec{t}) \Rightarrow \Psi; \quad \Phi, \forall\vec{z}(\phi \to \psi) \Rightarrow \Psi, \phi(\vec{t})}{\Phi, \forall\vec{z}(\phi \to \psi) \Rightarrow \Psi}$$

provided all the free variables in the premises occur free in the conclusion; here $\phi(\vec{t})$ denotes the result of substituting the respective members of the sequence $\vec{t}$ of terms for $\vec{z}$ in ϕ. (It might happen that ϕ does not actually contain some of the variables in $\vec{z}$ and thus some free variables in $\vec{t}$ may fail to occur in $\phi(\vec{t})$. Then these variables should occur elsewhere in the conclusion. The proviso is necessary to ensure soundness with respect to possibly empty domains.)

$$(\Rightarrow \forall) \qquad \frac{\Phi, \phi(\vec{y}) \Rightarrow \Psi, \forall\vec{z}(\phi \to \psi), \psi(\vec{y})}{\Phi \Rightarrow \Psi, \forall\vec{z}(\phi \to \psi)}$$

provided the variables in $\vec{y}$ do not occur free in the conclusion; $\phi(\vec{y})$ results by substituting $\vec{y}$ for $\vec{z}$.

Remarks *on the last two rules.* 1) *It is instructive to check that these two rules correspond to the following two "construction steps", or closure conditions, regarding H_1 and H_2 (c.f. discussion above):*

(i) *If $\forall\vec{z}(\phi \to \psi) \in H_1$, then for any $\vec{t}$ whose all (free) variables occur in formulas in*

H_1, either $\psi(\vec{t}) \in H_1$ or $\phi(\vec{t}) \in H_2$.

(ii) *If* $\forall \vec{z}(\phi \to \psi) \in H_2$, *then for some 'new' variables* $\vec{y}$, $\phi(\vec{y}) \in H_1$ *and* $\psi(\vec{y}) \in H_2$.

2) *Recalling the identification of* $\neg\phi$, $\phi \to \psi$ *and* $\forall x\phi$ *with particular cases of the compound operation* $\forall \cdot (\cdot \to \cdot\cdot)$, *by specialization we obtain the familiar rules (c.f. e.g. Feferman [1968] for* $\neg$, $\to$ *and* $\forall$ *from the last two rules.*

$$(\exists \Rightarrow) \qquad \frac{\Phi, \exists x\theta(x), \theta(y) \Rightarrow \Psi}{\Phi, \exists x\theta(x) \Rightarrow \Psi}$$

provided that the variable y does not occur free in the conclusion.

$$(\Rightarrow \exists) \qquad \frac{\Phi \Rightarrow \Psi, \exists x\theta(x), \theta(t)}{\Phi \Rightarrow \Psi, \exists x\theta(x)}$$

provided that every free variable in the premise occurs in the conclusion.

$$(\mathrm{Cut}_{T^=}) \qquad \frac{\Phi, \phi \Rightarrow \Psi \qquad \Phi \Rightarrow \Psi, \phi}{\Phi \Rightarrow \Psi}$$

provided each free variable in ϕ is free in the conclusion and ϕ is a substitution instance $\phi'(\vec{t})$ of a formula $\phi'(\vec{x})$ such that $\phi' \in \Phi' \cup \Psi'$ for some $\Phi' \Rightarrow \Psi'$ belonging to the set $T^=$.

The proof system defines the notion "σ is a formal consequence of T" in notation

$$T \vdash \sigma.$$

To this end, it is not necessary to introduce any notion of formal proof. Rather, we define the set of formal consequences of T as the least set closed under the given rules of inference. More precisely, let $\vdash_T$ be the smallest set of sequents in F such that all the axioms $(A1) - (A2)$ belong to $\vdash_T$, and whenever in an instance of any one of the rules of inference the premise (or, all the premises) belong to $\vdash_T$, so does the conclusion. Notice that with any given T, in the definition of $\vdash_T$ we consider the axioms $(A2)_{T^=}$ depending on T with the fixed given T and also, $(\mathrm{Cut})_{T^=}$ depending on T. Naturally, we will write

$$T \vdash \sigma \text{ equivalently to } \sigma \in \vdash_T .$$

The next thing is to formulate the basic rationale behind the proof system, viz. its *soundness*. We will do this here for the Boolean valued interpretation.

Let M be a Boolean valued model. We say that M is a model T and write $M \models T$, if every $\sigma \in T$ has value 1 in M. We write $T \overset{\mathrm{b}}{\models} \sigma$ to mean that every Boolean valued model of T is a model of σ.

Theorem 5.1.1 $T \vdash \sigma$ *implies that* $T \overset{\mathrm{b}}{\models} \sigma$.

The proof consists in a straightforward verification that all axioms have always value 1, and the whenever $M \models T$ and $M \models \sigma$ for (all) the premise(s) of an application of a rule of inference, then $M \models \sigma$ for the conclusion σ. Note that, in a roundabout way, Theorem 5.1.1 follows from our earlier soundness result concerning categories and from the discussion concerning Boolean valued models and $\mathrm{Sh}_{\mathcal{B}}$-valued functors. To indicate that indeed there is something to prove, we mention that the variable conditions (e.g. the one in the cut rule), will be used essentially in verifying soundness.

Theorem 5.1.2 (Boolean completeness theorem for $L_{\infty\omega}$)

$$T \overset{\mathrm{b}}{\models} \sigma \text{ implies } T \vdash \sigma.$$

In fact, given any fragment F which is a set *(as opposed to a* proper class*) and theory T in F, there is a complete Boolean algebra $\mathcal{B}$ and $\mathcal{B}$-valued model M such that for any sequent σ of F,*

$$M \models \sigma \quad \text{iff} \quad T \vdash \sigma.$$

Proof. We stipulate that the set of all variables be countable, more precisely, for each sort s, the set of variables of sort s be a denumerable infinite set, Var_s. (Hence, in case the language L is countable, the domain $|M|$ of our model defined below will also be countable.)

We start defining M by setting

$$|M|_s = \mathrm{Term}_s$$

where Term_s is the set of all terms of sort s. We define the operations corresponding to the operation symbols in L in the obvious way:

$$f^M(t_1, \dots, t_n) \underset{\mathrm{df}}{=} f t_1 \cdots t_n$$

for all suitable t_i and f.

The Boolean algebra $\mathcal{B}$ will be defined through a partially ordered set $\mathcal{P}$. We define P as the set of all sequents

$$p := \Phi \Rightarrow \Psi$$

such that

$$T \not\vdash \Phi \Rightarrow \Psi,$$

and we define for p as above and for

$$p' := \Phi' \Rightarrow \Psi'$$

the partial ordering on P by

$$p' \leq p \ (\Leftrightarrow \ p' \text{ extends } p) \underset{\mathrm{df}}{\Leftrightarrow} \Phi \subset \Phi' \text{ and } \Psi \subset \Psi'.$$

$\mathcal{P}$ is defined as the partially ordered set $\langle P, \leq \rangle$. We put $\mathcal{B}$ to be $\mathcal{P}^*$ (c.f. Chapter 4). $\mathcal{B}$ is a complete Boolean algebra.

For $p := \Phi \to \Psi$, we write Φ_p for Φ and Ψ_p for Ψ.

For an arbitrary formula $\phi \in F$ put

$$U_\phi \underset{\mathrm{df}}{=} \{p \in P : \phi \in \Phi_p\}.$$

Clearly, U_ϕ is an open set.

Similarly, let us define the open set

$$V_\phi \underset{\mathrm{df}}{=} \{p \in P : \phi \in \Psi_p\}.$$

To define M, put

$$\begin{aligned} R^M(t_1, \dots, t_n) \ &\underset{\mathrm{df}}{=} \ U^*_{Rt_1 \cdots t_n}, \\ t_1 \approx^M t_2 \ &\underset{\mathrm{df}}{=} \ U^*_{t_1 \approx t_2}, \end{aligned}$$

for terms $t_1, \ldots, t_n$ and the predicate symbol R satisfying the compatibility conditions concerning sorts.

To compete the definition of M, we define the membership functions (c.f. Chapter 4) $\| \cdot \| = \| \cdot \|_s : |M|_s \to \mathcal{B}$ by $\|t\| = U_t^*$ where

$$U_t = \{p \in P : \text{every (free) variable in } t \text{ occurs free in } \Phi_p \cup \Psi_p\}.$$

It follows that $\|t\| = \|\vec{x}\|$ where $\vec{x}$ is the sequence of free variables in t. Similarly, $\|\vec{t}\| = \|\vec{x}\|$, if $\vec{x}$ is the sequence of variables occurring free in at least one member of the sequence $\vec{t}$.

For an arbitrary formula $\theta \in L_{\infty\omega}$, $\|\theta\|$ denotes the value of θ in M when each free variable in θ denotes itself (every such variable being actually an element of the domain of M). Remember that now the (Boolean) value depends (through factors which are values of the membership functions) sensitively on which variables are being interpreted. Now, in the notation $\|\theta\|$ we understand that the variables being interpreted are *exactly* the ones free in θ, i.e. $\|\theta\|$ stands for

$$\|\theta[x_1 \text{ for } x_1, \ldots, x_n \text{ for } x_n]\|_M$$

with $x_1, \ldots, x_n$ exactly the distinct free variables of θ.

Merely on the basis that $|M|$ consists of terms with the operations indexed by the operation symbols in L defined in the trivial way, we can show by an easy induction that

$$\|\theta[x_1 \text{ for } x_1, \ldots, x_n \text{ for } x_n]\| = \|\vec{t}\| \cdot \|\theta(t_1, \ldots, t_n)\|$$

where on the right hand side, we have *substitution* of t_i for x_i. Here $\vec{t} = \langle t_1, \ldots, t_n \rangle$ and $x_1, \ldots, x_n$ include all the free variables of θ.

Lemma 5.1.3 $U_\phi^* \wedge V_\phi^* = 0$*, for any atomic* ϕ*.*

Proof. Since $U_\phi^* \wedge V_\phi^* = (U_\phi \cap V_\phi)^*$, it is enough to see that $U_\phi \cap V_\phi = \varnothing$. But it is clear that any element p of $U_\phi \cap V_\phi$ would be an axiom in group $(A1)$, contradicting the fact that every $p \in P$ is note derivable from T. □

Lemma 5.1.4 *For an arbitrary formula* ϕ*, we have*

$$U_\phi^* \leq \|\phi\| \tag{1}$$

$$\text{and} \quad V_\phi^* \wedge \|\phi\| = 0. \tag{2}$$

Proof. The proof is an induction on the complexity of ϕ. For an atomic formula $\phi = Pt_1 \ldots t_n$, recall the definition of the valuation $\|\phi\|$:

$$\|\phi\| \underset{\text{df}}{\Leftrightarrow} \|\vec{x}\| \cdot P^M(t_1, \ldots, t_n)$$

where $\|\vec{x}\| = \|x_1\| \cdots \|x_n\|$ and $\vec{x} = \langle x_1, \ldots, x_n \rangle$ are all the free variables in ϕ.

We have $P^M(t_1, \ldots, t_n) = U_{Pt_1 \cdots t_n}^* = U_\phi^*$ and clearly $\|\vec{x}\| \geq U_\phi^*$, hence

$$\|\phi\| = U_\phi^*.$$

This, together with 5.1.3, establishes the two claimed relations for $\phi = Pt_1 \cdots t_n$. Of course, the other type of atomic formula, $\phi := t_1 \approx t_2$ is handled similarly.

The rest of the proof consists of the induction steps according to the definition of formulas.

Case 1. $\phi = \bigwedge\Theta$.

The induction hypothesis is that for each $\theta \in \Theta$, $U^*_\theta \leq \|\theta\|$ and $V^*_\theta \cap \|\theta\| = \varnothing$. The proof of (1) for ϕ will use the rule $(\bigwedge \Rightarrow)$, i.e. the fact that the formal theorems derivable from T are closed under the rule $(\bigwedge \Rightarrow)$, whereas the proof of (2) will use the rule $(\Rightarrow \bigwedge)$.

Let $p \in U_{\bigwedge\Theta}$, $\theta \in \Theta$ and $p' \in P$, $p' \leq p$, $p' := \Phi \Rightarrow \Psi := \Phi, \Theta \Rightarrow \Psi$ (since $\bigwedge\Theta \in \Phi$). By the definition of P and $p' \in P$, we have that

$$T \not\vdash \Phi, \bigwedge\Theta \Rightarrow \Psi.$$

Hence "by the rule $(\bigwedge \Rightarrow)$",

$$T \not\vdash \Phi, \bigwedge\Theta, \theta \Rightarrow \Psi.$$

In other words, for $p'' = \Phi, \bigwedge\Theta, \theta \Rightarrow \Psi$ we have $p'' \in P$, hence obviously, $p'' \in U_\theta$ and $p'' \leq p'$. We have shown that

$$(\forall p' \leq p)(\exists p'' \leq p')p'' \in U_\theta,$$

i.e., that $p \in U^*_\theta$. Since $p \in U_\phi$ was arbitrary, we have that $U_\phi \subset U^*_\theta$, thus $U^*_\phi \leq U^*_\theta$. Since this holds for any $\theta \in \Theta$, we have

$$U^*_\phi \leq \bigwedge\nolimits_{\theta\in\Theta} U^*_\theta \leq \bigwedge\nolimits_{\theta\in\Theta}\|\theta\| = \|\bigwedge\Theta\| = \|\phi\|$$

where for the second inequality we used the induction hypothesis and for the last but one equality, the definition of $\|\cdot\|$. This establishes (1) in Case 1.

For (2), we first prove

$$V_{\bigwedge\Theta} \leq \bigvee\nolimits_{\theta\in\Theta} V^*_\theta. \tag{3}$$

Recall that $\bigvee_{\theta\in\Theta} V^*_\theta = (\bigcup_{\theta\in\Theta} V_\theta)^*$ Let $p \in V_{\bigwedge\Theta}$, $p' \leq p$, $p' \in P$, $p' := \Phi \Rightarrow \Psi := \Phi \Rightarrow \Psi, \bigwedge\Theta$. Consider $p''_\theta \underset{\text{df}}{=} \Phi \Rightarrow \Psi, \bigwedge\Theta, \theta$ for each $\theta \in \Theta$. It can happen that none of the p''_θ is in $\bigcup_{\theta\in\Theta} V_\theta$ only if none of the p''_θ is in P, i.e., if $T \vdash p''_\theta$ for each $\theta \in \Theta$. But then by $(\bigwedge \Rightarrow)$ we obtain $T \vdash p'$, contradicting $p' \in P$. Hence, there is $\theta \in \Theta$ such that $p''_\theta \in \bigcup_{\theta\in\Theta} V_\theta$. Also, $p''_\theta \leq p'$. We have shown

$$\forall p' \leq p \; \exists p'' \leq p' \; p'' \in \bigcup_{\theta\in\Theta} V_\theta,$$

i.e. $p \in (\bigcup_{\theta\in\Theta} V_\theta)^* = \bigvee_{\theta\in\Theta} V^*_\theta$. This means that $V_\bigwedge \subset \bigvee_{\theta\in\Theta} V^*_\theta$, showing (3).

Using (3), we argue as follows.

The induction hypothesis implies that

$$\bigvee\nolimits_{\theta\in\Theta} V^*_\theta \cap \bigwedge\nolimits_{\theta\in\Theta}\|\theta\| = \varnothing.$$

Since $\bigwedge_{\theta\in\Theta}\|\theta\| = \|\bigwedge\Theta\|$, this last equality, together with (3) implies (2) as desired.

Case 2, when $\phi = \bigvee\Theta$, is entirely similar to Case 1 and is left to the reader.

Case 3. $\phi = \exists x\theta(x) = \exists w\theta(w)$.

Let $p \in V$, $p' \in P$, $p' \leq p$, $p' := \Phi \Rightarrow \Psi := \Phi \Rightarrow \Psi, \exists x\theta(x)$ as before. Now, let t be a term of the same sort as x such that all variables in t occur in p, i.e., $p \in U_t$, i.e., $p \in U_t \cap U_\phi$. Putting $p'' := \Phi \Rightarrow \Psi, \exists x\theta(x), \theta(t)$, by rule $(\Rightarrow \exists)$ we obtain that $p'' \in P$, hence $p'' \in V_{\phi(t)}$. We have shown that for every $p \in U_t \cap U_\phi$

$$(\forall p' \leq p)(\exists p'' \leq p') \; p'' \in V_{\theta(t)}$$

i.e.

$$U_t \cap U_\phi \subset V^*_{\theta(t)}.$$

It follows that

$$\|t\| \cdot V^*_\phi \le V^*_{\theta(t)} \tag{4}$$

for an arbitrary term $t \in |M|_s$ where s is the sort of the variable x. Recall that by the remark we made on substitution, we have

$$\|\theta(t)\| = \|\theta[t \text{ for } x, x_1 \text{ for } x_1, \ldots, x_n \text{ for } x_n]\|$$

(recall that $\vec{x} = \langle x_1, \ldots, x_n \rangle$ are all the free variables of $\exists x\theta(x)$. Since the variable x actually occurs free in $\theta(x)$ (remember our restriction of forming $\exists x\theta(x)$) all free variables of t will actually occur free in $\theta(t)$, hence

$$\|\theta(t)\| \le \|t\|$$

or equivalently

$$\|t\| \cdot \|\theta(t)\| = \|\theta(t)\|.$$

Recalling the definition of $\|\cdot\|$, we now have

$$\begin{aligned} \|\exists x\theta(x)\| &= \|\exists x\theta(x)[x_i \text{ for } x_i] = \textstyle\bigvee_{t\in|M|_s} \|t\| \cdot \|\theta[t \text{ for } x, x_i \text{ for } x_i]\| \\ &= \textstyle\bigvee_{t\in|M|_s} \|t\| \cdot \|\theta(t)\| = \bigvee_{t\in|M|_s} \|\theta(t)\|. \end{aligned} \tag{5}$$

The formulas $\theta(t)$ have complexity smaller that that of $\exists x\theta(x)$, hence the induction hypothesis applies and we get

$$V^*_{\theta(t)} \wedge \|\theta(t)\| = 0.$$

Using (4) and the last statement, we get

$$\|t\| \cdot V^*_\phi \wedge \|\theta(t)\| = 0.$$

Since $\|\theta(t)\| \le \|t\|$, this implies

$$V^*_\phi \wedge \|\theta(t)\| = 0.$$

Using the expression (5) for $\|\phi\| = \|\exists x\theta(x)\|$, we get

$$V^*_\phi \wedge \|\phi\| = V^*_\phi \wedge \textstyle\bigvee_{t\in|M|_s} \|\theta(t)\| = \bigvee_{t\in|M|_s} (V^*_\phi \wedge \|\theta(t)\|) = 0$$

as required for (2).

The other part (1) of the claim is shown by using ($\exists \Rightarrow$) as follows.

Let $p \in U$, $p' \in P$, $p' \le p$. Let y be a free variable of the same sort as x such that y does not occur free in p' (there are altogether finitely many free variables in p'). Put $p'' := \Phi_{p'}, \theta(y) \Rightarrow \Psi_{p'}$. By rule ($\exists \Rightarrow$), we obtain that $p'' \in P$ and hence of course $p'' \in U_{\theta(y)}$. What we have shown is

$$(\forall p' \le p)(\exists p'' \le p')\ p'' \in \textstyle\bigcup_{t\in|M|_s} U_{\theta(t)},$$

hence, since $p \in U_\phi$ was arbitrary,

$$U_\phi \subseteq (\textstyle\bigcup_{t\in|M|_s} U_{\theta(t)})^* = \bigvee_{t\in|M|_s} U^*_{\theta(t)}.$$

The induction hypothesis tells us that $U^*_{\theta(t)} \leq \|\theta(t)\|$ and we know from the previous part of the proof that $\|\exists x\theta(x)\| = \bigvee_{t\in|M|_s}\|\theta(t)\|$. Hence

$$U^*_\phi \leq \bigvee_{t\in|M|_s} U^*_{\theta(t)} \leq \bigvee_{t\in|M|_s}\|\theta(t)\| = \|\exists x\theta(x)\|$$

as desired.

The last case, $\phi = \forall\vec{z}(\phi_1 \to \psi)$, is similar to, though a bit more complicated than, the case of the existential quantifier. We have to use both parts of the induction hypothesis, referring to U and V, to show each of the statements for ϕ. The details are omitted. □

Lemma 5.1.5 *For a formula allowed as a cut-formula in the cut-rule* $(\mathrm{Cut})_T$*, i.e.,* ϕ *a substitution instance of a formula in an axiom in* $T^=$*, we have*

$$U^*_\phi \vee V^*_\phi = \|\vec{x}\|$$

where $\vec{x}$ *is the sequence of free variables in* ϕ*.*

Proof. It is clear that the left side is contained in the right. Let $p \in U_{\vec{x}}$ and $p' \leq p$, $p' := \Phi \Rightarrow \Psi$. Hence, all of $\vec{x}$ occurs free in p'. Hence the following is a permissible application of $(\mathrm{Cut})_T$:

$$\frac{\Phi, \phi \Rightarrow \Psi \qquad \Phi \Rightarrow \Psi, \phi}{\Phi \Rightarrow \Psi}.$$

According to the definition of P, this implies that *either* there is $p'' \leq p'$, namely $\Phi, \phi \Rightarrow \Psi$, such that $p'' \in U_\phi$, *or* there is $p'' \leq p'$, namely $\Phi \Rightarrow \Psi, \phi$, such that $p'' \in V_\phi$. In other words, $U_{\vec{x}} \subseteq U_\phi \vee V_\phi$. This is sufficient for 5.1.5. □

Corollary 5.1.6 *For a formula* ϕ *as in* 5.1.5*,*

$$\|\phi\| = U^*_\phi = \|\vec{x}\| - V^*_\phi.$$

Proof. This follows immediately from 5.1.4 and 5.1.5.

End of proof of 5.1.2 We can now show that M is a *model* of $T^=$. Let $\sigma := \Theta \Rightarrow \Gamma$ be an axiom belonging to $T^=$, let $\vec{x}$ be the sequence of the fee variables in σ, and let $\vec{t}$ be a sequence of elements of the model with sorts matching those of $\vec{x}$. We claim that $\beta = 0$ in $\mathcal{B}$, where $\beta \underset{\mathrm{df}}{=} \|\bigwedge\theta[\vec{t}]\| - \|\bigvee\Gamma[\vec{t}]\|$. computing β we find:

$$\beta = \|\vec{t}\| \cdot \bigwedge_{\theta\in\Theta}\|\theta(t)\| \cdot \bigwedge_{\gamma\in\Gamma} - \|\gamma(t)\|$$

using $\|\phi[\vec{t}]\| = \|\vec{t}\| \cdot \|\phi(\vec{t})\|$.

If $\vec{x_\gamma}$ is the sequence of free variables in $\gamma(\vec{t})$, then $\|\vec{x_\gamma}\| \cdot (-\|\gamma(\vec{t})\|) = V^*_{\gamma(\vec{t})}$ by 5.1.6. Since clearly $\|\vec{x_\gamma}\| \geq \|\vec{t}\|$ for $\gamma \in \Gamma$,

$$\|\vec{t}\| \cdot (-\|\gamma(t)\|) = \|\vec{t}\| \cdot V^*_{\gamma(\vec{t})}$$

and

$$\beta = \|\vec{t}\| \cdot \bigwedge_{\theta\in\Theta} U^*_{\theta(\vec{t})} \cdot \bigwedge_{\gamma\in\Gamma} V^*_{\gamma(\vec{t})}.$$

Since each t in $\vec{t}$ does actually occur in a $\theta(\vec{t})$ or $\gamma(\vec{t})$, the factor $\|\vec{t}\|$ can be omitted and we obtain

$$\beta = \bigwedge_{\theta\in\Theta} U^*_{\theta(\vec{t})} \cdot \bigwedge_{\gamma\in\Gamma} V^*_{\gamma(\vec{t})} = (\bigcap_{\theta\in\Theta} U_{\theta(\vec{t})} \cap \bigcap_{\gamma\in\Gamma} V_{\gamma(\vec{t})})^*.$$

But any $p \in P$ that would belong to the intersection under the last $(\ )^*$, would be an *axiom* according to $(A2)_{T^=}$, hence we would have $T \vdash p$, contradicting $p \in P$. So the intersection is empty, and $\beta = 0$ as claimed. We thus obtain that

$$\|\bigwedge\Theta[\vec{t}]\| \leq \|\bigvee\Gamma[\vec{t}]\|$$

for any $\vec{t}$, i.e., $M \models \Theta \Rightarrow \Gamma$.

This completes proving that M is a model of $T^=$.

Since the equality axioms are included in $T^=$, our model M also satisfies the requirements concerning the interpretation of $\approx$.

By the soundness theorem, 5.1.1, it follows that $M \models \sigma$ for any σ such that $T \vdash \sigma$. It remains to show the converse.

Assume $T \nvdash \sigma$, $\sigma = \Phi \Rightarrow \Psi$ and let $\vec{x}$ be all the free variables in σ. But then σ is an element of P. Consider now the Boolean value $U^*_\sigma = \{p \in P : p \leq \sigma\}^*$. This is a nonzero value (since it is a *non-empty* set). By 5.1.4, we easily infer that $U^* \leq \|\vec{x}\| \cdot \|\bigwedge\Phi\|$ and $U^*_\sigma \cap \|\bigvee\Psi\| = 0$, and thus $\|\vec{x}\| \cdot \|\bigwedge\Phi\| \not\leq \|\bigvee\Psi\|$, and a fortiori

$$\|\sigma\| \underset{\mathrm{df}}{=} \bigwedge\nolimits_{t_1,\ldots,t_n}(\|t_1\|\cdots\|t_n\| \to (\|t_1\|\cdots\|t_n\| \cdot \|\bigwedge\Phi(t_1,\ldots,t_n)\| \to$$
$$\|\bigvee\Psi(t_1,\ldots,t_n)\|)) \neq 1_{\mathcal{B}}, \quad \text{i.e.,} \quad M \not\models \sigma.$$

This completes the proof of 5.1.2. □

We will now comment on an obvious variant of the above proof-system. Let us replace the restricted $(\mathrm{Cut})_T$ by the unrestricted (Cut) where ϕ is allowed to be any formula of F but the variable condition is retained. This rule, and hence the whole resulting system, is still sound, in fact even in a general categorical interpretation (c.f. Chapter 3, Theorem 3.2.8). Let us denote derivability in this new system by $\vdash'$. The we can construct a canonical $M' = M'_T$ quite the same way as M was constructed above, and we will actually have

$$\|\phi\|_{M'} = U^*_\phi = \|\vec{x}\| - V^*_\phi \qquad \text{for any } \phi,$$

by the argument for 5.1.5 above. (Here, of course we use notation signifying the *analogues* of previous entities, although not quite the same ones; e.g. even the partial ordering is different now!) Thus M' behaves in a more 'determined' way than M. Needless to say, M' also satisfies Theorem 5.1.2.

Let us now state the two-valued variants of the completeness theorem.

Theorem 5.1.7 *Assume* either *that* (i) *the fragment F is finitary, i.e. F contains no infintary disjunction or conjunction (i.e., $\bigwedge\Sigma \in F$ or $\bigvee\Sigma \in F$ implies that Σ is finite)* or *that* (ii) *F is countable (which implies that the language L can be taken to be countable too).*

Then $T \vdash \sigma$ if and only if $T \models \sigma$, where $T \models \sigma$ means that every set-valued model of T is a model of σ.

Proof. All the remaining work for this proof was done in Chapter 4. Consider the canonical $\mathcal{B}$-valued model M of T, constructed in the proof of 5.1.2. Assume $T \nvdash \sigma$. Hence $\|\sigma\|_M \neq 1_{\mathcal{B}}$. Then by 4.3.2 or 4.3.5, there is a set-valued model of T that does not satisfy σ. □

Corollary 5.1.8 (Compactness Theorem for finitary logic.) *Suppose that each finite subset of a set T of axioms in finitary logic $L_{\omega\omega}$ has an (ordinary* **Set***-) model. Then T has a model.*

Proof. In case of a fragment in the finitary logic, it is almost obvious that $T \vdash \sigma$ implies that $T' \vdash \sigma$ for some finite $T' \subset T$. Applying this and 5.1.7 for σ = false, the assertion follows. $\square$

§2 Completeness of a "one sided" system for coherent logic

A *coherent fragment* F of $L_{\infty\omega}$ is a fragment whose formulas are coherent i.e. built up from atomic formulas of L by using $\bigvee$ applied to arbitrary sets of formulas, $\bigwedge$ applied only to finite sets of formulas and $\exists$. A *(one-sided) sequent* of F is a sequent $\Theta \Rightarrow \psi$ with Θ a finite set of formulas in F and ψ a single formula in F.

Let T be a *theory in* F, i.e. a set of (one-sided) sequents of F. Next we define a formal system, relative to T; for a sequent σ, we write $T \vdash \sigma$ to denote the fact that σ is deducible in the formal system.

Axioms: $\quad \Theta \Rightarrow \psi \quad$ if $\psi \in \Theta$.

Rules of inference:

$(R\bigwedge_1)$ $$\frac{\Theta, \bigwedge\Sigma, \phi \Rightarrow \psi}{\Theta, \bigwedge\Sigma \Rightarrow \psi} \qquad \text{if } \phi \in \Sigma$$

$(R\bigwedge_2)$ $$\frac{\Theta, \bigwedge\Sigma \Rightarrow \psi}{\Theta \Rightarrow \psi} \qquad \text{if } \Sigma \subset \Theta$$

$(R\bigvee_1)$ $$\frac{\Theta, \phi, \bigvee\Sigma \Rightarrow \psi}{\Theta, \phi \Rightarrow \psi} \qquad \text{if } \Sigma \subset \Theta$$

and all free variables in Σ occur fee in the conclusion.

$(R\bigvee_2)$ $$\frac{\{\Theta, \bigvee\Sigma, \phi \Rightarrow \psi : \phi \in \Sigma\}}{\Theta, \bigvee\Sigma \Rightarrow \psi}$$

$(R\exists_1)$ $$\frac{\Theta, \phi(v/t), \exists x\phi(v/x) \Rightarrow \psi}{\Theta, \phi(v/t) \Rightarrow \psi}$$

here, of course, v and t are of the same sort as x is.

$(R\exists_2)$ $$\frac{\Theta, \exists x\phi(v/x), \phi \Rightarrow \psi}{\Theta, \exists x\phi(v/x) \Rightarrow \psi}$$

if v does not occur free in the conclusion.

(RT) $$\frac{\Theta, \Theta'(t_1, \ldots, t_n), \psi'(t_1, \ldots, t_n) \Rightarrow \psi}{\Theta, \Theta(t_1, \ldots, t_n) \Rightarrow \psi}$$

provided for some $\Theta'(v_1, \ldots, v_n) \Rightarrow \psi'(v_1, \ldots, v_n)$ belonging to $T^{=}$ (for $T^{=}$, c.f. last section) and for some terms $t_1, \ldots, t_n$, $\Theta'(t_1, \ldots, t_n)$ is the set of substitution instances $\theta(t_1, \ldots, t_n)$ of all $\theta(v_1, \ldots, v_n)$ in Θ', all the free variables in Θ' of ψ' are among $v_1, \ldots, v_n$, $\psi'(t_1, \ldots, t_n)$ is the result of substitution of the t_i for v_i in ψ' and finally, all free variables that occur in the premise occur in the conclusion.

(Comment: The rule (RT) depends on the fixed set T of sequents.)

We define (as expected) the set of *formal consequences* of T relative to the given system to be the smallest set $\mathcal{X}$ of sequents such that (i) each axiom is in $\mathcal{X}$, and (ii) whenever

$$\frac{\{\sigma_i : i \in I\}}{\sigma}$$

is an instance of a rule of inference and $\sigma_i \in \mathcal{X}$ for all $i \in I$, then also $\sigma \in \mathcal{X}$. We write $T \vdash \sigma$ for "σ is a formal consequence of T", i.e., for $\sigma \in \mathcal{X}$.

We will state a completeness theorem for the proof-system $T \vdash (\cdot)$ analogous to the one proved in the previous section. But for purposes of the later Chapter 7, we will state a slightly more general theorem.

For a theory T as above, a *T-consistency property* is defined to consist of items of the following kinds:

(i) a partially ordered set $\mathcal{P} = (P, \leq) = (P, \leq_{\mathcal{P}})$;

(ii) a function f defined on P such that for $p \in P$, $f(p)$ is a finite set of formulas in F;

(iii) a function Var defined on P such that for every $p \in P$, $\mathrm{Var}(p)$ is a finite set of free variables such that all free variables in $f(p)$ are in the set $\mathrm{Var}(p)$.

The following properties are required to be satisfied (p and q denote members of P):

(iv) if $q \leq p$, then $f(p) \subseteq f(q)$, and $\mathrm{Var}(p) \subseteq \mathrm{Var}(q)$;

(v) if $\bigwedge\Sigma \in f(p)$ and $\phi \in \Sigma$, then there is $q \leq p$ such that $\phi \in f(q)$;

(vi) if $\Sigma \subset f(p)$, then there is $q \leq p$ such that $\bigwedge\Sigma \in f(q)$;

(vii) if $\phi \in \Sigma$, all free variables in $\bigvee\Sigma$ belong to $\mathrm{Var}(p)$ and $\phi \in f(p)$, then there is $q \leq p$ such that $\bigvee\Sigma \in f(q)$;

(viii) if $\bigvee\Sigma \in f(p)$, then there are $\phi \in \Sigma$ and $q \leq p$ such that $\phi \in f(q)$;

(ix) if $\phi(v/t) \in f(p)$, then there is $q \leq p$ such that $\exists x \phi(v/x) \in f(q)$;

(x) if $\exists x \phi(v/x) \in f(p)$, then there is a free variable u and some $q \leq p$ such that $\phi(v/u) \in f(q)$;

(xi) if $\Theta'(v_1, \ldots, v_n) \Rightarrow \psi'(v_1, \ldots, v_n)$ belongs to the theory $T^{=}$ and $\Theta'(t_1, \ldots, t_n) \subseteq f(p)$ (compare (RT) above), then there is $q \leq p$ such that $\psi'(t_1, \ldots, t_n) \in f(q)$, *provided* all (free) variables in $t_1, \ldots, t_n$ belong to $\mathrm{Var}(p)$.
(*End of definition of consistency property*).

The conditions (v)-(xi) correspond to the rules of the above formal system. In fact, the following connection can be made. Define P to be the set of all sequents $\Theta \Rightarrow \psi$ such that $T \not\vdash \Theta \Rightarrow \psi$. Define the partial order $\leq$ on P by

$$\Theta' \Rightarrow \psi' \leq \Theta \Rightarrow \psi \qquad \text{iff} \qquad \Theta \subseteq \Theta' \text{ and } \psi' = \psi.$$

Let $f(\Theta \Rightarrow \psi) = \Theta$ and $\mathrm{Var}(\Theta \Rightarrow \psi) =$ the set of free variables in $\Theta \cup \{\psi\}$.

Proposition 5.2.1 *With these definitions, $\langle P, \leq, f, \mathrm{Var}(\cdot)\rangle$ is a T-consistency property.*

The verification of 5.2.1 is straightforward.

Next, we describe a Boolean-valued model constructed on the basis of a consistency property; it will be very similar to the construction of §1.

Let $\langle P, \leq, f, \mathrm{Var}(\cdot)\rangle$ be a T-consistency property. Let $\mathcal{P}^* = \mathcal{B}$ be the Boolean algebra of regular open subsets of $\mathcal{P} = (P, \leq)$ as before. Let the universe of the $\mathcal{B}$-valued model M consist of all terms of L, more precisely, let $|M|_s$ be the set of terms of sort s. The operations denoted by symbols in L are defined as before. We put $U_t = \{p \in P$:all variables in t belong to $\mathrm{Var}(p)\}$ and $\|t\| = \|t \in |M|_s\| = U_t^*$. Also, for any formula ϕ in F, we put $U_\phi = \{p \in P : \phi \in f(p)\}$, and we define

$$R^M(t_1, \ldots, t_n) = U^*_{Rt_1\cdots t_n}$$

and

$$\|t_1 \approx t_2\|_M = U^*_{t_1 \approx t_2}.$$

It follows as before that M is a properly defined $\mathcal{B}$-structure. Let us call M the *canonical model* associated with the consistency property.

Proposition 5.2.2 *Let M be the canonical model of a T-consistency property. With the notation above, we have*

(i) $\|\phi[\vec{t} \text{ for } \vec{v}]\|_M = U^*_{\phi(\vec{t} \text{ for } \vec{v})}$,

(ii) *M is a model of T.*

The proof is very similar to the proofs in §1 and the details are omitted.
By 5.2.1 and 5.2.2 we obtain

Corollary 5.2.3 (Completeness of the coherent system). (a) *For a theory T in F and a sequent σ of F,*

$$T \vdash \sigma \;\text{ iff }\; T \overset{\mathrm{b}}{\models} \sigma$$

for Boolean valued models.

(b) *For a theory T in a* countable or *in a* finitary *coherent fragment F and σ a sequent in F,*

$$T \vdash \sigma \;\text{ iff }\; T \models \sigma$$

for ordinary (set-) models.

Finally we wish to point out the principle behind finding the one-sided axiom system just like we did for the other axiom system. Assume F is a countable fragment of $L^g_{\infty\omega}$, T and σ are from F, $T \not\vdash \sigma$ and try to construct a model M of T not satisfying σ. We wish to construct M on the basis of a set H (instead of two sets H_1, H_2) of formulas such that the domain of M consists of the terms whose variables occur free in at least on formula in H and such that for any formula $\phi(\vec{x})$ in F with free variables in $|M|$

$$\phi(\vec{x}) \text{ is true in } M \;\Leftrightarrow\; \phi(\vec{x}) \in H.$$

(Compare the corresponding conditions for the other axiom system.) Now we again have that e.g. for $\Sigma \in F$ the following should be satisfied:

$$\bigwedge\Sigma \in H \;\text{ iff }\; \Sigma \subset H.$$

Instead of reformulating this condition with the aid of an H_2 as was done before, we are going to make sure directly that this equivalence holds; we will succeed because now Σ is *finite*. In particular, we again construct H as an increasing union $\bigcup_{n<\omega} H_n$ of finite sets H_n. Let σ, the fixed sequent given initially, be $\Theta \Rightarrow \psi$. We put $H_0 = \Theta$ and as an induction hypothesis we require

$$T \not\vdash H_n \Rightarrow \psi.$$

We will make sure that (1) holds in the following way. In one kind of step of the construction we will have that $\bigwedge\Sigma \in H_n$ and we want to put $H_{n+1} = H \cup \Sigma$; in another, we will have $\Sigma \subset H_n$ and $H_{n+1} = H_n \cup \{\bigwedge\Sigma\}$. Notice that since Σ is finite, if $\Sigma \subset H$ then for some n_0, $\Sigma \subset H_n$ for all $n \geq n_0$ hence we will have ample opportunity to perform the second kind of construction and make sure that $\bigwedge\Sigma \in H$. This possibility is not available for infinite Σ.

Now reflection shows that the possibility of performing these two kinds of constructions calls for exactly the two rules $(R\bigwedge_1)$ and $(R\bigwedge_2)$ given above.

Chapter 6

Existence theorems on geometric morphisms of topoi

§1 Preliminaries

Here we collect some simple facts and notation on Grothendieck topoi we will need. For more details, c.f. Chapter 1.

A *site* is a category together with a Grothendieck topology. In what follows, $\mathcal{C}$ will denote a site whose underlying category (also denoted by $\mathcal{C}$) is small and which has finite $\varprojlim$. A Grothendieck topology can be given by specifying a collection $\mathcal{G}_0(\mathcal{C})$ of families each of which is of the form $\{A_i \xrightarrow{f_i} A : i \in I\}$. Such a family can be called a *basic covering family.*

We note that an arbitrary collection $\mathcal{G}_0(\mathcal{C})$ of families in $\mathcal{C}$ of the form given above generates a smaller Grothendieck topology $\mathcal{G}(\mathcal{C})$ in which each family in $\mathcal{G}_0(\mathcal{C})$ is a *covering family*, i.e., $\mathcal{G}_0(\mathcal{C}) \subset \mathcal{G}(\mathcal{C})$ (c.f. Chapter 1, Section 1). Naturally, different collections of 'basic' covering families might generate the same topology. From our point of view, it is more natural to consider the site $\mathcal{C}$ to be given by the underlying category with $\mathcal{G}_0(\mathcal{C})$ rather than $\mathcal{G}(\mathcal{C})$.

$\widetilde{\mathcal{C}}$ denotes the category of all sheaves over the site $\mathcal{C}$; c.f. Chapter 1, Section 2. A Grothendieck topos is, by definition, a category equivalent to $\widetilde{\mathcal{C}}$ for some small site $\mathcal{C}$. We have the canonical functor $\varepsilon = \varepsilon_{\mathcal{C}}$

$$\varepsilon : \mathcal{C} \to \widetilde{\mathcal{C}}$$

that is left exact (i.e., preserves finite left limits) and continuous; a left exact functor $F : \mathcal{C} \to \mathcal{R}$ is *continuous* if for every basic covering family $\{A_i \xrightarrow{f_i} A : i \in I\} \in \mathcal{G}_0(\mathcal{C})$, we have that $\{FA_i \xrightarrow{Ff_i} FA : i \in I\}$ is an effective epimorphic family, i.e., $F(A) = \bigvee_{i \in I} \exists_{F(f_i)}(F(A_i))$; c.f. Proposition 3.3.3.

We will call a left exact and continuous functor $F : \mathcal{C} \to \mathcal{R}$ an *$\mathcal{R}$-model of $\mathcal{C}$*.

A geometric morphism $\mathcal{E}_1 \xrightarrow{u} \mathcal{E}_2$ is a pair of adjoint functors $\mathcal{E}_1 \underset{u_*}{\overset{u^*}{\rightleftarrows}} \mathcal{E}_2$ such that u^* is left exact and *continuous* in the sense that whenever for $\{A_i \xrightarrow{f_i} A : i \in I\}$ in $\mathcal{E}_1$ we have $A = \bigvee^{(\mathcal{E}_1)}_{i \in I} \exists_{f_i}(A_i)$ then also $F(A) = \bigvee^{(\mathcal{E}_2)}_{i \in I} \exists_{Ff_i}(FA_i)$ (for equivalent definitions, consult Theorem 1.3.11 in Chapter 1). u_* is determined by u^* up to isomorphism; a functor $u^*: \mathcal{E}_1 \to \mathcal{E}_2$ with the said properties will be called an *$\mathcal{E}_2$-model of $\mathcal{E}_1$*. Hence, geometric morphisms of $\mathcal{E}_2$ to $\mathcal{E}_1$ can essentially be identified with $\mathcal{E}_2$-models of $\mathcal{E}_1$.

$\widetilde{\mathcal{C}}$ has the following universal property, c.f. Corollary 1.3.14. For any Grothendieck topos $\mathcal{E}$ and any $\mathcal{E}$-model M of $\mathcal{C}$, there is an $\mathcal{E}$-model of $\widetilde{\mathcal{C}}$, $\widetilde{M}$, such that

$$\begin{tikzcd} \mathcal{C} \arrow[r, "\varepsilon"] \arrow[dr, "M"'] & \widetilde{\mathcal{C}} \arrow[d, "\widetilde{M}"] \\ & \mathcal{E} \end{tikzcd}$$

is commutative; $\widetilde{M}$ is determined up to isomorphism. Hence to construct a geometric morphism $\mathcal{E} \to \widetilde{\mathcal{C}}$ it is enough to construct an $\mathcal{E}$-model of $\mathcal{C}$.

Next we describe a logical formulation of some of the above notions. We start with the definition of a theory $T_{\mathcal{C}}$ associated with a site $\mathcal{C}$. $T_{\mathcal{C}}$ is formulated in the language $L_{\mathcal{C}}$ associated with the underlying category of $\mathcal{C}$ as given in Chapters 2 and 3; $T_{\mathcal{C}}$ will be a set of sequents in a coherent fragment of $(L_{\mathcal{C}})^g_{\infty\omega}$. $T_{\mathcal{C}}$ is defined to contain (i) all the 'axioms of category' (groups 1 and 2 before 2.4.5) corresponding to identity morphisms and commutative triangles in $\mathcal{C}$ (ii) all the axioms related to finite left limit diagrams in $\mathcal{C}$ (c.f. 2.4.5), (iii) all axioms of the form

$$a \approx a \Rightarrow \bigvee_{i \in I} \exists a (f_i(a_i) \approx a)$$

for a basic covering family $\{A_i \xrightarrow{f_i} A : i \in I\} \in \mathcal{G}_0(\mathcal{C})$.

On the basis of our earlier work, the following proposition in immediate.

Proposition 6.1.1 *M is an $\mathcal{R}$-model of $\mathcal{C}$ if and only if M is an $\mathcal{R}$-model of $T_{\mathcal{C}}$, for any $\mathcal{R}$.*

This gives our basic logical reformulation of the notion of an $\mathcal{E}$-model of $\mathcal{C}$, hence, ultimately, that of geometric morphism $\mathcal{E} \to \widetilde{\mathcal{C}}$.

In subsequent work, a certain subcategory of $\widetilde{\mathcal{C}}$ will be of good use. Let $\mathcal{R}$ be a *full* subcategory of $\widetilde{\mathcal{C}}$ such that $\mathcal{R}$ contains all objects $\varepsilon(A)$ for $A \in \mathrm{Ob}(\mathcal{C})$ and for each subobject $X \hookrightarrow \varepsilon(A)$, it contains an isomorphic copy of X, and conversely, every object in $\mathcal{R}$ is isomorphic to a subobject of some $\varepsilon(A)$, $A \in \mathrm{Ob}(\mathcal{C})$. $\mathcal{R}$ is determined up to equivalence, and it can be taken to be a small category. Moreover, it is easy to see that $\mathcal{R}$ is a *complete logical* category by which we mean that it has finite left limits, stable images and stable sups of arbitrary families of subobjects. Also, the inclusion functor $\mathcal{R} \to \widetilde{\mathcal{C}}$ is left exact and preserves images and arbitrary sups. $\mathcal{R}$ has a nice logical meaning as we now proceed to show.

Next we restate Lemma 1.3.8, part (ii).

Lemma 6.1.2 *If $X \xrightarrow{\xi} \varepsilon A$ is a subobject of εA in $\widetilde{\mathcal{C}}$, $\in \mathrm{Ob}(\mathcal{C})$, then there is a covering family $\{\varepsilon A_i \xrightarrow{g_i} X : i \in I\}$ in $\widetilde{\mathcal{C}}$ (i.e., $X = \bigvee_{i \in I} \exists_{g_i}(\varepsilon A_i)$) such that the compositions ξg_i are of the form $\xi g_i = \varepsilon(f_i)$ for some $A_i \xrightarrow{f_i} A$ in $\mathcal{C}$, for every $i \in I$.*

Now, $\varepsilon: \mathcal{C} \to \widetilde{\mathcal{C}}$, regarded as an interpretation of the language $L_{\mathcal{C}}$ is a *model* of $T_{\mathcal{C}}$. Also, we can regard ε as a functor from $\mathcal{C}$ into $\mathcal{R}$, with $\mathcal{R}$ defined above. We claim

Lemma 6.1.3 *Let A be an object of $\mathcal{C}$, 'a' a variable of sort A in the language $L_{\mathcal{C}}$. The subobjects $X \hookrightarrow \varepsilon A$ in $\widetilde{\mathcal{C}}$ are, up to isomorphism in $\widehat{\mathcal{C}}$, exactly the interpretations $\varepsilon_a(\phi(a))$ in $\widetilde{\mathcal{C}}$ (or in $\mathcal{R}$) of formulas $\phi(a)$ of the form $\bigvee_{i \in I} \phi_i(a)$ where $\langle \phi_i(a) : i \in I \rangle$ is* *an arbitrary family of coherent finitary formulas of the language $L_{\mathcal{C}}$ with the single free variable a. Also, $\varepsilon_x(\phi(\vec{x}))$ for each $\phi(\vec{x})$ in $L^g_{\infty\omega}(L_{\mathcal{C}})$ is, up to isomorphism, in $\mathcal{R}$.*

Proof. Here and below we use the notation $\underset{\sim}{A}_i(a)$ for the formula $\exists a_i(f_i(a_i) = a)$, with a morphism $A_i \xrightarrow{f_i} A$ given by the context. Let X be an object in $\mathcal{R}$, $A \in \mathrm{Ob}(\mathcal{C})$, $X \hookrightarrow \varepsilon A$. Using 6.1.2, we have $A_i \xrightarrow{f_i} A$ as described there. We see that X, as a subobject of $\varepsilon(A)$, equals to $\bigvee^{(\widetilde{\mathcal{C}})}_{i\in I} \exists^{(\mathcal{C})}_{\varepsilon(f_i)}(\varepsilon A_i) = \bigvee^{(\mathcal{R})}_{i} \exists^{(\mathcal{R})}_{\varepsilon(f_i)}(\varepsilon A_i)$. But the latter is $\varepsilon_a(\bigvee_{i\in I} \underset{\sim}{A}_i(a))$, with ε understood as an interpretation of $L_{\mathcal{C}}$ in either $\widetilde{\mathcal{C}}$ or in $\mathcal{R}$. The rest of 6.1.3 is clear. $\square$

Let X be an object of $\mathcal{R}$, $X \hookrightarrow \varepsilon(A)$. X is $\bigvee_{i\in I} \exists_{\varepsilon(f_i)}(\varepsilon A_i)$ as described in 6.1.3. Let M be an $\mathcal{E}$-model of $\mathcal{C}$, i.e., an $\mathcal{E}$-model of $T_{\mathcal{C}}$, and $\widetilde{M} : \widetilde{\mathcal{C}} \to \mathcal{E}$ the corresponding $\widetilde{\mathcal{C}}$-model. As an interpretation of the language L, M gives rise to the interpretation $M(\bigvee_{i\in I} \underset{\sim}{A}_i(a)) = \bigvee^{(\mathcal{E})}_{i\in I} M(\underset{\sim}{A}_i(a)) = \bigvee^{(\mathcal{E})}_{i\in I}(M(A_i) \to M(A))$. But by the commutative diagram defining $\widetilde{M}$ from M, this is exactly $\widetilde{M}(\bigvee_{i\in I} \exists_{\varepsilon(f_i)}(\varepsilon A_i)) = \widetilde{M}(X) \to \widetilde{M}(\varepsilon A)$ ($= M(A)$). In other words, the action of $\widetilde{M}$ on objects of $\mathcal{R}$ is described by interpretations of certain disjunctions by M.

An $\mathcal{E}$-model of $\widetilde{\mathcal{C}}$, $\widetilde{M} : \widetilde{\mathcal{C}} \to \mathcal{E}$ is said to be *conservative* if for subobjects X, Y of A in $\widetilde{\mathcal{C}}$, $\widetilde{M}(X) \leq \widetilde{M}(Y)$ in the subobject lattice of $\widetilde{M}(A)$ (if and) only if $X \leq Y$ in the subobject lattice of A.

Lemma 6.1.4 *For any $\mathcal{E}$-model of $\mathcal{C}$, $M : \mathcal{C} \to \mathcal{E}$, the associated $\mathcal{E}$-model of $\widetilde{\mathcal{C}}$, $\widetilde{M} : \widetilde{\mathcal{C}} \to \mathcal{E}$ is conservative iff for any family $\{A_i \xrightarrow{f_i} A : i \in I\}$ of morphisms of $\mathcal{C}$, if $\{MA_i \xrightarrow{Mf_i} MA : i \in I\}$ is a covering family in $\mathcal{E}$ then $\{\varepsilon A_i \xrightarrow{\varepsilon f_i} \varepsilon A : i \in I\}$ is a covering family in $\widetilde{\mathcal{C}}$. In other words, if the restriction of $\widetilde{M}$ to $\mathcal{R}$ is conservative, then so is widetildeM.*

The 'only if' part is clear. To prove the 'if' part, assume that $X \hookrightarrow C$, $Y \hookrightarrow C$ are subobjects of C in $\widetilde{\mathcal{C}}$. Recall (c.f. 1.3.7) that $\mathcal{C}$ *topologically generates* $\widetilde{\mathcal{C}}$ in the sense that each object of $\widetilde{\mathcal{C}}$ is covered by some family of objects of the form εA. Hence we have covering families $\{\varepsilon A_i \xrightarrow{f_i} X : i \in I\}$, $\{\varepsilon B_j, g_j, Y : j \in J\}$ in $\widetilde{\mathcal{C}}$.

Let $A_{ij} = \varepsilon A_i \times_C \varepsilon B_j$

$$\begin{tikzcd}
\varepsilon A_i \arrow[r] & X \arrow[r] & C \\
 & & Y \arrow[u] \\
A_{ij} \arrow[uu] \arrow[rr, "\text{p.b.}"] & & \varepsilon B_j. \arrow[u]
\end{tikzcd}$$

Now assume that $\widetilde{M}(X) \leq \widetilde{M}(Y)$ in the tops $\mathcal{E}$. Fix now an arbitrary $i \in I$. $\widetilde{M}(\varepsilon A_i)$ is covered in $\mathcal{E}$ by the family $\{\widetilde{M}(A_{ij}) \to \widetilde{M}(\varepsilon A_i) : j \in J\}$, by the assumption $\widetilde{M}(X) \leq \widetilde{M}(Y)$. Using 6.1.2, let us cover in $\widetilde{\mathcal{C}}$ each A_{ij} by a family $\{\varepsilon A^j_k \to A_{ij} : k \in K_j\}$ with some $A^j_k \in \mathrm{Ob}(\mathcal{C})$ such that $\varepsilon A^j_k \to A_{ij} \to \varepsilon A_i = \varepsilon(f^j_k)$ for some $A^j_k \xrightarrow{f^j_k} A_i$ in $\mathcal{C}$. Now, $\{\widetilde{M}(\varepsilon A^j_k) \to \widetilde{M}(A_{ij}) : k \in K_j\}$ is a covering family by the continuity of $\widetilde{M}$. By composition, $\{\widetilde{M}(\varepsilon A^j_k) \xrightarrow{\widetilde{M}(\varepsilon f^j_k)} \widetilde{M}(\varepsilon A_i) : j \in J, k \in K_j\}$, i.e. $\{M(A^j_k) \xrightarrow{M(f^j_k)} M(A_i) : j \in J, k \in K_j\}$ is a covering family in $\mathcal{E}$. Hence, by assumption $\{\varepsilon A^j_k \xrightarrow{\varepsilon f^j_k} \varepsilon A_i : j \in J, k \in K_j\}$

is a covering family in $\widetilde{\mathcal{C}}$. It follows that $\{A_{ij} \to \varepsilon A_i : j \in J\}$ is a covering family. Since this is true for each $i \in I$, it follows that $X \leq Y$ as required. $\square$

Finally, we will discuss how infs of families of subobjects of an arbitrary object of the topos $\widetilde{\mathcal{C}}$ can be controlled by those in $\mathcal{R}$, and similarly for $\forall_f(C)$. The main tool for deriving our formulas will be the fact that $\widetilde{\mathcal{C}}$ is topologically generated by $\mathcal{C}$, i.e., for any $B \in \mathrm{Ob}(\widetilde{\mathcal{C}})$ there is a *covering* family

$$\{A_i \xrightarrow{\alpha_i} B : i \in I\} \in \mathrm{Cov}(\widetilde{\mathcal{C}}) \tag{1}$$

such that $A_i = \varepsilon(A_i^0)$ for some $A_i^0 \in \mathrm{Ob}(\mathcal{C})$ (c.f. 1.3.7).

Let $(X_j \hookrightarrow B)_j$ be a family of subobjects of the object B in $\widetilde{\mathcal{C}}$. Below, $\bigwedge^{(C)}$ and $\bigvee^{(C)}$ denote inf and sup, respectively, in the lattice of subobjects of C. Let (1) be a covering family for B.

We will use the notation α^{-1} to denote pullback; e.g.

$$\begin{tikzcd} X_j \arrow[r, hook] & B \\ \alpha_i^{-1}(X_j) \arrow[u] \arrow[r, hook] \arrow[ur, phantom, "\text{p.b.}"] & A_i \arrow[u, "\alpha_i"'] \end{tikzcd}$$

Recall the following simple facts.

(i) $\alpha^{-1}\bigwedge^{(B)}_{j\in J}X_j = \bigwedge^{(A_i)}_{j\in J}\alpha_i^{-1}(X_j)$ (this is trivial to check),

(ii) a covering family when pulled back results in another covering family, hence $\bigwedge_{j\in J}X_j = \bigvee^{(B)}_{i\in I}\exists_{\alpha_i}(\alpha_i^{-1}\bigwedge_{j\in J}X_j)$.

Using now (i) too, we obtain the formula

$$\bigwedge_{j\in J}X_j = \bigvee^{(B)}_{i\in I}\exists_{\alpha_i}\bigwedge_{j\in J}\alpha_i^{-1}(X_j). \tag{2}$$

From this, it is easy to show the following

Lemma 6.1.5 (i) *Suppose that $\mathcal{E}$ is a Grothendieck topos and that $\widetilde{\mathcal{C}} \xrightarrow{M} \mathcal{E}$ is an $\mathcal{E}$-model of $\widetilde{\mathcal{C}}$ that preserves all infs on the level of $\mathcal{R}$, i.e. if*

$$X_j \hookrightarrow \varepsilon(A); \qquad j \in J$$

is a family of subobjects, then

$$\bigwedge^{(M(\varepsilon(A)))}_{j\in J}M(X_j) = M(\bigwedge^{\varepsilon(A)}_{j\in J}X_j).$$

Then $\widetilde{M}$ preserves all infs in $\widetilde{\mathcal{C}}$.

(ii) *Suppose that M preserves all* stably distriburive *(c.f. Chapter 3, Section 2) infs on the level of $\mathcal{R}$. Then it preserves all stably distributive infs in $\widetilde{\mathcal{C}}$.*

Proof. (AD (i)). Let $X_j \hookrightarrow B$, $j \in J$ be any family of subobjects in $\widetilde{\mathcal{C}}$. Let (1) be a covering family with $A_i = \varepsilon(A_i^0)$. We have

$$\bigwedge^{M(\varepsilon(A))}_{j\in J}M(X_j) = \bigvee^{(M(B))}_{i\in I}\exists_{M\alpha_i}\bigwedge^{(M(A_i))}_{j\in J}(M\alpha_i)^{-1}MX_j$$

by using the above formula (2) used in $\mathcal{E}$,

$$= M(\bigvee^{(B)}_{i\in I}\exists_{\alpha_i}\bigwedge^{(A_i)}_{j\in J}\alpha_i^{-1}X_j)$$

by using that M is a *model* and it preserves the infs $\bigwedge_{j\in J}^{(A_i)}\alpha_i^{-1}(X_j)$, $A_i = \varepsilon(A_i^0)$;

$$= M(\textstyle\bigwedge_{j\in J}^{B} X_j)$$

by our formula (2), now used in $\widetilde{\mathcal{C}}$.

(AD (ii)). Let X_j, etc., be as in (i) and assume that the inf $\bigwedge_{j\in J}^{(B)} X_j$ is stably distributive. It follows directly that the infs $\bigwedge_{j\in J}^{(A_i)}\alpha_i^{-1}(X_j)$ are also stably distributive. Hence the argument in (i) applies. $\square$

Next we turn to $\forall$. Let $\mathcal{E}$ be a Grothendieck topos, $A \xrightarrow{f} B$ a morphism in $\mathcal{E}$, $X \to A$ a subobject and let us consider $\forall_f(X) \to B$. First assume that $\{A_i \xrightarrow{\alpha_i} A : i \in I\}$ is a covering family. We claim that

$$\forall_f(X) = \textstyle\bigwedge_{i\in I}^{(B)} \forall_{f\alpha_i}(\alpha_i^{-1}(X)) \tag{6.1.6}$$

$$\begin{tikzcd}
A_i \arrow[r, "\alpha_i"] & A \arrow[r, "f"] & B \\
\alpha_i^{-1}(X) \arrow[u, hook] \arrow[r] & X \arrow[r] & \forall_f(X) \arrow[u, hook]
\end{tikzcd}$$

Recall that, by definition, we have $f^{-1}(\forall_f(X)) \le X$ and for any subobject $Y \hookrightarrow B$, $f^{-1}(Y) \le X$ implies $Y \le \forall_f(X)$.

Let $Y = \forall_f(X)$. We have

$$(f\alpha_i)^{-1}(Y) = \alpha_i^{-1}(f^{-1}(\forall_f(X))) \le \alpha_i^{-1}(X),$$

hence $\forall_f(X) = Y \le \forall_{f\alpha_i}(\alpha_i^{-1}(X))$, thus $\forall_f X \le \bigwedge_{i\in I}\forall_{f\alpha_i}(\alpha_i^{-1}(X))$. For the converse, let $Y \underset{\text{df}}{=} \bigwedge_{i\in I}\forall_{f\alpha_i}(\alpha_i^{-1}(X))$ and $X' \underset{\text{df}}{=} f^{-1}(Y)$. Then $\alpha_i^{-1}(X') \le (f\alpha_i)^{-1}(Y) \le (f\alpha_i)^{-1}(\forall_{f\alpha_i}(\alpha_i^{-1}(X)) = \alpha_i^{-1}(X)$, hence $X' = \bigvee_{i\in I}\exists_{\alpha_i}(\alpha_i^{-1}(X')) \le \bigvee_{i\in I}\exists_{\alpha_i}(\alpha_i^{-1}(X))$ $= X$; here we used the fact that $\{\alpha_i^{-1}(X') \to X' : i \in I\}$, $\{\alpha_i^{-1}(X) \to X : i \in I\}$ are covering families. From $f^{-1}(Y) = X' \le X$ it follows that $Y \le \forall_f(X)$, as required.

Secondly, we need another formula related to covering now B in the situation

$$\begin{tikzcd}
A \arrow[r, "f"] & B \\
X \arrow[u, hook] \arrow[r] & \forall_f(X) \arrow[u, hook]
\end{tikzcd}$$

i.e., we have a covering family $\{B_j \xrightarrow{\beta_j} B : j \in J\}$

$$\begin{tikzcd}
& A_j \arrow[rr, "f_j"] \arrow[dd] \arrow[rrd, phantom, "\text{p.b.}"] & & B_j \arrow[dd, "\beta_j"] \\
h_j^{-1}(X) = X_j \arrow[rr, "h_j"] \arrow[dd] \arrow[rd, phantom, "\text{p.b.}"] & & \forall_{f_j}(X_j) \arrow[ur] \arrow[ddd, bend right] & \\
& A \arrow[rr, "f"] & & B \\
X \arrow[rr] & & \forall_f(X) \arrow[ur] & \\
& & & \exists_{\beta_j}(\forall_{f_j}(X_j)) \arrow[uu]
\end{tikzcd}$$

With the notation of the diagram we claim

$$\forall_f(X) = \bigvee_{j\in J} \exists_{\beta_j}(\forall_{f_j}(X_j)) \tag{6.1.7}$$

Using the pull-back diagram

$$\begin{tikzcd} A_j \arrow[r, "f_j"] \arrow[d, "h_j"'] & B_j \arrow[d, "\beta_i"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

we have $f_j^{-1}\beta_j^{-1}(\forall_f(X)) = h_j^{-1}f^{-1}(\forall_f(X)) \le h_j^{-1}(X) = X_j$. Hence $\beta_j^{-1}(\forall_f(X)) \le \forall_{f_j}(X_j)$. Thus

$$\forall_f(X) = \bigvee_{j\in J}\exists_{\beta_j}\beta_j^{-1}(\forall_f(X)) \le \bigvee_{j\in J}\exists_{\beta_j}(\forall_{f_j}(X_j)),$$

showing one of the two required inequalities. For the other one, let $Y_j = \forall_{f_j}(X_j)$ and start with the equality $f^{-1}(\exists_{\beta_j}(Y_j)) = \exists_{h_j}(f_j^{-1}(Y_j))$ obtained from the pullback diagram exhibited before. It follows that

$$f^{-1}(\exists_{\beta_j}(Y_j)) = \exists_{h_j}(f_j^{-1}(\forall_{f_j}(X_j))) \le \exists_{h_j}(X_j) \le X$$

and hence $\exists_{\beta_j}(Y_j) \le \forall_f(X)$. We obtain $\bigvee_{j\in J}\exists_{\beta_j}(\forall_{f_j}(X_j)) \le \forall_f(X)$ as required.

Now we are ready to prove

Lemma 6.1.8 (i) *Suppose $\mathcal{E}$ is a Grothendieck topos, $\widetilde{\mathcal{C}} \xrightarrow{M} \mathcal{E}$ is an $\mathcal{E}$-model of $\widetilde{\mathcal{C}}$ such that* $(1)_{\text{(i)}}$ *for any $A \in \mathrm{Ob}(\mathcal{C})$ and any family $X_j \hookrightarrow \varepsilon(A)$, $j \in J$ of subobjects, M preserves $\bigwedge^{\varepsilon A}_{j\in J} X_j$, i.e., $M(\bigwedge^{\varepsilon A}_{j\in J} X_j) = \bigwedge^{(M(\varepsilon A))}_{j\in J} MX_j$ and* $(2)_{\text{(i)}}$ *for any $A \in \mathrm{Ob}(\mathcal{C})$ $B \in \mathrm{Ob}(\mathcal{C})$, $A \xrightarrow{f} B$ in $\mathcal{C}$ and $X \hookrightarrow \varepsilon A$ in $\widetilde{\mathcal{C}}$, we have that M preserves $\forall_{\varepsilon f}(X)$ i.e. $M(\forall_{\varepsilon f}(X)) = \forall_{M(\varepsilon f)} M(X_j)$. Then M preserves $\forall_f(X)$ for arbitrary $X \hookrightarrow A \xrightarrow{f} B$, all in $\widetilde{\mathcal{C}}$.*

(ii) *Modify the hypothesis of* (i) *so that in* $(1)_{\text{(i)}}$ *and* $(2)_{\text{(i)}}$ *only the stably distributive $\bigwedge$ and $\forall$ are considered. Then the conclusion is that M preserves all stably distributive $\forall$ in $\widetilde{\mathcal{C}}$.*

Proof. (AD (i)). First, consider the special case

$$X \hookrightarrow A \xrightarrow{f} \varepsilon B.$$

By 6.1.2, let $\{\varepsilon A_i^0 \xrightarrow{\alpha_i} A : i \in I\}$ be a covering family (in $\widetilde{\mathcal{C}}$) such that for $i \in I$, $f\alpha_i = \varepsilon f_i$ for some $A_i^0 \xrightarrow{f_i} B$ in $\mathcal{C}$; let $A_i = \varepsilon A_i^0$. Now we have

$$\begin{aligned} \forall_{Mf}(MX) &= \textstyle\bigwedge^{(M\varepsilon B)}_{i\in I} \forall_{M(f\alpha_i)}((M\alpha_i)^{-1}(X)) && \text{(by 6.1.6)} \\ &= \textstyle\bigwedge^{(M\varepsilon B)}_{i\in I} \forall_{M(\varepsilon f_i)}(M(\alpha_i^{-1}(X)) \\ &= \textstyle\bigwedge^{(M\varepsilon B)}_{i\in I} M(\forall_{\varepsilon f_i}(\alpha_i^{-1}(X))) && \text{(by } (2)_{\text{(i)}}) \\ &= M\textstyle\bigwedge^{(\varepsilon B)}_{i\in I} \forall_{\varepsilon f_i}(\alpha_i^{-1}(X)) && \text{(by } (1)_{\text{(i)}}) \\ &= M(\forall_f(X)) && \text{(by 6.1.6 again).} \end{aligned}$$

Next, consider the general case

$$X \hookrightarrow A \xrightarrow{f} B.$$

There is a covering family $\{\varepsilon B_j^0 \xrightarrow{\beta_j} B : j \in J\}$, with $B_j^0 \in \mathrm{Ob}(\mathcal{C})$; let $B_j = \varepsilon B_j^0$. Now we use (6.1.7). The diagram exhibited before the formula (6.1.7) now takes place in $\widetilde{\mathcal{C}}$. M transforms this into a diagram in $\mathcal{E}$, the two pull-backs still remain pull-backs in $\mathcal{E}$, and the covering of B will transform into a covering of MB. Therefore we can apply (6.1.7) in $\mathcal{E}$ to obtain

$$\forall_{Mf}(MX) = \bigvee\nolimits_{j\in J} \exists_{M\beta_j}(\forall_{Mf_j}(MX_j)).$$

This further equals

$$= \bigvee\nolimits_{j\in J} \exists_{M\beta_j}(M(\forall_{f_j}(X_j))$$

by the special case discussed above (since now $B_j = \varepsilon B_j^0$ comes from $\mathcal{C}$), and

$$\cdots = M(\bigvee\nolimits_{j\in J} \exists_{\beta_j} \forall_{f_j}(X_j))$$

because M is a model,

$$\cdots = M(\forall_f(X))$$

by (6.1.7) again.

(AD (ii)). The argument is roughly the same except that we have to keep track of distributivity. With the notation of the first part of the proof for (i) and the hypothesis that $\forall_f(X)$ is stably distributive, first of all, we have to show that the inf $\bigwedge_{i\in I} \forall_{f\alpha_i}(\alpha_i^{-1}X)$ is stably distributive (recall that $\varepsilon f_i = f\alpha_i$). The proof is straightforward but somewhat messy. Let $C_i = \forall_{f\alpha_i}(\alpha_i^{-1}X)$. Let $B' \xrightarrow{f} B$ be an arbitrary morphism in $\widetilde{\mathcal{C}}$ and $D \hookrightarrow B'$ a subobject of B' in $\widetilde{\mathcal{C}}$.

$$\begin{tikzcd}
 & & A_i \arrow[d, "\alpha_i"] & \alpha_i^{-1}(X) \arrow[l] \arrow[d] & \\
 & & A \arrow[ddl, "f"] & X \arrow[l] & \\
 & & & A_i' \arrow[d, "\alpha_i'"] & (\alpha_i)^{-1}(X') \arrow[l] \arrow[dd] \\
C_i \arrow[r, hook] & B & & & \\
g^{-1}(C_i) \arrow[u] \arrow[r, hook] & B' \arrow[u, "g"] & \text{p.b.} & A' \arrow[ll, "f'"] \arrow[uuul] & X' \arrow[l]
\end{tikzcd}$$

Consider the pullback A' as shown and the pullbacks A_i', etc., along $A' \to A$. What we have to show is $(\bigwedge_i g^{-1}(C_i)) \vee D = \bigwedge_i (g^{-1}(C_i) \vee D)$. We have $\bigwedge_i g^{-1}(C_i) = g^{-1} \bigwedge_i C_i = g^{-1}\forall_f(X) = \forall_{f'}(X')$. Now, $\forall_{f'}(X') \vee D = \forall_{f'}(X' \vee (f')^{-1}D)$ by the *stable* distributivity of $\forall_f(X)$. We can use the formula 6.1.6 to conclude that $\forall_{f'}(X' \vee (f')^{-1}(D)) = \bigwedge_i \forall_{f'\alpha_i}((\alpha_i')^{-1}(X' \vee (f')^{-1}(D))) = \bigwedge_i \forall_{f'\alpha_i}((\alpha_i')^{-1}(X') \vee (f'\alpha_i')^{-1}(D))$ since $\{A_i \xrightarrow{\alpha_i} A_i : i \in I\}$ is a covering family. Using the stable distributivity of $\forall_f(X)$ for the inside of the last expression, we ten

$$\cdots = \bigwedge\nolimits_i (\forall_{f'\alpha_i}((\alpha_i')^{-1}(X')) \vee D).$$

But $g^{-1}(C_i) = \forall_{f'\alpha_i}((\alpha_i)^{-1}(X'))$ for trivial reasons, so we get

$$\cdots = \bigwedge\nolimits_i (g^{-1}(C_i) \vee D)$$

as required.

There are two $\forall$'s that have to be verified to be stably distributive to complete the proof of (ii), but they are immediate. $\square$

Finally, there is a straight forward variant of 6.1.8 for Heyting implications $A \to B$ and actually, for the generalized $\forall$'s of the form $\forall_f(A_1 \to A_2)$ c.f. Chapter 2, Section 2. We will refer to this variant as 6.1.8$'$.

§2 Categorical completeness theorems

Our first result is an improved version of Barr's theorem, Barr [1974]. Let $\mathcal{E} = \widetilde{\mathcal{C}}$ be a Grothendieck topos.

Theorem 6.2.1 *There is a complete Boolean algebra $\mathcal{B}$ and a conservative* $\mathrm{Sh}_{\mathcal{B}}$*-model of $\mathcal{E}$*

$$\widetilde{M} : \mathcal{E} \to \mathrm{Sh}_{\mathcal{B}}$$

that preserves, in addition, all stably distributive infs and $\forall$'s in $\mathcal{E}$. (Here $\mathrm{Sh}_{\mathcal{B}}$ *is the category of all sheaves over $\mathcal{B}$, with the canonical topology.)*

Remark *The improvement over Barr* [1974] *is the part "in addition ... ".*

Proof. Here by $\forall$'s we mean expressions of the form $\forall_f(A_1 \to A_2)$ generalizing Heyting '$\to$' and ordinary $\forall_f(A_2)$, c.f. Chapter 2, Section 2. We first set up a theory T that reflects the logical properties of the category $\mathcal{R} = \mathcal{R}_{\mathcal{C}}$ (c.f. previous section). T will be formulated over the language $L_{\mathcal{C}}$.

Among the axioms of T, we include all the axioms of $T_{\mathcal{C}}$ (c.f. previous section). Then, along with passing to a larger fragment F of $L_{\infty\omega}$, we consider the axioms for *stably distributive* infs and for $\forall$'s in $\mathcal{R}$, as follows. First of all, we include in F all disjunctions $\underset{\sim}{X}(a) = \bigvee_{i\in I}\underset{\sim}{X}_i(a)$ for any family $\langle X = X_i \xrightarrow{\xi_i} A : i \in I\rangle$ of A in $\mathcal{C}$, for an arbitrary $A \in \mathrm{Ob}(\mathcal{C})$. (Thus we have a notation for all objects in $\mathcal{R}$.)

For a subobject $X \xhookrightarrow{\xi} \varepsilon A$, we will write $\underset{\sim}{X}(a)$ for one of the disjunctions $\bigvee_{i\in I}\underset{\sim}{X}_i(a)$ such that $X \xhookrightarrow{\xi} A = \varepsilon_a(\bigvee_{i\in I}\underset{\sim}{X}_i(a))$; here we mean interpretations by the interpretation $\varepsilon : L_{\mathcal{C}} \to \widetilde{\mathcal{C}}$ as in 6.1.3. By 6.1.3, for any $X \xhookrightarrow{\xi} \varepsilon A$, there is at least one such disjunction $\underset{\sim}{X}(a)$.

Let $X = \langle X^{(j)} : j \in J\rangle$ by an arbitrary family of subobjects in $\widetilde{\mathcal{C}}$ of εA, $A \in \mathrm{Ob}(\mathcal{C})$ such that the inf $\bigwedge_{j\in J} X^{(j)}$ is *stably distributive.* For any such family X, we take $Y = \bigwedge^{(\widetilde{\mathcal{C}})}_{j\in J}\underset{\sim}{X}^{(j)} \xhookrightarrow{\xi} A$ and put the following axioms into T:

$$\underset{\sim}{Y}(a) \Rightarrow \bigwedge_{j\in J}\underset{\sim}{X}^{(j)}(a)$$

$$\bigwedge_{j\in J}\underset{\sim}{X}^{(j)}(a) \Rightarrow \underset{\sim}{Y}(a)$$

(c.f. item 10 before 2.4.5; also recall that $\underset{\sim}{X}^j = \bigvee_{i\in I}\underset{\sim}{X_i}^{(j)}(a)$, with$\underset{\sim}{X_i}^{(j)} := \exists x(\xi_i^j(x) \approx a)$, $X_i^{(j)} \xhookrightarrow{\xi_i^j} A$ in $\mathcal{C}$.) we proceed similarly for certain $\forall$'s. Let $A \xrightarrow{f} B$ be a morphism in $\mathcal{C}$, $X \hookrightarrow \varepsilon A$, $Y \hookrightarrow \varepsilon A$ arbitrary subobjects in $\widetilde{\mathcal{C}}$, $X = \varepsilon_a(\underset{\sim}{X}(a)) = \varepsilon_a(\bigvee_{i\in I}\underset{\sim}{X}_i(a))$, and similarly for Y. Let $\forall_f(X \to Y)$ be $Z \hookrightarrow B$, $Z = \varepsilon_b(\underset{\sim}{Z}(b))$ and assume that $\forall_f(X \to Y)$ is stably distributive. We add to T the axioms $\underset{\sim}{Z}(b) \Leftrightarrow \forall a((fa \approx y \wedge \underset{\sim}{X}(a)) \to \underset{\sim}{Y}(a))$ for all such $\forall_f(X \to Y)$, c.f. item 11 before 2.4.5.

Finally, although they can be shown to be superfluous, for simplicity we add the axioms

$$\underset{\sim}{X}(a) \leq \underset{\sim}{Y}(a)$$

for subobjects X, Y of εA such that $X \leq Y$.

We take F to be the smallest fragment such that all sequents in T are sequents of F.

We claim the following properties of T.

(1) For arbitrary subobjects $X \overset{\xi}{\hookrightarrow} \varepsilon A$, $Y \overset{\eta}{\hookrightarrow} \varepsilon A$, in $\widetilde{\mathcal{C}}$, $X \leq Y$ in the ordering of subobjects of εA if and only if $T \vdash \underset{\sim}{X}(a) \Rightarrow \underset{\sim}{Y}(a)$.

(2) Let $M_0 : L_{\mathcal{C}} \to \mathcal{E}$ be an interpretation of the language $L_{\mathcal{C}}$ in some category $\mathcal{E}$ with finite $\varprojlim$ such that M_0 is a model of the full theory T. Then M_0, regarded as a functor $M : \mathcal{C} \to \mathcal{E}$, will be a *model* of $\mathcal{C}$, moreover, for the induced model $\widetilde{M} : \widetilde{\mathcal{C}} \to \mathcal{E}$, $\widetilde{M}$ will preserve all the stably distributive infs and $\forall$'s of $\widetilde{\mathcal{C}}$.

Proof of (1). The 'only if' part is trivial. Assume $T \vdash \underset{\sim}{X}(a) \Rightarrow \underset{\sim}{Y}(a)$. Notice that in the canonical interpretation $\varepsilon : L_{\mathcal{C}} \to \widetilde{\mathcal{C}}$, all axioms of T are *true*. For the part $T_{\mathcal{C}} \subset T$, this is so because ε is a *model* of the site $\mathcal{C}$. (c.f. preceding section). For the axioms related to infs and $\forall$'s, this is so since the axioms truly express the qualities concerned (c.f. First Main Fact 2.4.5). Now we will use the Soundness Theorem, 3.5.4. Notice that all formulas of the fragment F are stable with respect to ε; for conjunctions and universally quantified formulas this is true precisely because of the distributivity assumptions. By soundness, taken together with ε being a model of T, this implies that if $T \vdash \underset{\sim}{X}(a) \Rightarrow \underset{\sim}{Y}(a)$ then $\varepsilon \models \underset{\sim}{X}(a) \Rightarrow \underset{\sim}{Y}(a)$, i.e., $X \leq Y$, as required.

Proof of (2). Since $T_{\mathcal{C}} \subset T$, by 6.1.1 we have that $M : \mathcal{C} \to \mathcal{E}$ is a model of $\mathcal{C}$. All the special stability distributive infs and $\forall$'s that are mentioned in the hypotheses of Lemmas 6.1.5(ii) and 6.1.8(ii) and 6.1.8$'$ are taken care of directly by the axioms of T, i.e., they are preserved by 2.4.5. Then by the mentioned lemmas, the assertion follows.

Now we are ready to use the Completeness Theorem 5.1.2. Hence there is a complete Boolean algebra $\mathcal{B}$ and a $\mathcal{B}$-valued model M_0 such that for any sequent σ of the fragment F, $M_0 \models \sigma$ iff $T \vdash \sigma$. By Chapter 4, M_0 can be regarded as a categorical interpretation $M_0 : L_{\mathcal{C}} \to \mathrm{Sh}_{\mathcal{B}}$ and $M_0 \models \sigma$ in the $\mathcal{B}$-valued sense is equivalent to $M_0 \models \sigma$ in the categorical sense. By (2) above, M_0 gives rise to a $\mathrm{Sh}_{\mathcal{B}}$-model $\widetilde{M}$, $\widetilde{M} : \widetilde{\mathcal{C}} \to \mathrm{Sh}_{\mathcal{B}}$, that preserves all stably distributive fins and $\forall$'s in $\widetilde{\mathcal{C}}$. Finally, let us show that $\widetilde{M}$ is conservative. By Lemma 6.1.4, it is enough to show that if X, Y are subobjects of εA in $\widetilde{\mathcal{C}}$, $A \in \mathrm{Ob}(\mathcal{C})$, then $\widetilde{M}(X) \leq \widetilde{M}(Y)$ implies $X \leq Y$. But $\widetilde{M}(X) = M_0(\underset{\sim}{X}(a))$, $\widetilde{M}(Y) = M_0(\underset{\sim}{Y}(a))$ and $\widetilde{M}(X) \leq \widetilde{M}(Y)$ means that $M_0 \models \underset{\sim}{X}(a) \Rightarrow \underset{\sim}{Y}(a)$. Hence $T \vdash \underset{\sim}{X}(a) \Rightarrow \underset{\sim}{Y}(a)$. By (1) above, this implies that $X \leq Y$ as required. □

Remark *Barr's original theorem is 6.2.1 without the condition on infs and $\forall$'s. It is easy to see that if $\widetilde{M} : \mathcal{E} \to \mathrm{Sh}_{\mathcal{B}}$ is a conservative model preserving a specific inf, or $\forall$, in $\mathcal{E}$, then that in, or $\forall$, mud be stably distributive; hence in a sense 6.2.1 is optimal.*

There is a simplified version of Barr's theorem that talks about an arbitrary small category $\mathcal{R}$ with finite $\varprojlim$ instead of a topos $\mathcal{E}$, and about preserving those logical operations that can be carried out in $\mathcal{R}$. The proof of this result is just a simplification of 6.1.2.

Theorem 6.2.1$'$ *Let $\mathcal{R}$ be an arbitrary small category with finite $\varprojlim$. There is a complete Boolean algebra $\mathcal{B}$ and a functor $F : \mathcal{R} \to \mathrm{Sh}_{\mathcal{B}}$ such that* (i) *F is conservative, i.e. $A \leq B$ in $\mathcal{R}$ iff $F(A) \leq F(B)$ in $\mathrm{Sh}_{\mathcal{B}}$, and* (ii) *$F$ preserves all finite left limits, all* *(finte or infinite) stable sups, all stable images, (finite or infinite) stably distributive infs and stably distributive $\forall_f$'s that exist in $\mathcal{R}$.*

The next result we want to discuss is a theorem of Deligne, SGA4, Vol. 2, Exposé VI, Proposition 9.0, p. 336. We will call the site $\mathcal{C}$ *algebraic* if $\mathcal{C}$ has finite $\varprojlim$ (as always for us) and moreover, all the basic covering families $\{A_i \xrightarrow{f_i} A : i \in I\}$ in $\mathcal{G}_0(\mathcal{C})$ are finite, i.e., I is finite. A *coherent topos* is a category equivalent to some $\widetilde{\mathcal{C}}$, with an algebraic site $\mathcal{C}$ (c.f. loc. cit.).

Theorem 6.2.2 (Deligne loc. cit.) *"Every coherent topos has enough points". If $\mathcal{E}$ is a coherent topos, then there is a conservative model $\widetilde{M} : \mathcal{E} \to \mathbf{Set}^I$ into a (Boolean) topos of the form $\mathbf{Set}^I$, with I a* set. *Equivalently, there is a small family $\langle \widetilde{M}_i : i \in I \rangle$ of $\mathbf{Set}$-models of $\mathcal{E}$, $M_i : \mathcal{E} \to \mathbf{Set}$, such that for $X \hookrightarrow A$, $Y \hookrightarrow A$ in $\mathcal{E}$, $X \leq Y$ iff for all $i \in I$, $\widetilde{M}_i(X) \leq \widetilde{M}_i(Y)$.*

Proof. Let $\{A_i \xrightarrow{\alpha_i} A : i \in I\}$ be a family of morphisms in $\mathcal{C}$ such that $\{\varepsilon A_i \xrightarrow{\varepsilon\alpha_i} \varepsilon A : i \in I\}$ is *not* a covering family in $\widetilde{\mathcal{C}}$, i.e., the subobject $\bigvee_{i \in I} \exists_{\varepsilon\alpha_i} \varepsilon A_i$ of A is strictly less that εA. Consider the formulas $\underset{\sim}{A}_i(a) = \exists a_i(\alpha_i(a) \approx a)$. Let $T_{\mathcal{C}}$ be the theory associated with the site $\mathcal{C}$; not $T_{\mathcal{C}}$ is a finitary theory (every formula in $T_{\mathcal{C}}$ is finitary). Let the fragment F be the finitary coherent logic over the language $L_{\mathcal{C}}$. Let I' be a finite subset of I. We claim that $T_{\mathcal{C}} \not\vdash a \approx a \Rightarrow \bigvee_{i \in I'} \underset{\sim}{A}_i(a)$. Indeed, by the Soundness Theorem, the interpretation $\varepsilon : L_{\mathcal{C}} \to \widetilde{\mathcal{C}}$ would otherwise satisfy the sequent $a \approx a \Rightarrow \bigvee_{i \in I'} \underset{\sim}{A}_i(a)$ (since it satisfies $T_{\mathcal{C}}$), hence we would have $\varepsilon A = \bigvee_{i \in I'} \exists_{\varepsilon\alpha_i} A_i$ and a fortiori $\varepsilon A = \bigvee_{i \in I'} \exists_{\varepsilon\alpha_i} \varepsilon A_i$, that is not the case by our hypothesis; the claim is established.

Now, we can apply the two-valued Completeness Theorem 5.1.7(i). Hence, for any finite $I' \subset I$, there is an ordinary set-model of $T_{\mathcal{C}}$ that is not a model of $a \approx a \Rightarrow \bigvee_{i \in I'} \underset{\sim}{A}_i(a)$. Let c be a new individual constant not in $L_{\mathcal{C}}$ and consider the theory

$$T = T_{\mathcal{C}} \cup \{\underset{\sim}{A}_i(c) \Rightarrow \downarrow : i \in I\}$$

in an enlarged but still finitary logic. For any finite subset I' of I, the subset $T_{\mathcal{C}} \cup \{\underset{\sim}{A}_i(c) \Rightarrow \downarrow : i \in I'\}$ has a model N, since clearly, any model M of $T_{\mathcal{C}}$ that is not a model of $a \approx a \Rightarrow \bigvee_{i \in I'} \underset{\sim}{A}_i(a)$ can be made into such a model, by taking c^N to be an element of $M(A)$ that fails to satisfy the sequent $a \approx a \Rightarrow \bigvee_{i \in I'} \underset{\sim}{A}_i(a)$.

Using the compactness theorem for finitary logic (5.1.8), we conclude that the theory T has a model $M^* = (M, c^{M^*})$. Clearly, M is a model of $T_{\mathcal{C}}$. By the basic property of $T_{\mathcal{C}}$, 6.1.1, we have that $M : \mathcal{C} \to \mathbf{Set}$ is a model of $\mathcal{C}$, hence it gives rise to a model $\widetilde{M} : \widetilde{\mathcal{C}} \to \mathbf{Set}$ of M. We have

$$\widetilde{M}(\varepsilon A_i) = M(A_i) = \{x \in M(A) : M \models \underset{\sim}{A}_i[x \text{ for } a]\} \subset M(A)$$

and

$$\widetilde{M}(\textstyle\bigvee_{i \in I} \varepsilon A_i) = \bigcup_{i \in I} \widetilde{M}(\varepsilon A_i) \subset M(A).$$

Since for $x_0 = c^{M^*}$, $M \models \neg \underset{\sim}{A}_i[x_0 \text{ for } a]$ for all $i \in I$, we have that $x_0 \in M(A) - \widetilde{M}(\bigcup_{i \in I} \varepsilon A_i)$, hence

$$\widetilde{M}(\textstyle\bigvee_{i \in I} \varepsilon A_i) \neq \widetilde{M}(\varepsilon A) = M(A).$$

In conclusion, for any family $j = \{A_i \xrightarrow{\alpha_i} A : i \in I\}$ such that $\{\varepsilon A_i \xrightarrow{\varepsilon\alpha_i} \varepsilon A : i \in I\}$ is *not* a covering family in $\widetilde{\mathcal{C}}$, there is a model $\widetilde{M}_j : \widetilde{\mathcal{C}} \to \mathbf{Set}$ such that $\{M_j(A_i) \xrightarrow{M_j(\alpha_i)}$

$M_j(A) : i \in I\}$ is not a covering family in **Set** either. Let J be the set of all such j and let us define $M : \mathcal{C} \to \mathbf{Set}^J$ by $M = \langle M_j \rangle_{j\in J}$. It follows that M is a model of $\widetilde{\mathcal{C}}$ and also, that it satisfies the hypothesis of Lemma 6.1.4. Hence by 6.1.4, $\widetilde{M}$ is conservative. □

In the next result we conclude that certain other Grothendieck topoi also have enough points. The proof will be quite similar to the above proof of Deligne's theorem and in fact, it will be a bit simpler. It will exploit that we have a completeness theorem with set-valued models for countable fragments of $L_{\infty\omega}$ (c.f. 5.1.7(ii)). As far as we know, this result is new.

We will call a site $\mathcal{C}$ *separable* if (i) the underlying category $\mathcal{C}$ is countable, i.e., it has countably many objects and $\hom(A, B)$ is countable for each A, B in $\mathrm{Ob}(\mathcal{C})$, and (ii) the system of basic coverings $\mathcal{G}_0(\mathcal{C})$ is countable. Notice that if $\mathcal{C}$ is separable then the theory $T_\mathcal{C}$ is countable, in fact, it is a theory in a countable fragment F of $L^g_{\infty\omega}$. We will call a Grothendieck topos $\mathcal{E}$ *separable* if $\mathcal{E}$ is equivalent to $\widetilde{\mathcal{C}}$ for some separable site $\mathcal{C}$ with finite $\varprojlim$. We note without proof the following

Proposition 6.2.3 *For a Hausdorff topological space X, the category of sheaves* $\mathrm{Sh}(X)$ *over X is a separable topos if and only if X is homeomorphic to a complete separable metric space.*

Theorem 6.2.4 *A separable Grothendieck topos $\mathcal{E}$ has enough points, i.e., there is a conservative model $\widetilde{M} : \widetilde{\mathcal{C}} \to \mathbf{Set}^I$ with I a set.*

Proof. Let $j = \{A_i \xrightarrow{\alpha_i} A : i \in I\}$ be a family of morphisms in $\mathcal{C}$ such that $\{\varepsilon A_i \xrightarrow{\varepsilon\alpha_i} \varepsilon A : i \in I\}$ is *not* a covering family in $\widetilde{\mathcal{C}}$. Let F' be a countable fragment of $(L_\mathcal{C})^g_{\infty\omega}$ containing F and the formula $\bigvee_{i\in I} \underset{\sim}{A}_i(a)$, $\underset{\sim}{A}_i(a) = \exists\alpha_i(\alpha_i(a) = a)$; notice that I is countable. We claim that $T_\mathcal{C} \not\vdash \sigma_j$ where σ_j is $a \approx a \Rightarrow \bigvee_{i\in I} \underset{\sim}{A}_i(a)$. As before, this is a consequence of the soundness theorem, the fact that $\varepsilon : L_\mathcal{C} \to \mathcal{E} = \widetilde{\mathcal{C}}$ is an adequate interpretation of the full coherent logic $(L_\mathcal{C})^g_{\infty\omega}$, and our hypothesis on j. By 5.1.7(ii), there is a set-valued model M_j of $T_\mathcal{C}$ that is not a model of σ_j. Hence $M_j : \mathcal{C} \to \mathbf{Set}$ is a model of $\mathcal{C}$, $\widetilde{M}_j : \widetilde{\mathcal{C}} \to \mathbf{Set}$ is a model of $\widetilde{\mathcal{C}}$ and $\{M(A_i) \overset{M(\alpha_i)}{\longrightarrow} M(A) : i \in I\}$ is *not* a covering family in **Set**. Hence for $M = \langle M_j \rangle_{j\in J}$, with J the set of all j as considered above, $M : \widetilde{\mathcal{C}} \to \mathbf{Set}^J$ is a model that is conservative by Lemma 6.1.4. □

§3 Intuitionistic models

We will be interested in models that preserve all infs and all $\forall$'s in a given topos. Since the internal logic of topoi is intuitionistic, the results to be discussed are intimately related to the semantics of intuitionistic logic, c.f. Rasiowa-Sikorski [1963], Kripke [1963], Fitting [1969], [1973].

Theorem 6.3.1 *For any Grothendieck topos $\mathcal{E} = \widetilde{\mathcal{C}}$, there is a complete Heyting algebra $\mathcal{H}$ and a conservative model $\widetilde{M} : \widetilde{\mathcal{C}} \to \mathrm{Sh}_\mathcal{H}$ into the category of all sheaves over $\mathcal{H}$ such that M preserves, in addition, all infs and $\forall$'s in $\widetilde{\mathcal{C}}$.*

Proof. We will use the Boolean completeness theorem similarly as in the proof of 6.2.1 and we will extract a Heyting valued model from the Boolean valued model. We consider the theory $T_\mathcal{C}$, together wit the smallest fragment F containing all formulas in $T_\mathcal{C}$ and all infinite disjunctions $\bigvee_{i\in I} \underset{\sim}{A}_i(a)$, for families $\{A_i \xrightarrow{\alpha_i} A : i \in I\}$ of morphisms in $\mathcal{C}$; here $\underset{\sim}{A}_i(a) = \exists a_i\ \alpha_i(a_i) = a$. Let $\mathcal{B}$ be the Boolean algebra and M the $\mathcal{B}$-valued model constructed in the proof of 5.1.2 such that $M \models \sigma$ iff $T \vdash \sigma$, for any sequent

σ of F. We will use some properties of the model M constructed there beyond the actual statement of the completeness theorem. First of all, the elements of the partial domain $|M|_A$ corresponding to the sort A (= object A in $\mathcal{C}$) are the terms of $L_{\mathcal{C}}$ of sort A. Secondly, the variables as elements of the various partial domains $|M|_A$ are *generic* in the sense that if $\|\vec{x}\| \cdot \|\phi\| \leq \|\psi\|$, $\vec{x}$ is the sequence of all free variables in $\phi \Rightarrow \psi$ (recall that $\|\phi\|$ means the value of ϕ in M with the free variables of ϕ interpreted by themselves as elements of $|M|$), then in fact $T \vdash \phi \Rightarrow \psi$. In other words, a sequence $\vec{x}$ of variables as elements of $|M|$ have a property in M expressible in the form $\phi \Rightarrow \psi$ only if this property is shared by any tuple $\vec{a}$ of elements in $|M|$ with corresponding sorts.

To define the Heyting algebra $\mathcal{H}$, let us call an element of the Boolean algebra $\mathcal{B}$ *basic* if it is of the form $\|\phi\|_M$, with ϕ a formula of the fragment F. Let $\mathcal{H}$ be the set of elements of $\mathcal{B}$ which are obtained as the Boolean sup in $\mathcal{B}$ of any number of basic elements. $\mathcal{H}$ inherits a partial ordering from $\mathcal{B}$; we claim that $\mathcal{H}$ is a complete Heyting algebra with respect to this ordering. It is easy to check that, in fact, for any elements h_1, h_2, h_i of $\mathcal{H}$,

$$\begin{aligned} \textstyle\bigvee^{(\mathcal{H})}_{i\in I} h_i &= \textstyle\bigvee^{(\mathcal{B})}_{i\in I} h_i, \\ h_1 \wedge^{(\mathcal{H})} h_2 &= h_1 \wedge^{(\mathcal{B})} h_2, \\ \textstyle\bigwedge^{(\mathcal{H})}_{i\in I} h_i &= \textstyle\bigvee\{h : h \leq h_i \text{ for all } i \in I\}, \\ h_1 \to^{(\mathcal{H})} h_2 &= \textstyle\bigvee\{h : h \wedge h_1 \leq h_2\} \end{aligned}$$

and the laws of Heyting algebra are satisfied.

Next we define an $\mathcal{H}$-valued model M'. The domains of operations of the $\mathcal{H}$-valued structure M' are defined to be the same as those of M. By putting $\|\theta\|_{M'} = \|\theta\| = \|\theta\|_M$ for atomic formulas, we define equality on M' and also, the membership functions $\|a\|_{M'} = \|a = a\|_{M'} = \|a = a\|$. It is obvious that the basic conditions on these data are satisfied just because they are satisfied for M in the $\mathcal{B}$-valued sense. Next, since finite infs and arbitrary sups are computed in $\mathcal{H}$ just as in $\mathcal{B}$, and since evaluating formulas in $L^g_{\infty\omega}$ only these two operations are used, we clearly have that for every $\phi \in F$, $\|\phi\|_{M'} = \|\phi\|_M$. Hence, a sequent of $F \subset L_{\infty\omega}$ is satisfied in M' in the $\mathcal{H}$-valued sense just in case it is satisfied in M in the $\mathcal{B}$-valued sense. It follows that M' satisfies the theory $T_{\mathcal{C}}$, hence $M : \mathcal{C} \to \mathbf{Set}_{\mathcal{H}} = \mathrm{Sh}_{\mathcal{H}}$ is a model of $\mathcal{C}$. Since M' satisfies the same sequents in the fragment F as M and M is conservative, and a fortiori M satisfies the hypothesis of Lemma 6.1.4 (the condition after "iff"), it follows that $M' : \widetilde{\mathcal{C}} \to \mathrm{Sh}_{\mathcal{H}}$ is conservative.

We are left with having to verify the claims regarding $\forall$'s and infs. We will use 6.1.5(i) and 6.1.8(i) and thus, we have to verify the hypotheses of 6.1.5(i) and 6.1.8(i). Because of the similarity of the arguments, we leave it to the reader to verify the hypothesis of 6.1.5(i) which is identical to part $(1)_{\text{(i)}}$ of the hypothesis of 6.1.8(i) and we proceed to verify $(2)_{\text{(i)}}$ in 6.1.8(i). Let $A \xrightarrow{f} B$ be a morphism in $\mathcal{C}$ and $X \overset{\xi}{\hookrightarrow} \varepsilon A$ a subobject of εA in $\widetilde{\mathcal{C}}$. We know that $X \overset{\xi}{\hookrightarrow} \varepsilon A = \varepsilon(\underset{\sim}{X}(a))$ for some formula (infinite disjunction) $\underset{\sim}{X}(a)$. Similarly, for $Y = \forall_{\varepsilon f}(X \overset{\xi}{\hookrightarrow} \varepsilon A) \overset{\eta}{\hookrightarrow} \varepsilon B$, let $\underset{\sim}{Y}(b)$ be a formula of F such that $Y \overset{\eta}{\hookrightarrow} \varepsilon B = \varepsilon(\underset{\sim}{Y}(b))$. By item 11, 2.4.5, it is sufficient to show that M' as an $\mathcal{H}$-valued model satisfies the sequents

$$\underset{\sim}{Y}(b) \Rightarrow \forall a(fa \approx b \to \underset{\sim}{X}(a)),$$

$$\forall a(fa \approx b \to \underset{\sim}{X}(a)) \Rightarrow \underset{\sim}{Y}(b).$$

This is equivalent to saying that for any element $b = t = t(\vec{x}) \in |M'|_B$ (recall that elements of $|M|_B$ are terms!) we have

$$\|\underset{\sim}{Y}(t)\|_{M'} = \|\forall a(fa \approx t \Rightarrow \underset{\sim}{X}(a))\|_{M'}$$

with 'a' a variable not among the $\vec{x}$. By definition of $\mathcal{H}$-valued evaluation, the right-hand side equals to

$$(1) \qquad \bigvee\nolimits^{(\mathcal{B})}\{\|t\| \cdot \beta : \|fs = t\| \cdot \beta \leq \|\underset{\sim}{X}(s)\| \text{ for all } s \in |M|_A\}$$

where β ranges over basic elements of $\mathcal{B}$ (of $\mathcal{H}$). To show that the right-hand-side of the claimed equality is $\leq$ the left, assume that $\beta = \|\phi\|$ and that $\|fs = t\| \cdot \|\phi\| \leq \|\underset{\sim}{X}(s)\|$ for any $s \in |M|_A$. Then we can choose s to be a variable 'a' not occurring free either in t or in ϕ, and have

$$\|fa = t\| \cdot \|\phi\| \leq \|\underset{\sim}{X}(a)\|.$$

By the genericity of variables discussed above,

$$T_{\mathcal{C}} \vdash fa = t, \phi \Rightarrow \underset{\sim}{X}(a).$$

Since $\forall_{\varepsilon f}(X)$ exists, the rule

$$\frac{fa = t, \phi \Rightarrow \underset{\sim}{X}(a)}{\phi \Rightarrow \forall a[fa = t \to \underset{\sim}{X}(a)]}$$

is sound in the interpretation $\varepsilon : L_{\mathcal{C}} \to \widetilde{\mathcal{C}}$. Hence, since ε is a model of T, we have

$$\varepsilon_{\vec{y}}(\phi) \leq \varepsilon_{\vec{y}}(\forall a(fa = t \to \underset{\sim}{X}(a)));$$

here $\vec{y}$ is the sequence of variables occurring free in t or ϕ. Since $\varepsilon_b(\forall a(fa = b \to \underset{\sim}{X}(a))) = \forall_{\varepsilon f}(X) = Y$, and $\forall_{\varepsilon f}(X)$ is stable, we have by the Substitution Lemma 3.2.3 that $\varepsilon_{\vec{y}}(\forall a(fa = t \to \underset{\sim}{X}(a))) = \varepsilon_{\vec{y}}(\underset{\sim}{Y}(t))$. We have obtained that

$$\varepsilon_{\vec{y}}(\phi) \leq \varepsilon_{\vec{y}}(\underset{\sim}{Y}(t)).$$

Since the functor $\widetilde{M} : \widetilde{\mathcal{C}} \to \mathbf{Set}_{\mathcal{H}}$ preserves monomorphisms, we obtain that the subobject $\widetilde{M}(\varepsilon_{\vec{y}}(\phi))$ of $M(A) = \widetilde{M}(\varepsilon A)$ is $\leq$ the subobject $\widetilde{M}(\varepsilon_{\vec{y}}\underset{\sim}{Y}(t))$ of $M(A)$, all in $\mathbf{Set}_{\mathcal{H}}$. Translating this fact into the language talking about M' as an $\mathcal{H}$-valued model, this means that

$$\|\vec{y}\| \cdot \|\phi\| \leq \|\vec{y}\| \cdot \|\underset{\sim}{Y}(t)\|$$

or equivalently,

$$\|t\| \cdot \beta = \|t\| \cdot \|\phi\| \leq \|\underset{\sim}{Y}(t)\|.$$

This shows that each member of the sup in (1) is $\leq \|\underset{\sim}{Y}(t)\|$. This means that the sup itself is so too.

Conversely, taking $\beta = \|\underset{\sim}{Y}(t)\|$, we can show $\|fs = t\| \cdot \|\underset{\sim}{Y}(t)\| \leq \|\underset{\sim}{X}(s)\|$ for all $s \in |M|_A$; this establishes the other inequality. This finishes the verification of $(2)_{(\mathrm{i})}$ in 6.1.8(i). Note that, strictly speaking, we need to verify the more complicated 6.1.8$'$ (c.f. end of Section 1) to deal with generalized $\forall$. This completes what we had to say about the proof of 6.3.1. $\square$

In case there are enough point of $\mathcal{E}$, then there is a topological space T so that $\mathrm{Sh}_{\mathcal{H}}$ can be replaced by $\mathrm{Sh}(T)$, the category of sheaves over T, in 6.3.1. In fact, this space T is "the space of all points" of $\mathcal{E}$. Next we proceed to organize points into a space T.

Here we essentially redefine the space $\mathrm{esp}(\mathcal{E})$ introduced in Hakim [1972], p. 24. We do this in a differing terminology that is more suited to our prejudices. Let now $\mathcal{E} = \widetilde{\mathcal{C}}$ be an *arbitrary* Grothendieck topos. Let κ be the cardinality of the smallest fragment $F_{\mathcal{C}}$ such that $T_{\mathcal{C}}$ is a theory in $F_{\mathcal{C}}$. κ can be described as the maximum of κ_1, κ_2, κ_3 and $\aleph_0$ where κ_1 is the cardinality of the set of objects of $\mathcal{C}$, κ_2 is the maximum of cardinalities of the sets $\hom_{\mathcal{C}}(A, B)$, A, $B \in \mathrm{Ob}(\mathcal{C})$, and κ_3 is the cardinality of $\mathcal{G}_0(\mathcal{C})$. For instance, if $\mathcal{C}$ is a separable site, then $\kappa = \aleph_0$. Now we extend the language $L_{\mathcal{C}}$ by adding a set of new individual constants. For each $A \in \mathrm{Ob}(\mathcal{C})$, we take a set C_A of individual constants of cardinality κ. For $A \neq B$, $C_A \cap C_B = \varnothing$. we declare that each $c \in C_A$ is of sort A. For the language $L' \supset L_{\mathcal{C}}$ thus obtained, we have the set $\mathcal{T}_A$, the set of all closed (variable-free) terms of L' of sort A. Next we define our crucial notion, that of a C-structure. Roughly speaking, a C-structure is a structure of the similarity type $L_{\mathcal{C}}$ all of whose elements are denoted by some closed terms in L'. Precisely, a C-structure is given by (i) a set-valued structure M of the similarity type $L_{\mathcal{C}}$ (i.e. $M : L_{\mathcal{C}} \to \mathbf{Set}$), (ii) a family of functions F_A ($A \in \mathrm{Ob}(\mathcal{C})$) such that $\mathrm{dom}\, F_A \subset \mathcal{T}_A$, $\mathrm{rn}(F_A) = M(A)$ and if $t = fs$, t is of sort B, s is of sort A, and $s \in \mathrm{dom}\, F_A$, then we have $t \in \mathrm{dom}\, F_B$ and the compatibility relation $F_B(t) = f^M(F_A(s))$ holds.

In order that we have to deal with a (small) set of C-structures only, we require furthermore that the partial domains of the C-structures should be subsets of a fixed large enough set; in this way every C-structure without this restriction will be isomorphic to one with this restriction.

For a C-structure $(M, F_A)_{A \in \mathrm{Ob}(\mathcal{C})} = M$, the partial domains might be empty, but always, every element in them is *denoted* (via the F_A) by a term and we have that the set of terms that are actually used for denoting elements in M are closed under the operation symbols of $L_{\mathcal{C}}$.

For a formula $\phi(\vec{c})$ of the extended language $L'_{\infty\omega}$, $\vec{c} = (c_1, \ldots, c_n)$ being constants in $C = \bigcup A \in \mathrm{Ob}(\mathcal{C}) C_A$, $\mathrm{sort}(c_i) = A_i$, we can now define the set

$$\langle \phi(\vec{c}) \rangle$$

as the *set* of all C-structures $M = (M, F_A)_{A \in \mathrm{Ob}(\mathcal{C})}$ such that (i) $c_i \in \mathrm{dom}\, F_A$, and (ii) $M \models \phi[F_{A_1}(c_1), \ldots, F_{A_n}(c_n)]$ (i.e., the constants c_i are interpreted as $F_{A_i}(c_i)$ in the formula).

E.g. for the formula $c \approx c$, c of sort A, $\langle c \approx c \rangle$ is the set of all C-structures $(M, F_A)_{A \in \mathrm{Ob}(\mathcal{C})}$ such that $c \in \mathrm{dom}\, F_A$. $\langle c \approx c \rangle$ will be denoted by $\langle c \rangle$. For a sequence $\vec{c} = (c_1, \ldots, c_n)$ of constants, $\langle \vec{c} \rangle$ stands for $\langle c_1 \rangle \cap \langle c_1 \rangle \cap \cdots \cap \langle c_n \rangle$.

Let $F \subset F_{\mathcal{C}}$ be the smallest fragment of $(L_{\mathcal{C}})^g_{\infty\omega}$ such that all formulas in $T_{\mathcal{C}}$ are formulas of F and also, each formula $\bigvee_{i \in I} \underset{\sim}{A}_i(a)$ belongs to F, for any family $\{A_i \xrightarrow{\alpha_i} A : i \in I\}$ of morphisms in $\mathcal{C}$ and $\underset{\sim}{A}_i(a) = \exists a_i(\alpha_i(a_i) = a)$. Let F' be the fragment of $(L')^{\gamma}_{\infty\omega}$ obtained as the set of all substitution instances of formulas in F when we substitute constants in C for variables. We will be interested in *sentences* of F' only. Any sentence of F' has the form $\phi(\vec{c})$, with some $\phi(\vec{x})$ in F. It is easy to see that for a sequent $\phi(\vec{c}) \Rightarrow \psi(\vec{c})$ of F', $T_{\mathcal{C}} \vdash \phi(\vec{c}) \Rightarrow \psi(\vec{c})$ iff $T_{\mathcal{C}} \vdash \phi(\vec{x}) \Rightarrow \psi(\vec{x})$ provided that in $\vec{c} = (c_1, \ldots, c_n)$, $c_i \neq c_j$ for $i \neq j$. This is true basically because the axioms in $T_{\mathcal{C}}$ do not contain the new constant in $\mathcal{C}$.

Finally, the topological space T is defined as having the underlying set the set of all C-models and basic open sets all set of the form $\langle \phi(\vec{c}) \rangle$ with sentences $\phi(\vec{c})$ of the fragment F'.

For simplicity (though superfluously) we add all sequents $\phi(\vec{x}) \Rightarrow \psi(\vec{x})$ of F to $T_{\mathcal{C}}$ that are true in the interpretation $\varepsilon : L_{\mathcal{C}} \to \widetilde{\mathcal{C}}$; we denote the result by $T_{\mathcal{C}}$ too.

The basic fact that we need below is the following

Lemma 6.3.2 *The topos $\mathcal{E} = \widetilde{\mathcal{C}}$ has enough points in the sense of Theorems 6.2.2 and 6.2.4 if and only if the following condition* $(*)$ *holds:*

$(*)$ *For any sequent $\phi(\vec{c}) \Rightarrow \psi(\vec{c})$ of F', $T_{\mathcal{C}} \vdash \phi(\vec{c}) \Rightarrow \psi(\vec{c})$ iff $\langle \vec{c} \rangle \cap \langle \phi(\vec{c}) \rangle \subseteq \langle \psi(\vec{c}) \rangle$. (In case each constant in $\vec{c}$ actually occurs in both ϕ and ψ, the condition is $\langle \phi(\vec{c}) \rangle \subseteq \langle \psi(\vec{c}) \rangle$.)*

Proof. The 'if' part will not be needed below and its proof is left to the reader. Assume that $\widetilde{\mathcal{C}}$ has enough points. Suppose first that $T \vdash \phi(\vec{c}) \Rightarrow \psi(\vec{c})$. But then, since the rules of $\vdash$ are sound for the set-theoretic interpretation, it obviously follows that $\langle \vec{c} \rangle \cap \langle \phi(\vec{c}) \rangle \subseteq \langle \psi(\vec{c}) \rangle$. Suppose that $T \not\vdash \phi(\vec{c}) \Rightarrow \psi(\vec{c})$. Substituting distinct variables x_i for the c_i, we obtain that $T \vdash \phi(\vec{x}) \Rightarrow \psi(\vec{x})$, hence ε does not satisfy $\phi(\vec{x}) \Rightarrow \psi(\vec{x})$. In other words $\varepsilon_{\vec{x}}(\phi) \not\leq \varepsilon_{\vec{x}}(\psi)$, as subobjects of $\varepsilon(\vec{x}) = \varepsilon(A_i) \times \cdots \times \varepsilon(A_n)$, A_i the sort of x_i. Hence, since there are enough models of $\widetilde{\mathcal{C}}$, it follows that there is $M : L_{\mathcal{C}} \to \mathbf{Set}$ such that M is a model of $T_{\mathcal{C}}$ that is not a model of $\phi(\vec{x}) \Rightarrow \psi(\vec{x})$.

By the downward Lowenheim-Skolem theorem (c.f. [CK]) we can assume that each partial domain $|M|_A$ ($A \in \mathrm{Ob}(\mathcal{C})$) has cardinality at most κ. If $\vec{x} = \langle x_1, \ldots, x_n \rangle$, then there are elements $a_1, \ldots, a_n$ of $|M|$, $a_i \in |M|_{A_i}$, A_i the sort of x_i, such that $M \models \phi[\vec{a}]$ but $M \not\models \psi[\vec{a}]$. We turn M into a C-structure as follows.

We pick a subset $C'_A \subset C_A$ of cardinality at least that of $|M|_A$ and take an onto map $f_A : C'_A \to |M|_A$. (If $|M|_A$ is nonempty, we can take $C'_A = C_A$, but if $|M|_A$ is empty, we have to take $C'_A = \varnothing$.) We can always arrange the f_A such that $f_{A_i}(c_i) = a_i$, with c_i and a_i determined above. Then we can extend f_A in a unique way to a function F_A on a subset of $\mathcal{T}_A$ such that $(M, F_A)_{A \in \mathrm{Ob}(\mathcal{C})}$ will be a C-structure. By our arrangement, we now have that for M as a $\mathcal{C}$-structure, $c_i \in \mathrm{dom}\, F_{A_i}$ and $M \models \phi(\vec{c})$ but $M \models \neg\psi(\vec{c})$, in other words, $M \in \langle \vec{c} \rangle \cap \langle \phi(\vec{c}) \rangle - \langle \psi(\vec{c}) \rangle$, showing that $\langle \vec{c} \rangle \cap \langle \phi(\vec{c}) \rangle \not\subseteq \langle \psi(\vec{c}) \rangle$ as required. This proves 6.3.2.

Continuing with an arbitrary topos $\mathcal{E} = \widetilde{\mathcal{C}}$, we now set up a $\mathcal{H}$-valued model N of $\mathcal{E}$, where $\mathcal{H}$ is the complete Heyting algebra of all open subsets of the topological space T determined above. Let $|N|_A$ be $\mathcal{T}_A$, the set of all closed terms of L' of sort A and let the operations corresponding to the symbols in $L_{\mathcal{C}}$ be defined in the obvious way. By putting

$$\|t_1 \approx t_2\|^N = \langle t_1 \approx t_2 \rangle$$

we define the equality predicates on each $|N|_A$ and it is easy to check that this way we have defined a proper $\mathcal{H}$-valued structure of the similarity type $L_{\mathcal{C}}$. Next we notice the following easily verified facts:

$$\begin{aligned} \langle \phi_1 \wedge \phi_2 \rangle &= \langle \phi_1 \rangle \wedge \langle \phi_2 \rangle \\ \langle \textstyle\bigvee_{i \in I} \phi_i \rangle &= \textstyle\bigcup_{i \in I} \langle \phi_i \rangle \\ \langle \exists x \phi(x) \rangle &= \textstyle\bigcup_{t \in \mathcal{T}_A} \langle \phi(t \text{ for } x) \rangle. \end{aligned}$$

(A is the sort of the variable x.)

This shows that for any formula ϕ in F', for $\|\phi\|^N =$ the value of ϕ in N when each c_i in ϕ is evaluated as $c_i \in |N|_{A_i}$, we have

$$\|\phi\|^N = \langle \phi \rangle \tag{2}$$

Now assume that the topos $\mathcal{E} = \widetilde{\mathcal{C}}$ has enough points. It follows easily from (2) and 6.3.2 that $\widetilde{N} : \widetilde{\mathcal{C}} \to \mathrm{Sh}_{\mathcal{H}}$ for $N : \mathcal{C} \to \mathrm{Sh}_{\mathcal{H}}$ is a conservative model of $\widetilde{\mathcal{C}}$. Finally, we claim that N also preserves infs and $\forall$'s of $\widetilde{\mathcal{C}}$. The proof of this is exactly the same as the corresponding part in the proof of 6.3.1. In fact, now the constants play the same role as the free variables did in 6.3.1. Condition $(*)$ in 6.3.2 expresses the genericity of the constants in C. Using the formula (2) and condition $(*)$ in 6.3.2, we can perform the same computation as in 6.3.1 to verify the claims. Finally, we remark that $\mathrm{Sh}_{\mathcal{H}}$ is the same as what we call the category of sheaves over T, $\mathrm{Sh}_{\mathcal{H}} = \mathrm{Sh}(T)$. We have established:

Theorem 6.3.3 *Suppose that the topos $\mathcal{E} = \widetilde{\mathcal{C}}$ has enough points. Then there is a topological space and a conservative model $\widetilde{N}: \widetilde{\mathcal{C}} \to \mathrm{Sh}(T)$ such that $\widetilde{N}$ preserves all infs and $\forall$'s in $\widetilde{\mathcal{C}}$. (In fact, T is "the space of all points of $\mathcal{E}$" in a precise sense determined above.)*

By Theorems 6.2.2 and 6.2.4 now we have

Corollary 6.3.4 *The conclusion of 6.3.3 holds for any coherent topos and for any separable topos.*

Finally, we present a proof of an elegant theorem due to A. Joyal which is a variant of Kripke's completeness theorem, for intuitionistic logic (c.f. e.g. Fitting [1969]). For Joyal's theorem and applications of it, see also Joyal-Reyes [ ]. Before stating the theorem, we discuss a few auxiliary concepts.

In every Grothendieck topos, the operation of taking $\forall_f(X)$ for a subobject $X \hookrightarrow A$ and a morphism $A \xrightarrow{f} B$ can always be performed, and $\forall_f(X)$ is stable. Similarly, for Heyting implication, $X \to Y$, for subobjects X, Y of a given object A. We consider a special kind of Grothendieck topos, the functor category $\mathcal{E} = \mathbf{Set}^{\mathcal{K}}$ with an arbitrary category $\mathcal{K}$, and we compute $\forall_f(X)$ (and $X \to Y$) in $\mathcal{E}$.

Consider the diagram

$$\begin{tikzcd} F \arrow[r, hook, "i"] & G \arrow[d, "\nu"] \\ & H \end{tikzcd}$$

with i a monomorphism, in $\mathcal{E} = \mathbf{Set}^{\mathcal{K}}$; we are going to compute

$$\forall_\nu(F \overset{i}{\hookrightarrow} G).$$

$F(M) \overset{i_M}{\hookrightarrow} G(M)$ $(M \in \mathrm{Ob}(\mathcal{K}))$ is a monomorphism in **Set**; without loss of generality we can assume that i_M is a set-theoretic inclusion, for every $M \in \mathrm{Ob}(\mathcal{K})$. For the subobject $I = \forall_\nu(F \overset{i}{\hookrightarrow} G)$, again with set theoretic inclusions $I(M) \hookrightarrow H(M)$, we have the following formula: for any $M \in \mathrm{Ob}(\mathcal{K})$, for any $b \in H(M)$:

$(*)$ $b \in I(M) \Leftrightarrow$ for all morphisms $M \xrightarrow{g} N$ in $\mathcal{K}$, and for all $a \in G(N)$, if $\nu_N(a) = (Hg)(b)$, then $a \in F(N)$.

The effect of I on a morphism $M \xrightarrow{f} N$ is determined as the restriction of $H(f)$ to $I(M) \subset H(M)$; we have to check, however, that this restriction maps every element in its domain into $I(N)$. The reader can easily verify that this is so; when doing so he will realize why we have to consider 'all $M \xrightarrow{g} N$' in $(*)$, instead of just M (and $g = \mathrm{id}_M$). Furthermore, it is quite easy to check that the subobject $I \hookrightarrow H$ thus defined meets the requirements for being $\forall_\nu(F \to G)$; we omit the details. Computing Heyting implications is left to the reader.

The second concept we need is that of the *evaluation functor*. Let $\mathcal{E}, \mathcal{S}$ be categories and $\mathcal{K}$ a subcategory of the functor category $\mathcal{S}^{\mathcal{E}}$. The evaluation functor $\mathrm{ev}: \mathcal{E} \to \mathcal{S}^{\mathcal{K}}$

is defined as follows: (i) for an object X of $\mathcal{E}$, $\mathrm{ev}(X)$ is the functor $\mathcal{K} \to \mathcal{S}$ such that (a) for any $M \in \mathrm{Ob}(\mathcal{K})$, $\mathrm{ev}(X)(M) \underset{\mathrm{df}}{=} M(X)$ (note that M is a functor $M:\mathcal{E} \to \mathcal{S}$) and such that ($b$) for any $g:M \to N$ in $\mathcal{K}$, $\mathrm{ev}(X)(g) \underset{\mathrm{df}}{=} g_X : M(X) \to N(X)$ (note that g is a natural transformation $M \to N$); moreover (ii) for a morphism $X \xrightarrow{f} Y$ in $\mathcal{E}$, $\mathrm{ev}(f)$ is the natural transformation $\mathrm{ev}(X) \to \mathrm{ev}(Y)$ such that for $M \in \mathrm{Ob}(\mathcal{K})$ $(\mathrm{ev}(f))_M = M(f) : M(X) \to M(Y)$. It is easy to check that ev is indeed a functor.

We will apply the above in the situation when $\mathcal{E}$ is a coherent topos, $\mathcal{S} = \mathbf{Set}$, and $\mathcal{K}$ is the full subcategory $\mathrm{Mod}(\mathcal{E})$ of the functor category $\mathbf{Set}^{\mathcal{E}}$ whose objects are the $\mathbf{Set}$-models of $\mathcal{E}$, i.e., the continuous functors $\mathcal{E} \to \mathbf{Set}$ with respect to the canonical topologies on $\mathcal{E}$ and $\mathbf{Set}$. We will consider the functor $\mathrm{ev}:\mathcal{E} \to \mathbf{Set}^{\mathrm{Mod}(\mathcal{E})}$. Strictly speaking, we should and could consider a suitable small subcategory $\mathcal{K}$ of $\mathrm{Mod}(\mathcal{E})$ to make $\mathbf{Set}^{\mathcal{K}}$ locally small but this modification will be left to the reader to formulate.

Finally, in order to give a formulation which refers only to the coherent topos and not to a generating site, we use the notion of a coherent object. We will return to coherent objects in Chapter 9; here it suffices for us to know that for any coherent topos $\mathcal{E}$, we can present $\mathcal{E}$ as $\widetilde{\mathcal{C}}$ with an algebraic site $\mathcal{C}$ such that the coherent objects of $\mathcal{E}$ are exactly the objects isomorphic to some εA, $A \in \mathrm{Ob}(\mathcal{C})$, with $\varepsilon:\mathcal{C} \to \widetilde{\mathcal{C}} \simeq \mathcal{E}$ the representable sheaf functor and such that ε is full and faithful (c.f. Chapter 9).

Theorem 6.3.5 (A. Joyal) *For a coherent topos $\mathcal{E}$, the evaluation functor* $\mathrm{ev}:\mathcal{E} \to \mathbf{Set}^{\mathrm{Mod}(\mathcal{E})}$ *has the following properties:* (i) *it is a model of $\mathcal{E}$, (equivalently, it is ∞-logical),* (ii) *it is conservative,* (iii) *it preserves $\forall_f(X \hookrightarrow A)$ and $X \to Y$, whenever $X \hookrightarrow A$, $Y \hookrightarrow A$ are subobjects of a coherent object A, f is a morphism $A \to B$ between coherent objects.*

Proof. We will leave the proof of (i) and (ii) to the reader and concentrate on looking at $\forall_f(X)$. Let $\widetilde{\mathcal{C}}$ be a small algebraic site such that $\mathcal{E} = \widetilde{\mathcal{C}}$ and let $\varepsilon:\mathcal{C} \to \widetilde{\mathcal{C}}$ be the canonical functor.

Consider

$$\begin{tikzcd} & X \arrow[r, hook] & A' \arrow[d, "f'"] \\ \forall_f(X) = & Y \arrow[r, hook] & B' \end{tikzcd}$$

with $A' = \varepsilon A$, $B' = \varepsilon B$, $f' = \varepsilon f$, with $A \xrightarrow{f} B$ in $\mathcal{C}$, with $\varepsilon:\mathcal{C} \to \widetilde{\mathcal{C}}$. We are going to show that ev preserves $\forall_f(X)$. According to what was said above, this is enough to verify the claim made in (iii) about $\forall_f(X)$.

To verify that $\mathrm{ev}(Y) \hookrightarrow \mathrm{ev}(B')$ is $\forall_{\mathrm{ev}(f')}(\mathrm{ev}(X) \hookrightarrow \mathrm{ev}(A'))$, it is enough to verify that for any $\overline{M} \in \mathrm{Ob}(\mathrm{Mod}(\mathcal{E}))$, $(*)$ above is true where F is $\mathrm{ev}(X)$, G is $\mathrm{ev}(A')$, H is $\mathrm{ev}(B')$, ν is $\mathrm{ev}(f')$, with $I(M)$ replaced by $(\mathrm{ev}(Y))(\overline{M})$. Namely, as we said before, $(*)$ determines the subobject $I \hookrightarrow H$. We are left with the task of verifying the following:

$(**)$ $\quad b \in \overline{M}(Y) \;\Leftrightarrow\;$ for all morphisms $\overline{M} \xrightarrow{\overline{g}} \overline{N}$ in $\mathrm{Mod}(\mathcal{E})$ and for all $a \in \overline{N}(A')$, if $(\overline{N}(f'))(a) = \overline{g}_{B'}(b)$, then $a \in \overline{N}(X)$.

Next recall the following. Any model $\overline{M}:\mathcal{E} \to \mathbf{Set}$ can be represented as $\overline{M} = \widetilde{M}$ for a model $M:\mathcal{C} \to \mathbf{Set}$ with the commutative diagram

$$\begin{tikzcd} \mathcal{C} \arrow[r, "\varepsilon"] \arrow[dr, "M"'] & \widetilde{\mathcal{C}} \arrow[d, "\widetilde{M}"] \\ & \mathbf{Set}. \end{tikzcd}$$

Moreover, any model $M:\mathcal{C} \to \mathbf{Set}$ is a model of the theory $T_{\mathcal{C}}$ and conversely. We have $M(A') = M(A)$, etc, using $A' = \varepsilon A$, etc. Finally, the subobject $X \hookrightarrow \varepsilon A$ can be represented as the ε-interpretation of a (possibly infinite) disjunction $\bigvee_{i\in I}\underset{\sim}{X}_i(x)$ with $\underset{\sim}{X}_i$ finitary formulas of $(L_{\mathcal{C}})^g_{\omega\omega}$. Hence $\widetilde{M}(X \hookrightarrow \varepsilon A)$ will be $M_x(\bigvee_{i\in I}\underset{\sim}{X}_i(x)) \hookrightarrow M(A)$.

Taking these facts into account, $(**)$ translates into the following: here M is an arbitrary model $M:\mathcal{C} \to \mathbf{Set}$.

$(***)$ $b \in \widetilde{M}(Y) \Leftrightarrow$ for all natural transformations $M \xrightarrow{g} N$, with N a model $N:T_{\mathcal{C}} \to \mathbf{Set}$, and for all $a \in N(A)$, if $N(f)(A) = g_B(b)$ then $a \in N(\bigvee_{i\in I}\underset{\sim}{X}_i) = \bigcup_{i\in I} N(\underset{\sim}{X}_i)$.

Thus, in the righthand side of the equivalence $(***)$, we have achieved a reduction to a purely model theoretic condition concerning models of the theory $T_{\mathcal{C}}$. Of course, in proving the equivalence we have to relate this condition to the fact that $Y \hookrightarrow A'$ is $\forall_f(X)$ in the topos $\mathcal{E}$.

The left-to-right direction '$\Rightarrow$' in $(***)$ is trivial and is left as an exercise. For the other direction, we assume that $b \in M(B)$ but $b \notin \widetilde{M}(Y)$ and we construct

$(****)$ a model N of $T = T_{\mathcal{C}}$ together with a homomorphism (natural transformation) $g:M \to N$ and an element $a \in N(A)$ such that $N(f)(a) = g_B(b)$ and $a \notin \bigcup_{i\in I} \underset{\sim}{N}(X_i)$.

To this end, we employ the method of diagrams, one of the most commonly used methods in model theory. (More applications of the same method will appear in Chapter 7.) The method consists, roughly speaking, in translating the task of constructing the model N with additional items $g:M \to N$ and $a \in N(A)$ into showing that a certain set of axioms, in an extended language, can be simultaneously satisfied.

We introduce new individual constants denoting elements c of M: for each $c \in |M|_s$, let $\underline{c}_s$, or more briefly $\underline{c}$, be a new constant; we must have $\underline{c}_s = \underline{d}_t$ only if $s = t$ and $c = d$. Let L' be language $L = L_{\mathcal{C}}$ together with these constants $\underline{c}$, $c \in |M|$. We define $\mathrm{Diag}_+(M)$ (the "positive diagram" of M) as the set of all atomic sentences $\theta(\underline{c}_1,\dots,\underline{c}_n)$ of L' which are true in M when $\underline{c}_i$ is interpreted as c_i. E.g., if $(Mf)(c_1) \approx c_2$, with $f:A \to B$, then and only then we have the sentence $f(\underline{c}_1) \approx \underline{c}_2$ in $\mathrm{Diag}_+(M)$. The role of the positive diagram is contained in the following

Lemma 6.3.6 *Given a model N' of $\mathrm{Diag}_+(M)$, and its reduct N to L (obtained by forgetting the interpretations of symbols outside L), the map $g_A:M(A) \to N(A)$ defined by $c \mapsto (\underline{c})^{N'}$ give rise to a natural transformation $g = (g_A)_{A\in \mathrm{Ob}(\mathcal{C})} : M \to N$.*

The lemma will become obvious when one reflects that a natural transformation $M \to N$ is nothing but a homomorphism preserving the operations of the multisorted algebra M.

We write down the following set of axioms:

$$T' = T \cup \mathrm{Diag}_+(M) \cup \{f(\underline{a}) = \underline{b}\} \cup \{\neg\underset{\sim}{X}_i(\underline{a}) : i \in I\}$$

(as in model theory in general, we do not use Gentzen sequents to write down axioms). Here $\underline{a}$ is a *new* constant (not in L') of sort A, b is the specific element in $M(B) - \widetilde{M}(Y)$ we started with (hence $\underline{b}$ is in L') and the sentences $\underset{\sim}{X}_i(\underline{a})$ with the new constant $\underline{a}$ plugged in, come from the formulas $\underset{\sim}{X}_i$ referred to above. T' is formulated in finitary logic. We notice that if N' is a model of T' then, with N the reduct of N' to L, first of all, $N \models T\ (= T_{\mathcal{C}})$, and with g obtained as in Lemma 6.3.6, and with a $(\underline{a})^{N'}$ we will have all we wanted under $(****)$. So, it is sufficient to show that T' has a model!

At this point we refer to the Compactness Theorem, 5.1.8. Hence, it is enough to show that any finite subset T'' of T' has a model. Take a finite subset T'' of T'. Besides elements of T, T'' contains finitely many elements of $\mathrm{Diag}_+(M)$; without loss of generality, it contains $f(\underline{a}) \approx \underline{b}$, and it contains $\neg \underset{\sim}{X}_i(\underline{a})$ for a finite set I' of indices i. Let the conjunction of the formulas in $T'' \cap \mathrm{Diag}_+(M)$ be $\delta(\underline{b}, \underline{c}_1, \ldots, \underline{c}_n)$ and let $\phi(y)$ be the formula $\exists z_1 \cdots \exists z_n \delta(y, z_1, \ldots, z_n)$ with appropriate variables y, z_i. Because δ comes from the diagram of M, M *satisfies* ϕ at b, $M \models \phi[b]$.

We now assume that T'' has no model and derive a contradiction. The assumption implies that T together with the sentence

$$\delta(\underline{b}, \underline{c}_1, \ldots, \underline{c}_n) \wedge f\underline{a} \approx \underline{b} \wedge \textstyle\bigwedge_{i \in I'} \neg \underset{\sim}{X}_i(\underline{a})$$

cannot be satisfied by a model, i.e., that the sequent

$$\delta(\underline{b}, \underline{c}_1, \ldots, \underline{c}_n) \wedge f\underline{a} \approx \underline{b} \Rightarrow \textstyle\bigvee_{i \in I'} \underset{\sim}{X}_i(\underline{a})$$

is true in any model of T interpreting the constants $\underline{a}$, $\underline{b}$, $\underline{c}_i$. Replacing those constants by variables, and quantifying the c_i by $\exists$, we obtain

$$T \models \phi(y) \wedge f(x) \approx y \;\Rightarrow\; \textstyle\bigvee_{i \in I'} \underset{\sim}{X}_i(x);$$

here x, y are variables of sort A and B, respectively and $\models$ refers to logical consequence in terms of **Set**-models. Now we use the two-valued completeness theorem, e.g. 5.1.7(i), for $L^g_{\omega\omega}$. Notice that all sequents concerned are in the finitary coherent fragment $L^g_{\omega\omega}$. It follows that $T \vdash \phi(y) \wedge f(x) \approx y \;\Rightarrow\; \bigvee_{i \in I'} \underset{\sim}{X}_i(x)$ in either of the two complete axiomatizations we gave for $L^g_{\omega\omega}$ in Chapter 5.

Now we apply that the interpretation $\varepsilon : L_{\mathcal{C}} \to \mathcal{E} = \widetilde{\mathcal{C}}$ of the language is sound with respect to the axiomatization. Since ε is a model of $T = T_{\mathcal{C}}$, it follows that it is a model of $\phi(y) \wedge f(x) \approx y \;\Rightarrow\; \bigvee_{i \in I'} \underset{\sim}{X}_i(x)$, or equivalently, of $\phi(f(x)) \;\Rightarrow\; \bigvee_{i \in I'} \underset{\sim}{X}_i(x)$. Recall that ε interprets the sorts A and B as A', B', respectively and the operation f as $f' : A' \to B'$. Denoting the ε-interpretation of $\phi(y)$ by $Y' \hookrightarrow B'$, we have that the ε-interpretation of $\phi(f(x))$ is the pullback $(f')^{-1}(Y')$. The truth of the last sequent in $\mathcal{E}$, via its interpretation ε, means that the subobject $(f')^{-1}(Y')$ is $\leq$ the ε-interpretation of $\bigvee_{i \in I'} \underset{\sim}{X}_i(x)$. But the latter is obviously $\leq X \hookrightarrow A'$. Hence $(f')^{-1}(Y') \leq X$. By definition of the operation $\forall_f(\cdot)$ and of $\forall_f(X) = Y$, it follows that $Y' \leq Y$, hence $\widetilde{M}(Y') \subset M(Y)$. Now, as we noted above, $M \models \phi[b]$ for the particular element b. Since $M = \widetilde{M} \circ \varepsilon$, and Y' is the ε-interpretation of ϕ, this means that $b \in \widetilde{M}(Y') \subset \widetilde{M}(Y) \subset \widetilde{M}(B')$. We arrived at the *contradiction* $b \in \widetilde{M}(Y)$!

Recapitulating: this means that T'' does have a model as promised. Hence, so does T' and as we explained, this establishes that we can construct the items in $(****)$. By our previous reductions, this completes the proof of the theorem. $\square$

Chapter 7

Conceptual Completeness

§1 A completeness property of pretopoi

Recall that a logical category is one that has finite $\varprojlim$, stable images and stable sups of finite families of subobjects of any object (Definition 3.4.1). A *logical functor* between logical categories is one that preserves the logical structure, i.e. that preserves finite $\varprojlim$, images and finite sups (Definition 3.4.2). A *model* or *set-model* of a logical category $\mathcal{R}$ is a logical functor $M:\mathcal{R}\to\mathbf{Set}$. If we let $L_{\mathcal{R}}$ be the standard language associated with $\mathcal{R}$ (c.f. Chapter 2, Section 4), then any functor $M:\mathcal{R}\to\mathcal{S}$, in particular $M:\mathcal{R}\to\mathbf{Set}$, is an $\mathcal{S}$-interpretation of $L_{\mathcal{R}}$, or in particular in case $\mathcal{S}=\mathbf{Set}$, a many-sorted algebra of type $L_{\mathcal{R}}$: M is the family $\langle M(R), M(f)\rangle_{R\in\mathrm{Ob}(\mathcal{R}), f\in\mathrm{Mor}(\mathcal{R})}$ with $M(R)$ being a set (a *partial domain of* M) and

$$M(f):M(R)\to M(R')$$

an operation for $f:R\to R'$ in $\mathcal{R}$.

Let $\mathrm{Mod}(\mathcal{R})$ be the *category of all models of* $\mathcal{R}$: its objects are the models $M:\mathcal{R}\to\mathbf{Set}$ of $\mathcal{R}$ and $\mathrm{Mod}(\mathcal{R})$ is a full subcategory of the functor category $\mathbf{Set}^{\mathcal{R}}$. On an 'algebraic' level this means that the morphisms of $\mathrm{Mod}(\mathcal{R})$ are *homomorphisms* of (many-sorted) algebras: if $F:M\to M'$ is a morphism (i.e., a natural transformation of functors), then F is a family $(F_R)_{R\in\mathrm{Ob}(\mathcal{R})}$ such that

$$F_R:M(R)\to M'(R)$$

and such that F_R preserves the operations indexed by symbols f in $L_{\mathcal{R}}$:

$$\begin{tikzcd} R \arrow[d, "f"'] \\ R' \end{tikzcd} \qquad \begin{tikzcd} M(R) \arrow[r, "F_R"] \arrow[d, "M(f)"'] \arrow[dr, phantom, "\circlearrowright"] & M'(R) \arrow[d, "M'(f)"] \\ M(R') \arrow[r, "F_{R'}"'] & M'(R'). \end{tikzcd}$$

Next, recall that a functor $I:\mathcal{R}\to\mathcal{S}$ with a small logical category $\mathcal{R}$, is a logical functor jus in case I, as an $\mathcal{S}$-interpretation of the language $L_{\mathcal{R}}$ is an $\mathcal{S}$-*model* of the theory $T_{\mathcal{R}}$; $T_{\mathcal{R}}$ is formulated in the finitary coherent logic over $L_{\mathcal{R}}$, in other words, $T_{\mathcal{R}}$ is a finitary coherent theory (c.f. Theorem 3.5.3). In particular, $M:\mathcal{R}\to\mathbf{Set}$ is a model of $\mathcal{R}$ iff M is an ordinary model of the theory $T_{\mathcal{R}}$.

Throughout this chapter, we will assume that $\mathcal{R}$ and $\mathcal{S}$ are *small* logical categories and $I:\mathcal{R}\to\mathcal{S}$ is a logical functor. I^* will denote the functor

$$I^*:\mathrm{Mod}(\mathcal{S})\to\mathrm{Mod}(\mathcal{R})$$

induced by I by composition: for $M \in \mathrm{Ob}\,\mathrm{Mod}(\mathcal{S})$, $I^*(M) = M \circ I$ and for $F : M \to M'$ in $\mathrm{Mod}(\mathcal{S})$ $I^*(F) : I^*(M) \to I^*(M')$ is defined by

$$I^*(F)_R = F_{I(R)}, \quad \text{for } R \in \mathrm{Ob}(\mathcal{R}).$$

It is obvious that $M \circ I$, as a composition of logical functors, is again logical, hence a model of $\mathcal{R}$. It is equally obvious that $I^*(F)$ is a natural transformation $I^*(M) \to I^*(M')$ and that, in fact, I^* is a functor $\mathrm{Mod}(\mathcal{S}) \to \mathrm{Mod}(\mathcal{R})$.

We are going to prove results of the form that certain properties of I^* imply certain other properties of I. We will point out that those statements are essentially equivalent to model-theoretic ones, some of them well-known such as the Beth definability theorem. Also, we may point out what happens if $\mathcal{R}$ and $\mathcal{S}$ are Boolean algebras (regarded as categories in the in the usual way as partially ordered sets). Now, I is a Boolean homomorphism, $\mathrm{Mod}(\mathcal{R})$ is the *set* (discrete category) of all ultrafilters on $\mathcal{R}$ and I^* is the underlying function of the induced continuous map of the Stone spaces of $\mathcal{S}$ and $\mathcal{R}$. It is well-known that I^* is injective $\Rightarrow$ I is surjective and I^* is surjective $\Rightarrow$ I is injective (in fact, equivalences hold). The mentioned statements turn out to be special cases of new results below.

Next, we recall a consequence of the soundness and completeness theorems (3.5.4, 5.1.7). Let $R_1 \overset{f}{\hookrightarrow} R$ be a subobject of R in the category $\mathcal{R}$. For $M : \mathcal{R} \to \mathbf{Set}$, a model of $\mathcal{R}$, we identify $M(R_1)$ (as a subobject of $M(R)$) with the *subset* $\{r \in M(R) :$ there is $r_1 \in M(R_1)$ such that $M(f)(r_1) = r\}$ of $M(R)$. Let $R_1 \overset{f_1}{\hookrightarrow} R$, $R_2 \overset{f_2}{\hookrightarrow} R$ be two subobjects of R. By the above identification, $M(R_1) \leq M(R_2)$ in the ordering of subobjects of $M(R) \in \mathrm{Ob}(\mathbf{Set})$ iff $M(R_1) \subset M(R_2)$. We claim

Corollary 7.1.1 $R_1 \leq R_2$ *(in the ordering of subobjects of R) off $M(R_1) \leq M(R_2)$ for all models M of $\mathcal{R}$.*

For the proof, let $\underset{\sim}{R_1}(r)$ be the formula $\exists r_1 (f_1 r_1 \approx r)$ in the language $L_{\mathcal{R}}$, and similarly, $\underset{\sim}{R_2}(r)$. Then, $R_1 \leq R_2$ is equivalent to saying that $\mathcal{R}$ (the identical interpretation of $L_{\mathcal{R}}$) satisfies the sequent $\underset{\sim}{R_1}(r) \Rightarrow \underset{\sim}{R_2}(r)$. By the soundness theorem (3.5.4), this is equivalent to saying that $T_{\mathcal{R}} \vdash \underset{\sim}{R_1}(r) \Rightarrow \underset{\sim}{R_2}(r)$ and this (by completeness for set-models, 5.1.7) is equivalent to saying that $M \models \underset{\sim}{R_1}(r) \Rightarrow \underset{\sim}{R_2}(r)$ for each ($\mathbf{Set}$-)model $M \models T_{\mathcal{R}}$, hence to saying that $M \models \underset{\sim}{R_1}(r) \Rightarrow \underset{\sim}{R_2}(r)$ for each model of $\mathcal{R}$, i.e. to saying that $M(R_1) \subseteq M(R_2)$ for all models M of $\mathcal{R}$. $\square$

7.1.1 has an immediate consequence for the situation $I : \mathcal{R} \to \mathcal{S}$ introduced above.

Theorem 7.1.2 *With $\mathcal{R}$, $\mathcal{S}$, I as above, assume that $I^* : \mathrm{Mod}(\mathcal{S}) \to \mathrm{Mod}(\mathcal{R})$ is* surjective (on objects)*, i.e. for every $M \in \mathrm{Ob}(\mathrm{Mod}\,\mathcal{R})$ there is an $N \in \mathrm{Ob}(\mathrm{Mod}\,\mathcal{S})$ such that $I^*(N) \simeq M$ (isomorphisms in the category $\mathrm{Mod}(\mathcal{R})$). Then I is* conservative*, i.e. if R_1, R_2 are subobjects of R in $\mathcal{R}$ and $I(R_1) \leq I(R_2)$ (in the ordering of subobjects of $I(R)$), then $R_1 \leq R_2$. Hence, in particular, I is faithful.*

Proof. Suppose $R_1 \not\leq R_2$. By 7.1.1, there is $M \in \mathrm{Ob}(\mathrm{Mod}\,\mathcal{R})$ such that $M(R_1) \not\subset M(R_2)$. We have $N \in \mathrm{Ob}(\mathrm{Mod}\,\mathcal{S})$ such that $I^*(N) \simeq M$ and we can assume that $I^*(N) = M$. Then $M(R_i) = N(I(R_i))$ for $i = 1, 2$ and hence $N(I(R_1)) \not\subset N(I(R_2))$. By 7.1.1 again (now applied to $\mathcal{S}$), $I(R_1) \not\leq I(R_2)$. Q.E.D.

Faithfulness follows by considering the graph $G \hookrightarrow A \times B$ of a morphism $A \xrightarrow{f} B$. $\square$

The next result is related to the Beth definability theorem (c.f. CK).

Theorem 7.1.3 *Assume now in addition that $\mathcal{R}$ and $\mathcal{S}$ are* Boolean*, i.e. any subobject $A \hookrightarrow B$ of B has a (Boolean) complement $C \to B$ such that $A \wedge C = 0$ and $A \vee C =$*

B. *Assume* I^* *if* full with respect to isomorphisms, *i.e., if* $I^*(N_1) \xrightarrow{F} I^*(N_2)$ *is an isomorphism in* $\mathrm{Mod}(\mathcal{R})$*, then there is a morphism* $N_1 \xrightarrow{G} N_2$ *such that* $F = I^*(G)$*. Then* I *is* full with respect to subobjects*, i.e., if* $S \hookrightarrow I(R)$ *is a subobject of* $I(R)$ *in* $\mathcal{S}$*, then there is* $R' \hookrightarrow R$ *in* $\mathcal{R}$ *such that* $S \simeq I(R')$ *(isomorphism of* subobjects*).*

Proof. Recall Beht's theorem:

Beth' Theorem *Let* T *be a theory in full first order logic over a many sorted language* L'*, let* $I : L \to L'$ *be an interpretation of the language* L *in* L'*, and let* ϕ *be a formula with free variables* $\vec{v}$ *of* L' *each of which are of a sort* $I(r)$*,* r *a sort of* L*. Assume that whenever* N_1 *and* N_2 *are (ordinary set-)models of* T *such that* $I^*(N_1) = I^*(N_2)$*, then* $\phi^{N_1} = \phi^{N_2}$*. Then there is a formula* ψ *in the full first order logic over* L *such that* $T \vdash \forall\vec{v}[\phi \leftrightarrow (I(\psi))(\vec{v})]$

Remarks (i) *by an* interpretation $I : L \to L'$ *of languages, we mean here a map* I *that associates a sort* $I(s)$ *with each sort* s *of* L*, and associates an operation symbol* $I(f) : I(s_1) \to I(s_2)$ *with any operation symbol* $f : s_1 \to s_2$ *in* L*. (Here we deal with languages that do not have items other that sorts and unary operation symbols. In the usual definition of interpretation, one associates formulas of* L' *with the primitive symbols of* L*; the above more special definition is sufficient for us.)*

(ii) *For a formula* $\phi = \phi(\vec{v})$ *with free variables* $\vec{v}$*,* $\vec{v} = v_1, \ldots, v_n$*,* v_i *of sort* s_i *and a structure* N*,* $\phi^N = \{\vec{a} \in \times_{i=1}^{n} |N|_{s_i} : N \models \phi[\vec{a}]\}$*.*

(iii) *For an interpretation* $I : L \to L'$ *as in* (i)*, and a structure* N *for* L'*,* $I^*(N)$ *is defined in a natural way to be a structure for* L *(so that if* $I : \mathcal{R} \to \mathcal{S}$ *is a functor,* $L = L_{\mathcal{R}}$*,* $L' = L_{\mathcal{S}}$ *and* $N : \mathcal{S} \to \mathbf{Set}$*, then* $I^*(N)$ *has the same meaning as before, namely* $N \circ I$*.*

(iv) *For* $I : L \to L'$ *as before, and a formula* ψ *of* L*, by* $I(\psi)$ *we mean the result of replacing each variable* x *of sort* s *by a variable* y *of sort* $I(s)$ *(keeping distinct variables distinct) and each operation symbol* f *by* $I(f)$*.*

(v) *The assumption of the theorem is equivalent to saying that whenever* $N_1, N_2 \models T$ *and* $F : I^*(N_1) \xrightarrow{\sim} I^*(N_2)$*, then* $F(\phi^{N_1}) = \phi^{N_2}$*, or rather* $(F_{s_1} \times F_{s_2} \times \cdots \times F_{s_n})(\phi^{N_1} = \phi^{N_2})$ *where* s_i *is the sort of* v_i*,* $\vec{v} = \langle v_i : i = 1, \ldots, n\rangle$ *and* $F_s : N_1(s) \to N_2(s)$ *the* s*-part of the isomorphism* F*.*

(vi) *In the usual formulation of Beth's theorem,* L *is a sublanguage of* L' *and* I *is the inclusion map. In this case,* $I^*(N)$ *(for* N *a structure for* L'*) is the reduct* $N|L$ *of* N*, the result of forgetting interpretations of symbols in* $L' - L$*. (Unlike in one-sorted logic, the reduct* $N|L$ *may have a smaller domain than* N*.) Also, usually* ϕ *of the above formulation is replaced by a predicate symbol. It is easy to see that the above formulation follows from the "usual" one.*

Let us consider the language $L = L_{\mathcal{R}}$ and $L' = L_{\mathcal{S}}$; I then gives an interpretation of L in $L' : I : L \to L'$. Let the theory T be $T_{\mathcal{S}}$, in the language L'. Let the formula ϕ be $\underset{\sim}{S} :=$ $\underset{\sim}{S}(x) := \exists s\ fs \approx x$, for a subobject $S \xhookrightarrow{f} I(R)$, x a variable of sort $I(R)$. Let F be an isomorphism (in the ordinary sense, *or* in $\mathrm{Mod}(\mathcal{R})$, what is the same) $I^*(N_1) \xrightarrow{\sim} I^*(N_2)$. Then we have $N_1 \xrightarrow{G} N_2$ such that $F = I^*(G)$. Consider $g = G_{I(R)} = F_R : (I^*(N_1))(R) = N_1(I(R)) \to (I^*(N_2))(R) = N_2(I(R))$ and $g' = G_S : N_1(S) \to N_2(S)$. Identifying $N_1(S)$ and $N_2(S)$ with subsets of $N_1(I(R))$ and $N_2(I(R))$, respectively, as usual, it follows from the fact that G is a natural transformation for functors from the category $\mathcal{S}$ that

g' is the restriction of g to the set $N_1(S)$,hence that F_R *preserves* the formula $\underset{\sim}{S}$, i.e. $F_R(\underset{\sim}{S}^{N_1} \subset \underset{\sim}{S}^{N_2}$ ($\underset{\sim}{S}^{N_i}$ being $N_i(S)$, $i = 1$ and 2; $F_R(\underset{\sim}{S}^{N_1})$ is the image of S^{N_1} under F_R.)

Notice (this remark has an obvious general significance) that a model N of $\mathcal{S}$ will preserve Boolean complements: if $A \hookrightarrow B$ and $C \hookrightarrow B$ are in $\mathcal{S}$ and C is a Boolean complement of A, then $N(C)$ will be the usual set-theoretic complement of $N(A)$ with respect to $N(B)$. This is immediate from the fact that N preserves $\wedge$ and $\vee$. Taking S' to be a Boolean complement of S with respect to $I(R)$, we obtain (similarly as for S) that $F_R(\underset{\sim}{S'}^{N_1}) \subset \underset{\sim}{S'}^{N_2}$, *hence* that $F_R(\underset{\sim}{S}^{N_1}) = \underset{\sim}{S}^{N_2}$.

Hence we see that the hypothesis of Beth's theorem is satisfied. Thus we have a formula $\phi(r)$ of full first order logic over L (r a variable of sort R) such that $T \vdash \forall x[\underset{\sim}{S}(x) \leftrightarrow (I\phi)(x)]$. It follows that $S = (\underset{\sim}{S})^{\mathcal{S}} \simeq (I(\phi))^{\mathcal{S}}$. (Here and below we write $(\underset{\sim}{S})^{\mathcal{S}}$ for the interpretation of the formula in $\underset{\sim}{S}$ in the canonical language of $\mathcal{S}$ by the identical interpretation $\mathcal{S}$ of that language. In Chapter 2 and 3, we wrote this as $[\underset{\sim}{S}]$ but here we have to refer to more categories simultaneously.) The above mentioned fact about preserving Boolean complements extends obviously to any logical functor, hence to I. But then we have that $(I\phi)^{\mathcal{S}} = I(\phi^{\mathcal{R}}) = I(R')$ for $R' = \phi^{\mathcal{R}}$, hence $S \simeq I(R')$ as required.

The next result is similar to 7.1.3. Its source is a preservation theorem; we will give a proof for it in detail. The chief feature of the result is that it does not need the hypothesis of Booleanness but otherwise its hypothesis is stronger that that of 7.1.3.

Theorem 7.1.4 *If I^* is full, then I is full with respect to subobjects (for this phrase, c.f.* 7.1.3*). If in addition I^* is surjective on objects, then I is full.*

Theorem 7.1.4$'$ *Let I be an interpretation of languages: $I : L \to L'$, T a theory in first order logic over L'. Let $\phi(\vec{x})$ be a formula of $L'_{\omega\omega}$ such that the sort of each free variable x_i is of the form $I(r_i)$, r_i a sort in L. Assume that whenever $F : I^*(N_1) \to I^*(N_2)$ is a homomorphism, then $F_r(\phi^{N_1}) = \phi^{N_2}$. ($F_r$ is $F_{r_1} \times F_{r_2} \times \cdots \times F_{r_n} : N_1(I(r)) = N_1(I(r_1)) \times \cdots \times N_n(I(r_n)) \to N_2(I(r)) = N_2(I(r_1)) \times \cdots \times N_2(I(r_n))$; ϕ^{N_1} is a subset of $N_1(I(r))$ and ϕ^{N_2} is a subset of $N_2(I(r))$.) We briefly say that any L-homomorphism between models of T* preserves *ϕ.*

Under these hypotheses, there is a finitary coherent formula ψ (built up by finitary $\bigwedge$ and $\bigvee$, and $\exists$) of L such that $T \models \forall x[\phi(\vec{x}) \Leftrightarrow (I(\psi))(\vec{x})]$.

Proof. This theorem is very closely related to the Los-Tarski theorem on sentences preserved by substructures, and in fact, the proof of 7.1.4$'$ can be given along the same lines, notably by using the "method of diagrams".

To simplify notation, we'll assume that I is an inclusion, i.e. $L \subset L'$; now $I^*(N)$ is the *reduct* $N|L$. The general case in fact follows from this special one.

Let us introduce new individual constants $\vec{d}$ of sorts matching those of $\vec{x}$ (the free variables of ϕ).Expand the language L to $L(\vec{d})$ and L' to $L(\vec{d})$ to include the $\vec{d}$. Let $\phi_0 \underset{\text{df}}{=} \phi(\vec{d})$. Then the hypothesis is equivalent to saying that whenever N_1, N_2 are $L'(\vec{d})$-models of T, F is an $L(\vec{d})$-homomorphism of $N_1|L(\vec{d})$ into $N_2|L(\vec{d})$, then $N_1 \models \phi_0$ implies $N_2 \models \phi_0$. In other words, we have reduced the hypothesis to preservation of sentences. Notice that any coherent (= positive existential) $L(\vec{d})$ sentence is preserved by $L(\vec{d})$-homomorphisms; what we wan to show is essentially the converse of this fact.

We are going to work in the full finitary (Boolean) logic over $L'(\vec{d})$. Now we can dispense with Gentzen sequents and take theories as sets of sentences, by replacing $\Phi \Rightarrow \Psi$ by $\forall\vec{x}(\bigwedge\Phi \to \bigvee\Psi)$.

Define Σ to be the set of *negated positive existential* sentences $\neg\varepsilon$ of the language $L_1 = L(\vec{d})$ which are consequences of $T \cup \{\neg\phi_0\} : T \cup \{\neg\phi_0\} \models \varepsilon$.

Lemma *Let us assign a new individual constant $\underline{a}_s$ to each element $a \in |M|_s$ such that $\underline{a}_s = \underline{b}_s$ implies $a = b$ and $s = s'$. The* diagram Δ *of M is the set of atomic sentences in the language: L_1' plus the new individual constants $\underline{a}_s$, which are true in M when $\underline{a}_s$ is interpreted as a. The main fact about the diagram which can be seen immediately is that whenever N is a* model *of Δ, then the maps $F_s : a \mapsto (\underline{a}_s)^N$ form a homomorphism $F = (F_s)_s$ of M into N. Hence, once we know that $T \cup \{\neg\phi_0\} \cup \Delta$ has a model, we are done.*

By the compactness theorem, it is enough to see that for any finite subset Δ' of Δ,

$$T \cup \{\neg\phi_0\} \cup \Delta'$$

has a model. $\bigwedge\Delta'$ can be written as $\gamma(\underline{a}^1_{s_1}, \dots, \underline{a}^n_{s_n})$ where $\gamma(x_1, \dots, x_n)$ is the conjunction of atomic formulas of the language L_1. Assuming that the assertion is false amounts to saying that

$$T \cup \{\neg\phi_0\} \models \neg\exists x_1, \dots, x_n \gamma(x_1, \dots, x_n)$$

since the constants $\underline{a}^i s_i$ do not occur in $T \cup \{\neg\phi_0\}$. Denoting the formula on the right-hand-side by $\neg\varepsilon$, we obtain that $\neg\varepsilon \in \Sigma$ since ε is positive existential obviously. By assumption, this implies that $M \models \neg\varepsilon$. On the other hand, clearly $M \models \varepsilon$ since $M \models \gamma[a^1/x_1, \dots, a^n/x_n]$. This contradiction proves the Lemma.

Turning to the proof of 7.1.4′, assume now that ϕ_0 is preserved by L_1-homomorphisms. We claim that for Σ of the Lemma

$$T \cup \Sigma \models \neg\phi_0.$$

Indeed, let M be any L_1'-structure which is a model of $T \cup \Sigma$; we want to show that $M \models \neg\phi_0$. By the Lemma, there is $N \models T \cup \{\neg\phi_0\}$ and $F : M|L \to N|L$. By the preservation property, $M \models T \cup \{\phi_0\}$ and $N \models T$ imply $N \models \phi_0$; since we know $N \models \neg\phi_0$ and $N \models T$, and also $M \models T$, it follows that $M \models \neg\phi_0$, as was to be shown.

Finally, apply the compactness theorem to $T \cup \Sigma \models \neg\phi_0$ to obtain finitely many $\neg\varepsilon_1, \dots, \neg\varepsilon_n \in \Sigma$ (possibly $n = 0$) such that $T \models \neg\psi_0 \to \neg\phi_0$ where $\psi_0 = \bigvee\{\varepsilon_k : k = 1, \dots, n\}$ ($= \bigvee\varnothing = \uparrow$ if $n = 0$). Clearly, $T \models \neg\phi_0 \to \neg\psi_0$ since every formula $\neg\varepsilon$ in Σ was consequence of $T \cup \{\neg\phi_0\}$. We conclude that $T \models \phi_0 \;\Leftrightarrow\; \psi_0$; obviously, ψ_0 is a coherent L_1-formula.

Now, $\psi_0 = \psi(\vec{d})$ for a coherent L-formula $\psi(\vec{x})$ and we have $T \models \forall x[\phi \Leftrightarrow \psi]$, as required. □

Proof of 7.4.1 This in entirely similar to the derivation of 7.1.3 form Beth's theorem. We put $L = L_{\mathcal{R}}$ $L' = L_{\mathcal{S}}$, $T = T_{\mathcal{S}}$ and, for a subobject $S \to I(R)$, $\underset{\sim}{S}(x)$ (with x of sort $I(R)$) the usual formula such that $(\underset{\sim}{S})^{\mathcal{S}} = S$. Similarly as in the proof of 7.1.3, the hypothesis of I^* being full implies that any Mod $\mathcal{R}$-morphism

$$F : I^*(N_1) \to I^*(N_2)$$

preserve the formula $\underset{\sim}{S}$:

$$F_R((\underset{\sim}{S})^{N_1}) = (\underset{\sim}{S})^{N_2}.$$

So the hypothesis of 7.1.4′ is satisfied. Hence there is a coherent $L_{\mathcal{R}}$ formula ψ such that $T_{\mathcal{S}} \vdash \forall x[\underset{\sim}{S}(x) \Leftrightarrow (I\psi)(x)]$, hence $S \simeq (I\psi)^{\mathcal{S}} = I(\psi^{\mathcal{R}}) = I(R')$ for $R' = \psi^{\mathcal{R}}$; here we used that ψ is coherent and hence I respects ψ: $(I\psi)^{\mathcal{S}} = I(\psi^{\mathcal{R}})$.

To show the second assertion in 7.1.4, let $g : I(R_1) \to I(R_2)$ and let us consider the graph of $g : G \hookrightarrow I(R_1) \times I(R_2) = I(R_1 \times R_2)$. By the main part of 7.1.4, we have some subobject $F \hookrightarrow R_1 \times R_2$ such that $I(F) = G$. Now, by 7.1.2, I is conservative. Since G

is functional, it follows easily from conservativeness that F is functional. By Theorem 2.4.4, there is $f : R_1 \to R_2$ such that F is the graph of f. It follows from $I(F) = G$ that $I(f) = g$. $\square$

Definition 7.1.5 *An object S of $\mathcal{S}$ is* finitely covered by $\mathcal{R}$ via I *if there are finitely many objects $R_1, \dots, R_n$ in $\mathcal{R}$, there are subobjects $S_i \hookrightarrow I(R_i)$ in $\mathcal{S}$, and there are morphisms $S_i \xrightarrow{f_i} S$ in $\mathcal{S}$ such that $S = \bigvee_{i=1}^{n} \exists_{f_i}(S_i)$.*

Theorem 7.1.6 *Assume that I^* is faithful. Then every S in $\mathrm{Ob}(\mathcal{S})$ is finitely covered by $\mathcal{R}$ via I.*

Theorem 7.1.6′ *Let $I : L \to L'$ and T be as before, S a sort of L'. Suppose whenever F and G are L'-homomorphisms $M \to N$ of models M, N of T, then $I^*(F) = I^*(G)$ implies $F = G$. Then there are finitely many (finitary) coherent formulas $\phi_i(\vec{r}^{\,i}, s)$ of the language L' where the variables $\vec{r}^{\,i}$ have sorts $I(R^i)$, $\vec{R}^i$ sorts of L, such that*

(i) *"$\phi_i(\vec{r}^{\,i}, s)$ defines a partial function $\vec{r}^{\,i} \mapsto s$ in T" i.e.,*
$T \models \forall \vec{r}^{\,i} \forall s \forall s' [\phi_i(\vec{r}^{\,i}, s) \wedge \phi_i(\vec{r}^{\,i}, s') \to s = s']$ *and*

(ii) *"these functions cover S", i.e.*

$$T \models \forall s \bigvee_i \exists \vec{r}^{\,i} \phi_i(\vec{r}^{\,i}, s).$$

Proof. Assume that the conclusion does not hold. Consider the set Φ of all (finitary) coherent formulas $\phi(\vec{r}, s)$ of L', with $\vec{r}$ "coming from L" as above, for which condition (i) above holds. Let d be a new individual constant of sort S. Consider the set of sentences

$$T' = T \cup \{\neg \exists \vec{r} \phi(\vec{r}, d) : \phi \in \Phi\}.$$

T' is finitely consistent; otherwise we would have, after all, a system of formulas satisfying both (i) and (ii) above. By the compactness theorem, we have a model of T', say (M, c) where c interprets d We have that $c \in |M|_S$ is not in the set ("range of ϕ")

$$\{b \in |M|_S : M \models \exists \vec{r} \phi(\vec{r}, y)[b/y]\}$$

for any coherent ϕ that satisfies (i) above.

Next, we will construct N, F and G such that $F, G : M \to N$, and $F \neq G$, in fact $F_S(c) \neq G_S(c)$, but $I^*(F) = I^*(G)$. Let us introduce new individual constants $\underline{a}$ (more precisely, $\underline{a}_{I(R)}$) for each $a \in |M|_{I(R)}$ and for each sort $I(R)$, R in L; moreover, *two* individual constants $\underline{b}$ and $\underline{\underline{b}}$ for each $b \in |M|_{S'}$ and for each other sort S' in L' (again, $\underline{b}$ is $\underline{b}_{S'}$ and $\underline{\underline{b}}$ is $\underline{\underline{b}}_{S'}$ in fact). In particular, c will get two names, $\underline{c}$ and $\underline{\underline{c}}$.

Let $\mathrm{Diag}_{-}(M)$ denote the set of atomic sentences $\tau(\underline{\vec{a}}, \underline{\vec{b}}, \underline{c})$, with $\underline{\vec{a}}$ and $\underline{\vec{b}}$ new constants as explained, which are true in M when $\underline{a}$, $\underline{b}$, $\underline{c}$ denote a, b, c, respectively. Similarly, let $\mathrm{Diag}_{=}(M)$ be the set of atomic $\tau(\underline{\vec{a}}, \underline{\underline{\vec{b}}}, \underline{\underline{\vec{c}}})$ with similar condition on truth.

Finally, let us consider the set of sentences

$$T'' = T \cup \mathrm{Diag}_{-}(M) \cup \mathrm{Diag}_{=}(M) \cup \neg \underline{c} \approx \underline{\underline{c}}.$$

The point of T'' is, of course, that if N is a model of T'' then the maps

$$\begin{array}{rcll} F : a & \mapsto & (\underline{a})^N & \\ b & \mapsto & (\underline{b})^N & \quad (\text{including } c \mapsto (\underline{c})^N) \end{array}$$

and

$$\begin{array}{rcll} G : a & \mapsto & (\underline{a})^N & \\ b & \mapsto & (\underline{\underline{b}})^N & \text{(including } c \mapsto (\underline{\underline{c}})^N) \end{array}$$

will be L'-homomorphisms; moreover, we will have $I^*(F) = I^*(G)$ (since $F(a) = G(a)$ and $F(c) \neq G(c)$. *Hence*, it is sufficient to show that T'' is finitely consistent. A finite subset of T'' is a subset of a set of the form

$$T \cup \{\neg \underline{c} \approx \underline{\underline{c}}\} \cup \{\tau_i(\underline{\vec{a}}_i, \underline{\vec{b}}_i, \underline{c}) : i = 1, \dots, n\} \cup \{\tau_i'(\underline{\vec{a}'}_i, \underline{\underline{\vec{b}'}}_i, \underline{\underline{c}}) : i = 1, \dots, n\}.$$

By further enlarging the set we can assume that our set is

$$T \cup \{\neg \underline{c} \approx \underline{\underline{c}}\} \cup \{\tau_i(\underline{\vec{a}}_i, \underline{\vec{b}}_i, \underline{c}) : i = 1, \dots, n\} \cup \{\tau_i(\underline{\vec{a}'}_i, \underline{\underline{\vec{b}'}}_i, \underline{\underline{c}}) : i = 1, \dots, n\}.$$

with the *same* formulas $\tau_i(\vec{x}, \vec{y}, z)$ and elements $\vec{a}_i$, $\vec{b}_i$ appearing in both set-formations. Taking conjunction $\gamma = \bigwedge\{\tau_i : i = 1, \dots, n\}$, we are reduced to consider the consistency of $T \cup \{\gamma(\underline{\vec{a}}, \underline{\vec{b}}, \underline{c})\} \cup \{\gamma(\underline{\vec{a}}, \underline{\underline{\vec{b}}}, \underline{\underline{c}})\}$.

Suppose this last set is inconsistent. Then also $T \cup \{\exists \vec{y}\ \gamma(\underline{\vec{a}}, \vec{y}, \underline{c}) \wedge \exists \vec{y}\ \gamma(\underline{\vec{a}}, \vec{y}, \underline{\underline{c}}) \wedge \neg \underline{c} \approx \underline{\underline{c}}\}$ is inconsistent. Put $\phi(\vec{r}, z) := \exists \vec{y}\ \gamma(\vec{r}, \vec{y}, z)$ with $\vec{r}$ and z being variables. The last mentioned inconsistency amounts to saying that $\phi(\vec{r}, z)$ defines a partial function $\vec{r} \mapsto z$ in T (c.f. (i) above) and clearly, ϕ is coherent. But, since by the definition of $\mathrm{Diag}_{-}(M)$, we also have $M \models \gamma(\underline{\vec{a}}, \underline{\vec{b}}, \underline{c})$, hence $M \models (\exists \vec{r}\ \phi(\vec{r}, z))[c/z]$, contradicting the property of c that it does not lie in the range of such a partial function.

We have shown that T'' is consistent, and hence as explained above, also that $F, G : M \to N$ exist such that $N \models T$, $F \neq G$ and $I^*(F) = I^*(G)$. Since this is the negation of the assumption of 7.1.6′, this completes the proof of 7.1.6′. $\square$

The proof of 7.1.6 is immediate on the basis of 7.1.6′. We put (as before) $L = L_{\mathcal{R}}$, $L' = L_{\mathcal{S}}$, $T = T_{\mathcal{S}}$, we have $I : L \to L'$. The hypotheses of 7.1.6′ are satisfied, so we have a system of formulas $\phi_i(\vec{r}^{\,i}, s)$ with properties (i) and (ii). We put $R_i = R_1^i \times \cdots \times R_{k_i}^i$ where $\vec{r}^{\,i} = \langle r_1^i, \dots, r_{k_i}^i \rangle$, r_j^i is of sort $I(R_j^i)$. Then $(\exists s\ \phi_i(\vec{r}^{\,i}, s))^{\mathcal{S}}$ is a subobject S_i of $I(R_i)$. $(\phi_i(\vec{r}^{\,i}, s))^{\mathcal{S}}$ is a subobject of $I(R_i) \times S$ and can be regarded as a subobject of $S_i \times S$. Condition (i) tells us that $(\phi_i(\vec{r}^{\,i}, s))^{\mathcal{S}}$ as a subobject of $S_i \times S$ is functional, hence there is a (unique) morphism $S_i \xrightarrow{f_i} S$ whose graph is $(\phi_i(\vec{r}^{\,i}, s))^{\mathcal{S}}$. Finally, condition (ii) clearly says that $\bigvee \exists_{f_i}(S_i) = S$. $\square$

Lemma 7.1.7 *Let $\mathcal{R} = \mathcal{P}$ be a pretopos (c.f. Definition* 3.4.3*), $\mathcal{P} \xrightarrow{I} \mathcal{S}$ logical as before. Assume that I is full with respect to subobjects (c.f.* 7.1.3*) and conservative (c.f.* 7.1.2*). Then every $S \in \mathrm{Ob}(\mathcal{S})$ finitely covered by $\mathcal{R}$ via I is in the essential image of $\mathcal{R}$: there is R such that $S \simeq I(R)$.*

Remark *This lemma is a simplified version of* 1.4.11.

PROOF. We have finitely many $R_i \in \mathrm{Ob}(\mathcal{R})$, $(i = 1, \dots, n)$, $S_i \to I(R_i)$ and $S_i \xrightarrow{f_i} S$ such that $S = \bigvee_i \exists_{f_i}(S_i)$. Since I is full with respect to subobjects, we may assume that $S_i = I(R_i)$. Let us form the disjoint sum $R \underset{\text{df}}{=} \coprod_i R_i$.

Then $I(R) = \coprod_i I(R_i)$ (c.f. the remark preceding 3.4.6) and the family of maps f_i induces a map $I(R) \xrightarrow{f} S$ such that $\exists_f I(R) = \bigvee_i \exists_{f_i} I(R_i) = S$, in other words, f is an effective epimorphism. Let

$$S' \overset{p_1}{\underset{p_2}{\rightrightarrows}} I(R) \xrightarrow{f} S$$

be the kernel-pair of f, i.e., the diagram

$$\begin{tikzcd} S' \arrow[r, "p_1"] \arrow[d, "p_s"'] & I(R) \arrow[d, "f"] \\ I(R) \arrow[r, "f"'] & S \end{tikzcd}$$

is a pullback diagram. We have that $S' \overset{\langle p_1,p_2\rangle}{\hookrightarrow} I(R) \times I(R) = I(R \times R)$ is a mono, hence S' (as a subobject of $I(R \times R)$) is isomorphic to $I(R')$ for some R', subobject of $R \times R : R' \underset{\rho_2}{\overset{\rho_1}{\rightrightarrows}} R$. We can assume that $S' = I(R)$, $p_i = I(\rho_i)$ for $i = 1, 2$. $S' \underset{p_2}{\overset{p_1}{\rightrightarrows}} I(R)$, being a kernel pair, is an equivalence relation on $I(R)$. By the conservativeness of I, the axioms defining the equivalence relation $I(R)$ are 'reflected' by I, hence ti follows that $R' \underset{\rho_2}{\overset{\rho_1}{\rightrightarrows}} R$. is an equivalence relation on R. Now, let us form the quotient R/R': $R' \underset{\rho_2}{\overset{\rho_1}{\rightrightarrows}} R \xrightarrow{\rho} R/R'$ in the pretopos $\mathcal{R}$. I preserves images, hence

$$I(R) \xrightarrow{I(\rho)} I(R/R')$$

is an effective epimorphism. Since (ρ_1, ρ_2) is the kernel pair of ρ and I preserves finite $\varprojlim$, (p_1, p_2) is the kernel pair of $I(\rho)$. But an effective epimorphism is the coequalizer of its kernel pair. Hence $I(R) \xrightarrow{I(\rho)} I(R/R')$ is a coequalizer of p_1 and p_2. For the same reason, so is $I(R) \xrightarrow{f} S$. Hence $S \simeq I(R)$. □

Theorem 7.1.8 (Main result). *Suppose $\mathcal{P}$ is a pretopos and $I : \mathcal{P} \to \mathcal{S}$ is a logical functor. If $I^* : \mathrm{Mod}(\mathcal{S}) \to \mathrm{Mod}(\mathcal{R})$ is an equivalence of categories, then so is I.*

Proof. By 7.1.2, I is conservative and in particular, I is faithful. By 7.1.4, I is full with respect to sub objects and also, I is full. By 7.1.6 and 7.1.7, every object S in $\mathcal{S}$ is in the essential image or I. Thus, I is an equivalence. □

§2 Infinitary generalizations; preliminaries

There are generalizations of Section 1 that concern models preserving infinite sups. As we have seen, from the logical point of view, these are models of infinitary sentences. As is well konwn, Beth's definability theorem and many other results (notably preservation theorems) have suitable generalizations to $L_{\infty\omega}$ and to certain fragments of $L_{\infty\omega}$ (c.f. Keisler [1971], Makkai [1969]). Actually, in this context mainly countable fragments of $L_{\infty\omega}$ (of $L_{\omega_1\omega}$) have been considered in the literature by by using techniques of Mansfield [1972], we can easily get full generalizations for Boolean valued models so that the results on ordinary models and $L_{\omega_1\omega}$ are immediate consequences. Below we will illustrate this by giving infinitary generalizations of 7.1.4 and 7.1.6, and thus, of 7.1.8 as well. These will be generalizations in the proper sense, i.e., the original results will follow from them. However, the proofs of the infinitary versions are considerably less transparent so it has seemed worthwhile to give the proofs of Section 1 as well.

In Makkai [1969], consistency properties are the main tool for proving preservation theorems. It is an interesting technical point that, unlike in the previous section, model-theoretical results as such don not seem to be quite sufficient for the present purposes and thus certain modifications of the original notion of consistency property, etc., are necessary. It is easy to point out the reasons already at this stage. We have seen above that the model-theoretical content of 7.1.4 is the preservation theorem saying (forgetting the role of T for simplicity) that a sentence is preserved by (into) homomorphisms iff it is logically equivalent to a positive existential sentence. There is a natural generalization of this theorem for $L_{\omega_1\omega}$: in this, the positive existential sentences are those that are built up using (countable) $\bigvee$ and $\bigwedge$, and $\exists$. The reason why this cannot be directly applied is that models do not preserve the infinitary $\bigwedge$ (and if they are required to, they won't exist). The required generalization of 7.1.4 does follow from the model-theoretical case if the categories are Boolean: in this case $\bigwedge$ can be expressed in terms of $\bigvee$ and $\neg$: "$\bigwedge = \neg\bigvee\neg$", and as we have pointed out, models of categories preserve (Boolean) $\neg$. It is interesting that infinite infs (and other things) play a role in the following results although the models will not be assumed to preserve them. We have not been able to see exactly what is necessary of our assumptions concerning infs, etc., although there is a certain naturalness of the proofs suggesting that naturalness of the assumptions.

Definition 7.2.1 (c.f. 3.4.3). *Let κ be a regular cardinal. A κ-logical category $\mathcal{R}$ is a logical category in which the sup of any family F of subobjects of a given object such that* $\mathrm{card}(F) < \kappa$ *exists, and is stable under pullbacks.*

Let us denote by $L^g_{\kappa\omega}$ the coherent fragment of $L_{\infty\omega}$ consisting of those coherent formulas of $L_{\infty\omega}$ in which for each subformula of the form $\bigvee\Sigma$, Σ has cardinality $< \kappa$.

Proposition 7.2.2 *For a category $\mathcal{R}$ with finite $\varprojlim$, $\mathcal{R}$ is κ-logical if and only if the fragment $L^g_{\kappa\omega}$ is stable with respect to any interpretation $I : L \to \mathcal{R}$ (c.f.* 3.5.4*).*

PROOF. Obvious.

The right context in which infinitary logic, more precisely, the syntax of infinitary logic should be considered is that of admissible sets. Admissible sets are "partial set-theoretical universes": transitive sets satisfying certain weak axioms that amount to closure conditions for being able to perform certain set-theoretical operations of a recursive-constructive nature.

Below, we will establish our results in the context of admissible sets. However, the reader may choose to systematically ignore this level of generality and read our proofs as referring directly to $L^g_{\kappa\omega}$, with κ a regular cardinal. In doing so, for "$\mathcal{A}$-logical" or "weakly $\mathcal{A}$-logical" read "κ-logical" (7.2.1 above), for "$\mathcal{A}$-finite" read "having power $< \kappa$". Moreover, in this case ignore references to "$\mathcal{A}$-recursive", "$\mathcal{A}$-rec. en." and ignore any distinction between A and $\overline{A}$, I and $\overline{A}$, etc. There is only one place where the proof becomes actually simpler for the κ-logical case; we will indicate this below.

Nevertheless, there is a loss when one restricts attention to the κ-logical case. To wit, the only κ for which non-trivial κ-logical categories can be countable is $\kappa = \aleph_0$. On the other hand, there are many countable admissible sets that give rise to countable $\mathcal{A}$-logical categories with genuine infinitary sups. For these, the Rasiowa-Sikorski lemma provides ordinary **Set**-models as opposed to the general case when we only have Boolean-valued ones.

A point to emphasize here is that our results below depend strongly on the fact that our categories have "sufficiently many" sups among others. One way to make this hold is to require the existence of *all* sups of sets of cardinality less than a given κ. But this is a crude way that immediately excludes from the scope of the validity of the results all but a few trivial countable categories. There is a way to formulate the results in such a way that there are many countable cases where they hold but this requires a fine formulation of the requirement of "sufficiently many" sups, namely, the formulation with admissible sets.

For admissible sets, we refer to Barwise [1975] and Keisler [1971]. Notably, we have in mind admissible sets with urelements. Also, speaking about an admissible set $\mathcal{A}$, we actually mean a structure with some arbitrary relations $R_1, \ldots, R_\ell$ (besides $\in$ and the set of urelements) and accordingly, e.g. $\mathcal{A}$-recursively enumerable means definable by a Σ-formula using also the relations R_i besides $\in$.

Let $\mathcal{A}$ be an admissible set.

Definition 7.2.3 *Let $\mathcal{R}$ be a category. An* $\mathcal{A}$-recursive representation *of $\mathcal{R}$ consists of $\mathcal{A}$-recursive sets $|\mathcal{R}|_{\text{ob}}$, $|\mathcal{R}|_{\text{morph}}$; $\mathcal{A}$-recursive functions* $\text{Dom} : |\mathcal{R}|_{\text{morph}} \to |\mathcal{R}|_{\text{ob}}$, $\text{Codom} : |\mathcal{R}|_{\text{morph}} \to |\mathcal{R}|_{\text{ob}}$, $\text{Id} : |\mathcal{R}|_{\text{ob}} \to |\mathcal{R}|_{\text{morph}}$ *and* $\text{Comp} : (|\mathcal{R}|_{\text{morph}})^2 \to |\mathcal{R}|_{\text{morph}}$ *(or defined at least for the right pairs for "morphisms" $\in |\mathcal{R}|_{\text{morph}}$) and surjective maps* $\overline{(\cdot)} : |\mathcal{R}|_{\text{ob}} \to \text{Ob}(\mathcal{R})$, $\overline{(\cdot)} : |\mathcal{R}|_{\text{morph}} \to \text{morph}(\mathcal{R})$ *which carry the above functions* Dom, Codom, Id *and* Comp *into the domain, codomain, etc. functions in the category: e.g.* $\overline{\text{Id}(r)} = \text{Id}_r$ *if $r \in |\mathcal{R}|_{\text{ob}}$, etc. We furthermore require that the relation*

$$\{\langle r_1, r_1 \langle \in (|\mathcal{R}|_{\text{morph}})^2 : \overline{r_1} = \overline{r_2} \text{ in } \mathcal{R}\}$$

is $\mathcal{A}$-r.e.

An $\mathcal{A}$-recursively presented category *is a category together with an $\mathcal{A}$-recursive presentation of it.*

For an $\mathcal{A}$-recursively presented category, when we speak of $\mathcal{A}$-finite or $\mathcal{A}$-recursive, etc. families of objects or morphisms, we mean a family of names $\subset |\mathcal{R}|_{\text{ob}} \cup |\mathcal{R}|_{\text{morph}}$ which is $\mathcal{A}$-finite, or $\mathcal{A}$-recursive, etc. This involves a certain measure of abuse of language since at the same time, we will talk about, say, the sum etc. of the family considered; here of course the objects *denoted* by the names are understood.

For an $\mathcal{A}$-recursively presented category $\mathcal{R}$, the language $L_{\mathcal{R}}$ will be redefined such that the sorts of $L_{\mathcal{R}}$ are the elements of $|\mathcal{R}|_{\text{ob}}$, and the operation symbols of $L_{\mathcal{R}}$ are the elements of $|\mathcal{R}|_{\text{morph}}$, the Dom and Codom functions giving the sorting of operation symbols. Of course, the map $r \mapsto \overline{r}$ takes the place of the identical interpretation. Notice that $L_{\mathcal{R}}$ is an $\mathcal{A}$-recursive language.

Definition 7.2.4 *An $\mathcal{A}$-logical category is an $\mathcal{A}$-recursively presented category $\mathcal{R}$ such that*

(i) *$\mathcal{R}$ has finite left limits which can be computed $\mathcal{A}$-recursively: there is an $\mathcal{A}$-recursive function F that, applied to a finite diagram consisting of names in $|\mathcal{R}|_{\text{ob}} \cup |\mathcal{R}|_{\text{morph}}$ of objects and morphisms in $\mathcal{R}$, gives the names of a left limit diagram of the given diagram. E.g., if $R_1, R_2 \in |\mathcal{R}|_{\text{ob}}$, then $F(\langle R_1, R_2 \rangle)$ is a 5-tuple $\langle R_1, R_2, R, \pi_1, \pi_2 \rangle$ such that the diagram*

$$\begin{tikzcd} \overline{R_1} & & \overline{R_2} \\ & \overline{R} \arrow[ul, "\overline{\pi_1}"] \arrow[ur, "\overline{\pi_2}"'] & \end{tikzcd}$$

is a product of the two objects $\overline{R_1}$ and $\overline{R_2}$; similarly for other finite left limits.

(ii) *$\mathcal{R}$ has $\mathcal{A}$-finite sups that can be computed $\mathcal{A}$-recursively: there is an $\mathcal{A}$-recursive function G such that if $a \underset{\text{df}}{=} \langle R_i \xrightarrow{f_i} R : i \in I \rangle$ is an $\mathcal{A}$-finite family ($R_i, R \in |\mathcal{R}|_{\text{ob}}, \text{Dom}(f_i) = R_i, \text{Codom}(f_i) = R$), $\overline{f_i}$ is a monomorphism for $i \in I$, then $G(a) = \langle R', f \rangle$ such that $\overline{R'} \xhookrightarrow{\overline{f}} \overline{R}$ is the sup of the family $\overline{R_i} \xhookrightarrow{\overline{f_i}} \overline{R}$ of subobjects of $\overline{R}$. Moreover, the $\mathcal{A}$-finite sups are stable under pullbacks.*

(iii) *$\mathcal{R}$ has stable images that can be computed $\mathcal{A}$-recursively (this can be made precise similarly to (ii)).*

Let $L^g_{\mathcal{A}}$ be the coherent $\mathcal{A}$-fragment of $L_{\infty\omega}$; the formulas of $L^g_{\mathcal{A}}$ are those of $L^g_{\infty\omega}$ that are elements of $L_{\mathcal{A}}$. Equivalently, $L^g_{\mathcal{A}}$ is the smallest set of formulas of $L_{\infty\omega}$ belonging to $\mathcal{A}$ closed under finite $\wedge$, $\exists x$ and $\mathcal{A}$-finite disjunction: if $\Sigma \subset L^g_{\mathcal{A}}$ is an $\mathcal{A}$-finite set of formulas having only finitely many free variables altogether, then $\bigvee\Sigma \in L^g_{\mathcal{A}}$.

Proposition 7.2.5 *Let $I : L \to \mathcal{R}$ be an interpretation of the language L in $\mathcal{R}$, let L be an $\mathcal{A}$-recursive language, $\mathcal{R}$ $\mathcal{A}$-recursively presented and let I be induced by an interpretation $I' : L \to L_{\mathcal{R}}$ of languages ($I(\sigma) = \overline{I'(r)}$) where I' is $\mathcal{A}$-recursive.*

If $\mathcal{R}$ is $\mathcal{A}$-logical, then

(a) *every formula of $L^g_{\mathcal{A}}$ is "adequately interpreted by I" i.e., $L^g_{\mathcal{A}}$ is stable with respect to I, c.f. terminology before* 3.5.4. *In fact there is an $\mathcal{A}$-recursive function H such that, if ϕ is a formula in $L^g_{\mathcal{A}}$ with free variables among $x_1, \dots, x_n$ of sorts $R_1, \dots, R_n$, respectively, then $H(\langle\phi; \langle x_1, \dots, x_n\rangle\rangle)$ is $\langle R', f\rangle$ where $\overline{R'} \xrightarrow{\overline{f}} \overline{R_1} \times \dots \times \overline{R_n}$ is a monomorphism and in fact, $I_{\vec{x}}(\phi) \simeq \overline{R'} \hookrightarrow \overline{R_1} \times \dots \times \overline{R_n}$, and*

(b) *the predicate $I \models (\cdot)$ for sequents σ of $L^g_{\mathcal{A}}$ is $\mathcal{A}$-r.e.*

The proof of (a) is an easy application of some general principles on admissible sets (definition by recursion); we will not give details.

For the proof of (b), observe that for subobjects $\overline{A} \overset{\overline{f}}{\hookrightarrow} \overline{B}$, $\overline{C} \overset{\overline{g}}{\hookrightarrow} \overline{B}$, $\overline{A} \le \overline{C}$ is equivalent to the existence of $h \in |\mathcal{R}|_{\text{morph}}$ such that $\overline{g}\overline{h} = \overline{f}$. Given a sequent σ, $I \models \sigma$ is equivalent to $\overline{A} \le \overline{C}$ for some recursively computed A, B, C, f and g (such that f and g are monomorphisms), hence $I \models \sigma \;\Leftrightarrow\; \exists h[h \in |\mathcal{R}|_{\text{morph}}$ and $\overline{g}\overline{h} = \overline{f}]$ which is $\mathcal{A}$-r.e. by our assumptions.

Example 7.2.6 This is how to subsume κ-logical categories under the notion of 7.2.4. Let $\mathcal{R}$ be a κ-logical category. Consider $|\mathcal{R}| = \text{Ob}(\mathcal{R}) \cup \text{morph}(\mathcal{R})$ as a set of urelements and let $\mathcal{A} = H_{|\mathcal{R}|}(\kappa)$ be the set of sets with support $\subset |\mathcal{R}|$ with are hereditarily of power less than κ. $\langle\mathcal{A}, \in\rangle$ together with any predicates on $\mathcal{A}$ will be an admissible set. Add finitely many predicates so that for the resulting $\mathcal{A}$, the conditions of 7.2.4 will be trivially satisfied. For the presentation, take the identical one: $|\mathcal{R}|_{\text{ob}} = \text{Ob}(\mathcal{R})$, $|\mathcal{R}|_{\text{morph}} = \text{morph}(\mathcal{R})$, $\overline{r} = r$. By the axiom of choice, there is a function F assigning left limits to finite diagrams. Adjoin $F \subset \mathcal{A} \times \mathcal{A}$ to $\langle\mathcal{A}, \in\rangle$; then F will be trivially $\mathcal{A}$-recursive in the resulting $\mathcal{A}$. Similarly, adjoin a function to $\langle\mathcal{A}, \in\rangle$ to make (ii) and (iii) of 7.4.2 true. – Observe that the $\mathcal{A}$-finite subsets of $|\mathcal{R}|$ are exactly those of power $< \kappa$; hence we will have that $\mathcal{R}$ is $\mathcal{A}$-logical.

Summarized, the admissible set to be considered with κ-logic is the set of sets hereditarily of power $< \kappa$. With suitably adding (arbitrary) predicates, we can arrange that any prescribed predicate becomes $\mathcal{A}$-recursive.

It turns out that for the purposes of our results below, the conditions of 7.2.4 can, at least apparently, be considerably relaxed. We have not investigated yet how much weaker the notion to be considered next actually is.

Definition 7.2.7 *A category $\mathcal{R}$ is called* weakly $\mathcal{A}$-logical *(w-$\mathcal{A}$-logical) if it is $\mathcal{A}$-recursively presented, and satisfies the following two conditions* (i) *and* (ii)*.*

(i) *$\mathcal{R}$ is logical in the usual sense and for the* finitary *fragment* $(L_{\mathcal{R}})^g_{\omega\omega}$, *the truth of sequents σ of this fragment in the "identical" interpretation $r \mapsto \overline{r}$, called $\mathcal{R}$, is $\mathcal{A}$-r.e.: $\{\sigma$ sequent of $(L_{\mathcal{R}})^g_{\omega\omega} : \mathcal{R} \models \sigma\}$ is $\mathcal{A}$-r.e.*

(ii) *For any $\mathcal{A}$-finite family $\langle A_i \xrightarrow{f_i} A : i \in I\rangle \in \mathcal{A}$ of subobjects the sup $B \xrightarrow{f} A$* exists *and is stable under pullbacks.*

Remark *So, the main difference is that recursive computability of sups is not required.*

Definition 7.2.8 *An* $\mathcal{A}$-logical functor $I : \mathcal{R} \to \mathcal{S}$ *between w-*$\mathcal{A}$*-logical categories* $\mathcal{R}$ *and* $\mathcal{S}$ *is given by an* $\mathcal{A}$*-recursive map* $I : L_{\mathcal{R}} = |\mathcal{R}|_{\text{ob}} \cup |\mathcal{R}|_{\text{morph}} \to L_{\mathcal{S}} = |\mathcal{S}|_{\text{ob}} \cup |\mathcal{S}|_{\text{morph}}$ *such that* I *induces the map* $\overline{I} : \overline{r} \to \overline{I(r)}$ *and* $\overline{I}$ *is an* $\mathcal{A}$*-logical functor* $\mathcal{R} \to \mathcal{S}$ *i.e.,* $\overline{I}$ *is left exact and preserves* $\mathcal{A}$*-finite sups.*

Definition 7.2.9 (c.f. 3.4.5). *A* κ *logical functor* $I : \mathcal{R} \to \mathcal{S}$ *between* κ*-logical categories* $\mathcal{R}$ *and* $\mathcal{S}$ *is a logical functor that preserves sups of powers less than* κ*.*

Example 7.2.6 (continued). Let $\mathcal{R}$, $\mathcal{S}$ and $I\mathcal{R} \to \mathcal{S}$ be κ-logical. We consider $\mathcal{A} = H_{|\mathcal{R}| \cup |\mathcal{S}|}(\kappa)$ and adjoin some predicates including $I \subset \mathcal{A} \times \mathcal{A}$ itself, and we obtain that for the resulting $\mathcal{A}$, $I : \mathcal{R} \to \mathcal{S}$ is an $\mathcal{A}$-logical functor.

§3 Infintary generalizations

In what follows we assume that $\mathcal{R}$ and $\mathcal{S}$ are weakly $\mathcal{A}$-logical categories, with $\mathcal{A}$ a fixed admissible set. We assume that $I : \mathcal{R} \to \mathcal{S}$ is an $\mathcal{A}$-logical functor.

Remark *For* $I : \mathcal{R} \to \mathcal{S}$ $\mathcal{A}$*-logical, the actual functor is what was denoted by* $\overline{I}$ *above. Usually, we will not make a notational distinction between* I *and* $\overline{I}$*; we will denote* $\overline{I}$ *by* I *as well.*

Let $\mathcal{B}$ be any complete Boolean algebra. The *$\mathcal{B}$-valued $\mathcal{A}$-models of* $\mathcal{R}$ are the left exact and $\mathcal{A}$-continuous functors

$$M : \mathcal{R} \to \mathrm{Sh}_{\mathcal{B}};$$

$\mathcal{A}$-continuous means that M preserves $\mathcal{A}$-finite sups.

Notice that a $\mathcal{B}$-valued $\mathcal{A}$-model is just like an $\mathcal{A}$-logical functor except that there is no recursiveness condition on M. In particular, a *$\mathcal{B}$-valued κ-model* is just a κ-logical functor

$$M : \mathcal{R} \to \mathrm{Sh}_{\mathcal{B}}\,.$$

The category $\mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{R})$ of $\mathcal{B}$-valued $\mathcal{A}$-models of $\mathcal{R}$ is the full subcategory of the functor category $(\mathrm{Sh}_{\mathcal{B}})^{\mathcal{R}}$ whose objects are the $\mathcal{B}$-valued $\mathcal{A}$-models of $\mathcal{R}$. Given $\mathcal{A}$-logical $I : \mathcal{R} \to \mathcal{S}$, we have

$$I^* = I^*_{\mathcal{B}} : \mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{S}) \to \mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{R})$$

defined by composition: for $N \in \mathrm{Ob}(\mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{S}))$, $I^*(N)$ is defined by

$$I^*(N)(R) = N(I(R))$$

and

$$I^*(N)(R \xrightarrow{f} R') = N(I(R)) \xrightarrow{N(I(f))} N(I(R')).$$

It is easy to see that, since I is $\mathcal{A}$-logical (so, in particular, is induced by an $\mathcal{A}$-recursive map) and N is an $\mathcal{A}$-model, $I^*(N)$ also will be an $\mathcal{A}$-model.

With $\mathcal{B} = 2$ the two-valued algebra, we obtain the category of *$\mathcal{A}$-models*, $M : \mathcal{R} \to$ **Set**, $\mathrm{Mod}_{\mathcal{A}}(\mathcal{R})$.

Let $\mathcal{R}$ be a w-$\mathcal{A}$-logical category. Consider the language $L_{\mathcal{R}}$ associated with $\mathcal{R}$ and consider $T_{\mathcal{R}}$, the theory in $(L_{\mathcal{R}})^g_{\mathcal{A}}$ representing the $\mathcal{A}$-logical structure of $\mathcal{R}$:

$$T_{\mathcal{R}} = T^{(\text{ex})}_{\mathcal{R}} \cup T$$

where $T^{(\text{ex})}_{\mathcal{R}}$ is the internal theory of $\mathcal{R}$ as a finitary logical category, i.e. the theory in the coherent finitary logic over $L_{\mathcal{R}}$ consisting of the axioms corresponding to the left limit diagrams and images in $\mathcal{R}$, (note however that the symbols to be used are the names in $|\mathcal{R}|_{\text{ob}} \cup |\mathcal{R}|_{\text{morph}}$ instead of objects and morphisms themselves) and T is the set of all "axioms of sups" (c.f. item 8 in 2.4.5) corresponding to sup diagrams:

$$
\begin{tikzcd}
\overline{A_i} \arrow[r, hook, "\overline{f_i}"] \arrow[d, phantom, "\vdots"] & \overline{A} \\
 & \bigvee_i \overline{A_i} = \overline{B} \arrow[u, hook, "\overline{f_i}"]
\end{tikzcd}
$$

with $\langle A_i \xrightarrow{f_i} A : i \in I \rangle$ $\mathcal{A}$-finite.

We let $F_{\mathcal{R}}$ (or $F^g_{\mathcal{R}}$, with g signifying 'coherent') be the smallest fragment such that $T_{\mathcal{R}}$ is a theory in $F_{\mathcal{R}}$: the formulas of $F_{\mathcal{R}}$ are the formulas in sequents in $T_{\mathcal{R}}$, all the subformulas of those, and all substitution instances of the previous ones. The main property of formulas in $F_{\mathcal{R}}$ is that infinite disjunction is applied only to finitary formulas: if $\Sigma \in F_{\mathcal{R}}$, then every element of Σ is a finitary formula. This property, together with the fact that $\mathcal{R}$ is weakly-$\mathcal{A}$-logical, is sufficient to ensure that the formula is *adequately interpreted* by the "identical" interpretation $\mathcal{R} : r \to \overline{r}$, i.e., that the fragment $F_{\mathcal{R}}$ is stable in $\mathcal{R}$ (c.f. Chapter 3, Section 2).

Recall that by 3.2.8 we have

$(*)$ for any sequent σ of $F^g_{\mathcal{R}}$,

$$T \vdash \sigma \;\Leftrightarrow\; \mathcal{R} \models \sigma.$$

Hence, by the completeness theorem (5.1.1, 5.1.2) saying that

$$T \vdash \sigma \;\Leftrightarrow\; T \overset{\mathrm{b}}{\models} \sigma.$$

we have, similarly to 7.1.1,

Corollary 7.3.1 *For subobjects R_1 and R_2 of R, $R_1 \leq R_2$ iff $M(R_1) \leq M(R_2)$ for all complete Boolean algebras $\mathcal{B}$ and all $\mathcal{B}$-valued $\mathcal{A}$-models M of $\mathcal{R}$.*

In case $\mathcal{A}$ is countable *(hence $F_{\mathcal{R}}$ is countable), we have $R_1 \leq R_2$ iff $M(R_1) \leq M(R_2)$ for all $M \in \mathrm{Ob}(\mathrm{Mod}_{\mathcal{A}}(\mathcal{R}))$.*

Proof. Similar to that of 7.1.1.

Theorem 7.3.2 *With $I : \mathcal{R} \to \mathcal{S}$ $\mathcal{A}$-logical, $\mathcal{R}$, $\mathcal{S}$ w-$\mathcal{A}$-logical, assume that $I^*_{\mathcal{B}} : \mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{S}) \to \mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{R})$ is surjective on objects, for any complete Boolean algebra $\mathcal{B}$ (c.f.* 7.1.2*). Then I is conservative (c.f.* 7.1.2 *again). In particular, I is faithful.*

In case $\mathcal{A}$ is countable*, it is enough to assume the hypothesis for the 2-element Boolean algebra, i.e., that $I^* : \mathrm{Mod}_{\mathcal{A}}(\mathcal{S}) \to \mathrm{Mod}_{\mathcal{A}}(\mathcal{R})$ is surjective on objects.*

Proof. Similar to that of 7.1.2.

Theorem 7.3.3 *Assume that, in addition, $\mathcal{R}$ and $\mathcal{S}$ are Boolean (c.f.* 7.1.3*) and I^* is full with respect to isomorphisms, for any complete Boolean algebra $\mathcal{B}$. Then I is full with respect to subobjects. If $\mathcal{A}$ is countable, it is enough to assume that $I^* : \mathrm{Mod}_{\mathcal{A}}(\mathcal{S}) \to \mathrm{Mod}_{\mathcal{A}}(\mathcal{R})$ is full with respect to isomorphisms.*

Proof. In case $\mathcal{R}$ and $\mathcal{S}$ are in fact $\mathcal{A}$-logical, the proof is similar to that of 7.1.3, on the basis of Beth's theorem for logics on admissible sets. The application of the theorem to logic depends on the fact that every formula of the full admissible logic $L_{\mathcal{A}}$ is interpretable; this we know by 7.2.5.

In case $\mathcal{R}$ and $\mathcal{S}$ are only assumed to be w-$\mathcal{A}$-logical, the proof would be similar to the two proofs we present below; since we are mainly interested in the $\mathcal{A}$-logical case, we will not give this proof here.

Remark *For* 7.3.2*, the admissibility of* $\mathcal{A}$ *does not play any essential role; in* 7.3.3 *however, it is an essential assumption.*

For our proof of the result that corresponds to 7.1.4, we need an additional assumption. We don't know if 7.3.5 is true without this assumption though we suspect it isn't in general.

Assumption 7.3.4 (i) $\mathcal{R}$ has $\mathcal{A}$-finite infs: for any $\mathcal{A}$-finite family $\langle \overline{R_i} \xhookrightarrow{\overline{f_i}} \overline{R} : i \in I\rangle$ of subobjects, the inf $\bigwedge(\overline{R_i} \xhookrightarrow{\overline{f_i}} \overline{R})$ exists in $\mathcal{R}$.

(ii) I preserves $\mathcal{A}$-finite infs: $I(\bigwedge_i \overline{R_i}) \xhookrightarrow{I(\overline{f})} I(\overline{R})$ is the inf of the family

$$\langle I(\overline{R_i}) \xhookrightarrow{I(\overline{f_i})} I(\overline{R}) : i \in I\rangle,$$

for $\langle R_i \xhookrightarrow{f_i} R : i \in I\rangle$ $\mathcal{A}$-finite, and for $\bigwedge_i \overline{R_i} \xhookrightarrow{\overline{f}} \overline{R}$ the inf of $\langle R_i \xhookrightarrow{f_i} R\rangle$.

Remark *The corresponding (via* 7.2.6*) conditions for the κ-logical case have to do with info's of families of subobjects of power* $< \kappa$. Hence*, for* $\kappa = \omega$*, the condition is part of the left-exactness condition and hence automatically satisfied. This is the reason why the finitary* 7.1.4 *will follow from* 7.3.5*.*

Also, let us note that in case $\mathcal{R}$ is Boolean and $\mathcal{A}$-logical, moreover I is an $\mathcal{A}$-logical functor, then $\mathcal{R}$ and I satisfy 7.3.4. Namely, if $\langle \overline{R_i} \xhookrightarrow{\overline{f_i}} \overline{R} : i \in I\rangle$ is an $\mathcal{A}$-finite family of subobjects in $\mathcal{R}$, then $\bigwedge_i \overline{R_i}$ exists and it is $\neg\bigvee\neg\overline{R_i}$ where $\neg$ is Boolean complement (and, as it is easy to see, $\neg(\cdot)$ can be recursively computed) and I will preserve $\neg$ and $\bigvee$.

Theorem 7.3.5 *Assume* $\mathcal{R}$*,* $\mathcal{S}$ *and* I *are as before and assume* 7.3.4*. Assume that* (i) $I^*_{\mathcal{B}}$ *is full for any non-trivial complete Boolean algebra* $\mathcal{B}$*,* or (ii) *that* $\mathcal{A}$ *is countable and* $I^* : \mathrm{Mod}_{\mathcal{A}}(\mathcal{S}) \to \mathrm{Mod}_{\mathcal{A}}(\mathcal{R})$ *is full,* or (iii) *that each set in* $\mathcal{A}$ *is finite (i.e., the logic* $L_{\mathcal{A}}$ *is finitary) and* I^* *as in* (ii) *is full.* Then I *is full with respect to subobjects. If in addition,* I^* *is surjective on objects, then* I *is full.*

Proof of 7.3.5 We will work with the fragments $F_{\mathcal{R}}$ and $F_{\mathcal{S}}$ defined above, and apply consistency properties as defined in Chapter 5, Section 2. Assume that $S \xhookrightarrow{f} I(R)$ is a subobject of $I(R)$; if $\underset{\sim}{S}(a)$ is $\exists s(fs = a)$ (a is a variable of sort $I(R)$), then $S \xhookrightarrow{f} I(R) = (\underset{\sim}{S}(a))^{\mathcal{S}}$. Assume, by reductio ad absurdum, that S is not in the essential image of I. We will construct a non-trivial complete Boolean algebra $\mathcal{B}$, $\mathcal{B}$-valued models $N_1 : \mathcal{S} \to \mathrm{Sh}_{\mathcal{B}}$ and $N_2 : \mathcal{S} \to \mathrm{Sh}_{\mathcal{B}}$ and a natural transformation (homomorphism) $F : I^*(N_1) \to I^*(N_2)$ such that $F_{\mathcal{R}}$ does not preserve $\underset{\sim}{S}$: *it is not the case that*

$$\|y = F_{I(R)}(x)\| \cdot \|\underset{\sim}{S}(x)\|_{N_1} \leq_{\mathcal{B}} \|\underset{\sim}{S}(y)\|_{N_2}$$

for all $x \in N_1(I(R))$. This will suffice, since, as it is easy to see similarly to the final part of the proof of 7.1.4, the assumption on fullness implies that $\underset{\sim}{S}(a)$ *is* preserved by all $F : I^*(N_1) \to I^*(N_2)$.

We fix the free variables a_0 and $\underline{a}_0$ of sort $I(R)$. Let us define the partial ordering $(P, \leq)$ as follows. We put P to be the set of triples

$$p = \langle \Delta_{N_1}, \Delta_{N_2}, \Delta_F \rangle$$

such that: (i)-(iii) below hold: (i) $\Delta_{N_1}, \Delta_{N_2}$ are finite sets of formulas of the fragment $F_{\mathcal{S}}$,

(ii) Δ_F is a finite sequence

$$\langle \langle a_i, \underline{a}_i \rangle : i < n \rangle$$

of pairs $\langle a, \underline{a} \rangle$ of terms. For any $i < n$, the sort of a_i equals the sort of $\underline{a}_i$ and equals $I(R_i)$ for some $R_i \in |\mathcal{R}|_{\mathrm{ob}}$. a_0 and $\underline{a}_0$ are the variables fixed above. Moreover, every free variable occurring in some a_i $(i < n)$ occurs in $\Delta_{N_1} \cup \{\underset{\sim}{S}(a_0)\}$.

(iii) There does not exist a formula θ such that

$(*)$ θ is in the finitary logic $(L_{\mathcal{R}})^g_{\omega\omega}$, having n distinct free variables $x_0, \ldots, x_{n-1}$; and

$$\begin{aligned} T_{\mathcal{S}} &\vdash \Delta_{N_1}, \underset{\sim}{S}(a_0) \Rightarrow I(\theta)(\vec{a}), \\ T_{\mathcal{S}} &\vdash \Delta_{N_1}, I(\theta)(\vec{\underline{a}}) \Rightarrow \underset{\sim}{S}(\underline{a}_0), \end{aligned}$$

where $I(\theta)(\vec{a})$ is obtained from $I(\theta)$ by substituting $I(x_i)$ by a_i for $i < n$; similarly for $I(\theta)(\vec{\underline{a}})$.

(We will say that θ *blocks* p (*from being an element of* P if all the conditions after $(*)$ are satisfied. Hence (iii) says that there is no θ blocking p).

Remark *Intuitively, thinking of two-valued models, Δ_{N_1} is an approximation of the 'full description' of N_1 (set of all sentences with names for elements in N_1 which are true in N_1), and similarly for Δ_{N_2}. Δ_F is an approximation of the homomorphism, with a_i being mapped onto $\underline{a_i}$ by the homomorphism.*

Notice that $p = \langle \varnothing, \varnothing, \langle a_0, \underline{a}_0 \rangle \rangle$ belongs to P; if, on the contrary, there existed a θ such that

$$\begin{aligned} T_{\mathcal{S}} &\vdash \underset{\sim}{S}(a_0) \Rightarrow (I(\theta))(a_0) \\ T_{\mathcal{S}} &\vdash I(\theta)(\underline{a}_0) \Rightarrow \underset{\sim}{S}(\underline{a}_0) \end{aligned}$$

then we would clearly have

$$S = (\underset{\sim}{S}(a_0))^{\mathcal{S}} = (I(\theta))^{\mathcal{S}} = I(\theta^{\mathcal{R}});$$

contrary to the indirect hypothesis of the proof.

We define the partial ordering $\leq$ of P by component wise inclusion:

$$p \leq q \Leftrightarrow \Delta^{(p)}_{N_1} \supset \Delta^{(q)}_{N_1} \text{ and } \Delta^{(p)}_{N_2} \supset \Delta^{(q)}_{N_2} \text{ and } \Delta^{(p)}_F \supset \Delta^{(q)}_F;$$

here we use the notation $p = \langle \Delta^{(p)}_{N_1}, \Delta^{(p)}_{N_1}, \Delta^{(p)}_F \rangle$.

We will derive two consistency properties from $(P, \leq)$ for the construction of the two models N_1 and N_2. Put $f_1(p) = \Delta^{(p)}_{N_1}$, $\mathrm{Var}_1(p) = \mathrm{Var}(\Delta^{(p)}_{N_1}) \cup \{a_0\}$ ($\mathrm{Var}(\Delta)$ is the set of free variables in Δ); and $f_2(p) = \Delta^{(p)}_{N_2}$ and $\mathrm{Var}_2(p) = \mathrm{Var}(\Delta^{(p)}_{N_2}) \cup \bigcup_{i<n} \mathrm{Var}(a_i)$ (here a_i come from Δ_F as above). The two consistency properties are

$$\begin{aligned} \Gamma_1 &= (P, \leq, f_1, \mathrm{Var}_1) \\ \text{and } \Gamma_2 &= (P, \leq, f_2, \mathrm{Var}_2). \end{aligned}$$

Lemma 7.3.6 Γ_1 *and* Γ_2 *are* $T_{\mathcal{S}}$*-consistency properties.*

1. Γ_1 *is a consistency property.*

Property (iv) (in the definition of consistency properties) is obvious. Each of the rest of the properties (v)-(xiv) calls for finding a suitable $q \le p$, for a given $p \in P$. In each case q will be found so that $\Delta^{(q)}_{N_2} = \Delta^{(p)}_{N_2}$ and $\Delta^{(q)}_F = \Delta^{(p)}_F$, in other words, only $\Delta^{(q)}_{N_1}$ changes. Except for the case (viii), $\Delta' = \Delta^{(q)}_{N_1}$ (and thus q) will be defined outright (we put $\Delta = \Delta^{(q)}_{N_1}$: for

$$\begin{aligned}
\text{(v):}\quad & \Delta' = \Delta \cup \{\phi\},\\
\text{(vi):}\quad & \Delta' = \Delta \cup \{\textstyle\bigwedge\Sigma\},\\
\text{(vii):}\quad & \Delta' = \Delta \cup \{\textstyle\bigvee\Sigma\},\\
\text{(ix):}\quad & \Delta' = \Delta \cup \{\exists x\, \phi(x/v)\},\\
\text{(x):}\quad & \Delta' = \Delta \cup \{\phi(u/v)\} \text{ where } u \text{ is a free variable of the sort of}\\
& v \text{ such that } u \text{ does not occur in } \Delta,\\
\text{(xi):}\quad & \Delta' = \Delta \cup \{\psi'(t_1, \ldots, t_n)\},\\
\text{(xii):}\quad & \Delta' = \Delta \cup \{t \approx t\},\\
\text{(xiii):}\quad & \Delta' = \Delta \cup \{t_1 \approx t_2\},\\
\text{(xiv):}\quad & \Delta' = \Delta \cup \{\phi(t_2)\}.
\end{aligned}$$

(We will handle (viii) separately below.)

We have to show that q thus defined is in fact in P; then the fact that $q \le p$ will be clear by the definition of $\le$. (i) and (ii) in the definition of being an element of P are obviously satisfied for q. Now we assume that (iii) in that definition is not satisfied, i.e. that some θ blocks q from being an element of P. The way we set up things now ensures that in each of the above cases (with (viii) excluded of course so far), this will imply that the same θ will block p, contrary to the assumption $p \in P$. This fact, for each of the above cases, will be the consequence of the presence of the *corresponding* rule of inference of the one-sided system defining the notion $T_{\mathcal{S}} \vdash (\cdot)$. For example, for condition (vii) we have the assumption in (vii) that the free variables in $\bigvee\Sigma$ belong to $\mathrm{Var}_1(p) = \mathrm{Var}(\Delta) \cup \{a_0\}$. So, from the fact that θ blocks q, i.e. $T_{\mathcal{S}} \vdash \Delta, \bigvee\Sigma, \underset{\sim}{S}(a_0) \Rightarrow I(\phi)(\vec{a})$ it follows that

$$T_{\mathcal{S}} \vdash \Delta, \underset{\sim}{S}(a_0) \Rightarrow I(\phi)(\vec{a})$$

by the rule $R\bigvee_1$; note that $\phi \in \Sigma$ and $\phi \in \Delta$!

There is nothing to check for the other condition in $(*)$ (involving Δ_{N_2}) since there is no change in this respect in q as compared to p. We conclude that θ blocks p, as promised.

It remains to handle the verification of condition (viii). We assume that the claim of (viii) for Γ_1 does not hold. Let $\phi \in \Sigma$ and define q_ϕ to agree with p except in $\Delta^{(\cdot)}_{N_1}$ and we put $\Delta^{(q_\phi)}_{N_1} = \Delta \cup \{\phi\}$. We have that for every $\phi \in \sigma$ there is some θ_ϕ blocking q_ϕ from being a member of P (otherwise $q_\phi \le p$ would verify (viii)). This fact means that (putting $T = T_{\mathcal{S}}$)

$$T \vdash \Delta, \phi, \underset{\sim}{S}(a_0) \Rightarrow I(\theta_\phi)(\vec{a})$$

and

$$T \vdash \Delta_{N_2}, I(\theta_\phi)(\underline{\vec{a}}) \Rightarrow \underset{\sim}{S}(a_0)$$

for any $\phi \in \Sigma$; here $\vec{a}$ and $\underline{\vec{a}}$ refer to Δ_F as before. Notice that $T \vdash \sigma$ for any sequent of the fragment $F_{\mathcal{S}}$ is equivalent to $\mathcal{S} \models \sigma$.

The next thing to show would naturally be that the θ_ϕ can be chosen so that the family $\langle\theta_\phi : \phi \in \Sigma\rangle$ is $\mathcal{A}$-finite; this would make sure that $\bigvee_{\phi\in\Sigma}\theta_\phi$ will be interpretable in $\mathcal{R}$. In the κ-logical case (κ regular infinite cardinal), there is no problem: since $\mathrm{card}(\sigma) < \kappa$, it follows that $\mathrm{card}\{\theta_\phi; \phi \in \Sigma\} < \kappa$, hence $\bigvee_{\phi\in\Sigma}\theta_\phi$ is interpretable (and $\{\theta_\phi : \phi \in \Sigma\}$ is $\mathcal{A}$-finite). In the general case, we have to use the admissibility of $\mathcal{A}$ and we have to do something a bit more complicated. Of course, the argument is essentially as in similar proofs in Makkai [1969] and Keisler [1971]. Nevertheless, we will give the details of the proof; this is the only place where understanding admissible sets is essential.

Consider the predicate $P(\cdot,\cdot)$ of two variables on the set $\mathcal{A}$: $P(\phi,\theta) \underset{\mathrm{df}}{\Leftrightarrow} \phi$ belongs to Σ, θ is a formula of the finitary language $(L_{\mathcal{S}})^g_{\omega\omega}$ with free variables in the fixed set $\mathrm{Var}(\Delta_F^{(p)})$, and

$$S \models \Delta, \underset{\sim}{S}(a_0), \phi \Rightarrow I(\theta)(\vec{a})$$

and

$$S \models \Delta_{N_2}, I(\theta)(\underline{\vec{a}}) \Rightarrow \underset{\sim}{S}(a_0).$$

Notice that each element of Σ is finitary (see above the discussion of the fragment $F_{\mathcal{S}}$). The sequents displayed may contain some infinitary formulas but they are fixed and can be replaced by some finitary ones which are equivalent to them in $\mathcal{S}$. (If, e.g. $\bigvee_{i\in I}\psi_i$ occurs in Δ, with free variables exactly $u_1, \dots, u_k$ then we look at $S = (\bigvee_{i\in I}\psi_i)^{\mathcal{S}} \overset{\overline{f}}{\hookrightarrow} \overline{U_1} \times \cdots \times \overline{U_k} = \overline{U}$ and projections $\overline{\pi_i} : \overline{U} \to \overline{U_i}$ and take the formula $\psi(u_1, \dots, u_k)$

$$\exists s\, \exists u\, [\textstyle\bigwedge_{i=1}^k \pi_i(u) = u_i \wedge f(s) = u];$$

then $(\bigvee\psi_i)^{\mathcal{S}} = (\psi)^{\mathcal{S}}$, ψ has the same free variables as ψ_i, and we can thus replace ϕ_i by ψ in the sequent without changing truth in $\mathcal{S}$.) The variable formulas ϕ and $I(\theta)$ are finitary. By the definition of "weak-$\mathcal{A}$-logical" and the fact that I is $\mathcal{A}$-recursive, it follows that the predicate P is $\mathcal{A}$-r.e. ($\underset{\sim}{\Sigma}$ on $\mathcal{A}$).

Now we have that

$$\mathcal{A} \models \forall\phi \in \Sigma\ \exists\theta\ P(\phi,\theta).$$

By Σ-collection, there is a transitive $w \in \mathcal{A}$ such that

$$\mathcal{A} \models \forall\phi \in \Sigma\ \exists\theta \in w\ P^{(w)}(\phi,\theta); \tag{1}$$

here we have used P to denote a Σ_1-formula of the language $\mathcal{A}$ defining the predicate P on $\mathcal{A}$ and $P^{(w)}$ to denote its relativization to w.

Put $\Theta_\phi = \{\theta \in w : \mathcal{A} \models P^{(w)}(\phi,\theta)\}$; by Δ-comprehension, $\Theta_\phi \in \mathcal{A}$, and in fact the family $\langle\Theta_\phi : \phi \in \Sigma\rangle \in \mathcal{A}$.

Moreover

(i) every element $\theta \in \Theta_\phi$ satisfies $P(\phi,\theta)$: since P is a Σ_1-formula and w is a transitive set, we have $\mathcal{A} \models P^{(w)}(\phi,\theta) \Rightarrow \mathcal{A} \models P(\phi,\theta)$.

(ii) by (1), Θ_ϕ is non-empty for each $\phi \in \Sigma$.

Put $\Theta = \bigcup_{\phi\in\Sigma}\Theta_\phi$; $\Theta \in \mathcal{A}$. Let $\theta' = \bigvee\Theta = \bigvee_{\phi\in\Sigma,\theta\in\Theta_\phi}\theta$ and let θ_0 be a finitary formula with the same free variables as θ' that is $\mathcal{R}$-equivalent to θ_0, i.e. $(\theta_0)^{\mathcal{R}} = (\theta')^{\mathcal{R}}$.

Since I is $\mathcal{A}$-logical,

(iii) $(I(\theta_0))^{\mathcal{S}} = \bigvee^{(\mathcal{S})}_{\phi\in\Sigma,\theta\in\Theta_\phi}(I(\theta))^{\mathcal{S}}$.

For showing that θ_0 is actually a block for p as required, consider first the following inference:

$$\frac{\{\overbrace{\Delta_{N_1}, \bigvee\Sigma}^{\Delta_{N_1}}, \phi, \underset{\sim}{S}(a_0) \Rightarrow I(\theta)(\vec{a}) : \phi \in \Sigma, \theta \in \Theta_\phi\}}{\underbrace{\Delta_{N_1}, \bigvee\Sigma}_{\Delta_{N_1}}, \underset{\sim}{S}(a_0) \Rightarrow (\bigvee_{\phi\in\Sigma,\theta\in\Theta_\phi} I(\theta))(\vec{a})}$$

Notice that since Θ_ϕ are non-empty ((ii) above),

$$\bigvee\Sigma = \bigvee_{\phi\in\Sigma,\theta\in\Theta_\phi}\phi$$

so the displayed inference is sound, as easily seen by the definition and stability of the sups involved.

Since (by (i) above) θ is a block for q_ϕ if $\phi \in \Sigma$ and $\theta \in \Theta_\phi$, the premises are true in $\mathcal{S}$. Hence, so is the conclusion, which shows, together with (iii), that θ_0 satisfies the first condition for being a block for p.

Secondly, the sound inference

$$\frac{\{\Delta_{N_2}, I(\theta)(\vec{a}) \Rightarrow \underset{\sim}{S}(\underline{a}_0) : \phi \in \Sigma, \theta \in \Theta_\phi\}}{\Delta_{N_2}, (\bigvee_{\phi\in\Sigma,\theta\in\Theta_\phi} I(\theta))(\vec{a}) \Rightarrow \underset{\sim}{S}(\underline{a}_0)}$$

shows that θ_0 also satisfies the second condition for being a block for p.

This contradicts the fact that there is no block for p since $p \in P$. This contradiction shows that our assumption that Γ_1 does not satisfy (viii) cannot hold.

This completes showing that Γ_1 is a consistency property.

2. Γ_2 *is a consistency property.*

The only difference as compared to Γ_1 is in the case of condition (viii) in the definition of a consistency property.

Now, we have $\bigvee\Sigma \in \Delta_2 \underset{\text{df}}{=} \Delta^{(p)}_{N_2}$. We put q_ϕ (for $\phi \in \Sigma$) to agree with p on $\Delta_1 = \Delta_{N_1}$ and Δ_F and we put

$$\Delta^{(q)}_{N_2} = \Delta_2 \cup \{\phi\}.$$

Assume there is θ_ϕ blocking q_ϕ for each $\phi \in \Sigma$.

Using the admissibility of $\mathcal{A}$, next we show as in the proof for Γ_1 that there is an $\mathcal{A}$-finite family

$$\langle\Theta_\phi : \phi \in \Sigma\rangle$$

of *non-empty* sets Θ_ϕ such that for every $\phi \in \Sigma$, each $\theta \in \Theta_\phi$ blocks q_ϕ. We view this family as indexed by pairs $\langle\theta, \phi\rangle$, i.e., we consider the function

$$\langle\theta, \phi\rangle \mapsto \theta \quad (\phi \in \Sigma, \theta \in \Theta_\phi).$$

Consider $\theta' = \bigwedge_{\phi\in\Sigma,\theta\in\Theta_\phi}\theta = \bigwedge\bigcup_{\phi\in\Sigma}\Theta_\phi$; by assumption 7.3.4 the inf exists. Let θ_0 be a finitary formula $\mathcal{R}$-equivalent to the conjunction, with the same free variables. Then, of course, the I images are $\mathcal{S}$-equivalent. By 7.3.4(ii) it then follows that

$$(I(\theta_0))^{\mathcal{S}} = \bigwedge^{(\mathcal{S})}_{\phi\in\Sigma,\theta\in\Theta_\phi}(I(\theta))^{\mathcal{S}}.$$

The claim is that θ_0 is a block for p. We know that

$$\mathcal{S} \models \Delta_{N_1}, \underset{\sim}{S}(a_0) \Rightarrow I(\theta)(\vec{a}) \quad \text{for} \quad \phi \in \Sigma, \theta \in \Theta_\phi.$$

By the definition of infs, we have that

$$\mathcal{S} \models \Delta_{N_1}, \underset{\sim}{S}(a_0) \Rightarrow (\bigwedge_{\phi\in\Sigma,\theta\in\Theta_\phi} I(\theta))(\vec{a})$$

(more precisely, the automatic stability of infs is involved here too), hence

$$\mathcal{S} \models \Delta_{N_1}, \underset{\sim}{S}(a_0) \Rightarrow I(\theta_0)(\vec{a}). \tag{1}$$

On the other hand, we have

$$\mathcal{S} \models \Delta_{N_2}, \phi, I(\theta)(\underline{\vec{a}}) \Rightarrow \underset{\sim}{S}(\underline{a}_0) \quad \text{for} \quad \phi \in \Sigma, \theta \in \Theta_\phi. \tag{2}$$

Since each Θ_ϕ is *non-empty*, we have that for every $\phi \in \Sigma$ *there is* $\theta \in \Theta_\phi$ such that (2) holds and $(I(\theta))^{\mathcal{S}} \leq (I(\theta_0))^{\mathcal{S}}$ (since θ occurs as a member of the conjunction). So we have that for every $\phi \in \Sigma$

$$\mathcal{S} \models \Delta_{N_2}, \phi, I(\theta_0)(\underline{\vec{a}}) \Rightarrow \underset{\sim}{S}(\underline{a}_0).$$

Hence, by the definition and stability of the sup $(\bigvee\Sigma)^{\mathcal{S}}$, we have

$$\mathcal{S} \models \Delta_{N_2}, \bigvee\Sigma, I(\theta_0)(\underline{\vec{a}}) \Rightarrow \underset{\sim}{S}(\underline{a}_0).$$

$$\text{i.e.}\ \ \mathcal{S} \models \Delta_{N_2}, I(\theta_0)(\underline{\vec{a}}) \Rightarrow \underset{\sim}{S}(\underline{a}_0) \tag{2}$$

$$(\text{since } \bigvee\Sigma \in \Delta_{N_2})$$

(1) and (2) together show that θ_0 in fact is a block for p.

This completes the proof of 7.3.6.

7.3.6 enables us to construct the models N_1, N_2 as the canonical molders derived from the consistency properties Γ_1, Γ_2 by the "model existence theorem" 5.2.2. In particular, we have the Boolean value-algebra $\mathcal{B} = \mathcal{P}^* = (P, \leq)^*$ of regular open subsets of $\mathcal{P}$, common to Γ_1 and Γ_2, hence to N_1 and N_2. The domain of N_1 and N_2 consist of he set of terms of the language $L_{\mathcal{S}}$. We have that N_1 and N_2 are models of $T_{\mathcal{S}}$, hence $N_1 : \mathcal{S} \to \mathrm{Sh}_{\mathcal{B}}$, $N_2 : \mathcal{S} \to \mathrm{Sh}_{\mathcal{B}}$ are $\mathcal{B}$-valued $\mathcal{A}$-models of $\mathcal{S}$. Also, we have that

$$\|\phi\|_{N_1} = (U_1)^*_\phi \quad \text{where}$$

$$(U_1)_\phi = \{p \in P : \phi \in \Delta^{(p)}_{N_1}\},$$

$$\text{and} \quad \|\phi\|_{N_2} = (U_2)^*_\phi$$

$$\text{where} \quad (U_2)_\phi = \{p \in P : \phi \in \Delta^{(p)}_{N_2}\}.$$

Now we turn to the homomorphism F which of course also has been "built in" into P. F is to be construed as a natural transformation $F : I^*(N_1) \to I^*(N_2)$ of functors $\in (\mathrm{Sh}_{\mathcal{B}})^{\mathcal{R}}$. We define $F_R : N_1(I(R)) \to N_2(I(R))$ $(R \in \mathrm{Ob}(\mathcal{R}))$, a morphism of $\mathrm{Sh}_{\mathcal{B}}$ as follows. Recall that $|N_1(I(R))| = |N_2(I(R))| = \mathcal{X} = \mathcal{X}_R$ is the set of terms of sort $I(R)$.

For $a, \underline{a} \in \mathcal{X}$, we put $U_{\langle a,\underline{a}\rangle} = \{p \in P : \langle a, \underline{a}\rangle \in \Delta^{(p)}_F\}$ and $F_R(a, \underline{a}) =$ (also denoted as $\|\underline{a} = F_R(a)\|$) $\underset{\text{df}}{=}$

$$\bigvee_{a'\in\mathcal{X}} \bigvee_{\underline{a}'\in\mathcal{X}} \|a = a'\| \cdot \|\underline{a} = \underline{a}'\| \cdot U^*_{\langle a',\underline{a}'\rangle}.$$

Lemma 7.3.7 *Given* $p \in P$, $p = (\Delta_{N_1}, \Delta_{N_2}, \Delta_F)$, *and* $\langle \vec{s}, \vec{t}\rangle = \langle\langle s_0, t_0\rangle, \ldots, \langle s_k, t_k\rangle\rangle$ *a subsequence of* Δ_F, suppose *that* $\phi(\vec{x}) = \phi(x_0, \ldots, x_k)$ *is an atomic formula of* $L_{\mathcal{R}}$ *such* *that* $(I(\phi))(\vec{s}) \in \Delta_{N_1}$. *Then for* $q = (\Delta_{N_1}, \Delta_{N_2} \cup \{(I\phi)(\vec{t})\}, \Delta_F)$ *we have* $q \in P$ *and hence* $q \leq p$.

Proof. Assume, on the contrary, that $q \notin P$. Then there is θ blocking q, i.e., we have

$$T \vdash \Delta_{N_1}, \underset{\sim}{S}(a_0) \Rightarrow \theta(\vec{s}\,')$$

$$T \vdash \Delta_{N_2}, (I\theta)(\vec{t}), \theta(\vec{t}\,') \Rightarrow \underset{\sim}{S}(\underline{a}_0)$$

where $\Delta_F(\vec{s}\,', \vec{t}\,')$. Of course, we can write $(I\phi)(\vec{t}\,')$ for $(I\phi)(\vec{t})$ and $(I\phi)(\vec{s}\,')$ for $(I\phi)(\vec{s}) \in \Delta_{N_1}$. Now we have

$$T \vdash \Delta_{N_1}, \underset{\sim}{S}(a_0) \Rightarrow (I\phi)(\vec{s}\,') \wedge (I\theta)(\vec{s}\,')$$

and

$$T \vdash \Delta_{N_2}, (I\phi)(\vec{t}\,') \wedge (I\theta)(\vec{t}\,') \Rightarrow \underset{\sim}{S}(a_0)$$

showing that $\phi(\vec{x}) \wedge \theta$ is a block for p, contradicting $p \in P$. Notice that since ϕ is atomic, $\phi(\vec{x}) \wedge \theta$ is in fact a legitimate finitary formula of $L^g_{\mathcal{R}}$.

Lemma 7.3.8 *For sequences* $\vec{s}$ *and* $\vec{t}$ *of terms, let us write* $\|\vec{t} = F(\vec{s})\|$ *for* $\|t_0 = F_{R_0}(s_0)\| \cdots \|t_k = F_{R_k}(s_k)\|$ *where the sort of* s_i *and* t_i *is* $I(R_i)$. *Let* $\phi(\vec{x})$ *be an atomic formula of* $L_{\mathcal{R}}$. *Then we have that*

$$\|(I\phi)(\vec{s})\|_{N_1} \cdot \|\vec{t} = F(\vec{s})\| \leq \|(I\phi)(\vec{t})\|_{N_2}.$$

Proof. Easy on the basis of 7.3.7 and follows a pattern established before.

Lemma 7.3.9 *We have the following as required for* F *being a natural transformation* $I^*(N_1) \to I^*(N_2)$.

$$\text{(i)} \quad \|t = t'\|_{N_2} \cdot \|t = F(s)\| \leq \|t' = F(s)\|$$

$$\text{(ii)} \quad \|s = s'\|_{N_1} \cdot \|t = F(s)\| \leq \|t = F(s')\|$$

$$\text{(iii)} \quad \|t = F(s)\| \cdot \|t' = F(s)\| \leq \|t = t'\|_{N_2}.$$

Proof. (i) and (ii) are easy consequences of the definitions of $\|t = F(s)\|$ (the subscript R to F has been suppressed). (iii) is a consequence of 7.3.8 applied to an atomic formula $x = x'$.

Lemma 7.3.10 *For given* $p = (\Delta_{N_1}, \Delta_{N_2}, \Delta_F) \in P$, $\Delta_F = (\vec{s}, \vec{t})$, *and any term* s *of the sort* $I(R)$ *for some* $R \in |\mathcal{R}|_{\text{ob}}$, *such that each free variable of* s *is in* $\text{Var}_1(p) = \text{Var}(\Delta_{N_1}) \cup \{a_0\}$, *let* b *be a variable of sort* $I(R)$ *that does not occur in* $\text{Var}(\Delta_{N_2}) \cup \{\underline{a}_0\} \cup \text{Var}(\vec{t})$ *(b exists since the latter set if finite). Then* $q \underset{\text{df}}{=} (\Delta_{N_1}, \Delta_{N_2}, \Delta_F \cup \{\langle s, b\rangle\})$ *belongs to* P *and hence* $q \leq p$.

Proof. Suppose $\theta = \theta(\vec{x}, x)$ blocks q, i.e.

$$T \vdash \Delta_{N_1}, \underset{\sim}{S}(a_0) \Rightarrow (I\theta)(\vec{s}, s)$$

and

$$T \vdash \Delta_{N_2}, (I\theta)(\vec{t}, b) \Rightarrow \underset{\sim}{S}(a_0).$$

Then firstly,

$$T \vdash \Delta_{N_1}, \underset{\sim}{S}(a_0) \Rightarrow \exists x((I\theta)(\vec{s}, x)).$$

This is because $T \vdash \sigma \;\Leftrightarrow\; \mathcal{S} \models \sigma$ and the following inference:

$$\frac{\Gamma \Rightarrow \psi}{\Gamma \Rightarrow \exists x\ \psi(x)}$$

is a valid rule in $\mathcal{S}$ *provided* each free variable occurring in s occurs in Γ. This is left and an easy exercise; the restriction con be easily seen to be essential (this rule is a variant of the rule ($\Rightarrow \exists$) in Chapter 5). Now, the restriction on the free variables is assumed in the hypothesis of the lemma. Secondly, since b is a "new" variable, we also can infer

$$T \vdash \Delta_{N_2}, \exists x((I\theta)(\vec{t}, x)) \Rightarrow \underset{\sim}{S}(a_0)$$

This shows that $\exists x\theta$ blocks p, contrary to $p \in P$. Notice that θ is a formula of $(L^g_{\mathcal{R}})_{\omega\omega}$, so is $\exists x\theta$, as required for blocking.

Lemma 7.3.11 *$\|s = s\|_{N_1} = \bigvee_{t\in\mathcal{X}} \|t = F_R(s)\|$ for any $R \in |\mathcal{R}|_{\rm ob}$ and $\mathcal{X}$ the set of terms of sort $I(R)$, as required for F being a natural transformation.*

Proof. Easy on the basis of 7.3.10.

Lemma 7.3.12 *The following diagram*

$$\begin{array}{ccc} R_1 & \qquad I^*(N_1)(R_1) = N_1(I(R_1)) \xrightarrow{F_{R_1}} N_2(I(R_1)) \\ {\scriptstyle f}\downarrow & \qquad\qquad {\scriptstyle N_1(I(f))}\downarrow \qquad\qquad\qquad \downarrow{\scriptstyle N_2(I(f))} \\ R_2 & \qquad\qquad\qquad N_1(I(R_2)) \xrightarrow[F_{R_2}]{} N_2(I(R_2)) \end{array}$$

commutes (as required for F being a natural transformation $I(N_1) \to I(N_2)$).

Proof. The assertion is equivalent (by the definition of composition in $\mathrm{Sh}_{\mathcal{B}}$ to the identity

$$\begin{aligned} &\bigvee_{s' \text{ of sort } I(R_2)} \|s' = (If)(s)\|_{N_1} \cdot \|t = F_{R_2}(s')\| \\ = &\bigvee_{s'' \text{ of sort } I(R_1)} \|s'' = F_{R_1}(s)\| \cdot \|t = (If)(s'')\|_{N_2} \end{aligned}$$

to show this, first we deduce

$$\|s' = (If)(s)\|_{N_1} \cdot \|t = F_{R_2}(s')\| \le \bigvee_{s''} (\|s'' = F_{R_1}(s)\| \cdot \|s' = (If)s\|_{N_1} \cdot \|t = F_{R_2}(s')\|)$$

from 7.3.11. By 7.3.8, the contents of the parentheses is $\le \|t = (If)s''\|_{N_2}$ (apply 7.3.8, to the atomic formula $y = (If)x$ and ϕ). So, since $\|s'' = F_{R_1}(s)\|$ is itself a factor in the product in the parentheses, we conclude

$$\|s' = (If)(s)\|_{N_1} \cdot \|t = F_{R_2}(s')\| \le$$

the right hand side of the claimed identity, hence the '$\le$ part' of the equality is shown. The other part is similar.

Lemma 7.3.13 (i) $\|\underline{a}_0 = F_R(a_0)\| = 1_{\mathcal{B}}$

(ii) $\|\underset{\sim}{S}(a_0)\|_{N_1} = 1_{\mathcal{B}}$

(iii) $\|\underset{\sim}{S}(\underline{a}_0)\|_{N_2} = 0_{\mathcal{B}}$

(iv) *Assuming that I is* not *full with respect to subobjects, $0_{\mathcal{B}} \neq 1_{\mathcal{B}}$,*
i.e., $\mathcal{B}$ is nontrivial.

PROOF. (i) is left as an exercise. (iv) is a consequence of the remark we made after the definition of $\mathcal{P}$ that under the assumption we have $(\varnothing, \varnothing, \langle\langle a_0, \underline{a}_0\rangle\rangle) \in P$, thus P is non-empty.

Let us show (iii). This will be a consequence of the fact that $\underset{\sim}{S}(\underline{a}_0) \notin \Delta^{(p)}_{N_2}$ for any $p \in P$. In fact, if $\underset{\sim}{S}(\underline{a}_0) \in \Delta^{(p)}_{N_2}$, then obviously, the trivial formula $\uparrow = \bigwedge\varnothing$ would block p from being a member of P.

The proof of (ii) is similar.

SUMMARY OF THE PROOF OF 7.3.5 Assume, contrary to the assertion of the theorem, that I is not full with respect to subobjects. We construct the complete Boolean algebra $\mathcal{B} = \mathcal{P}^*$ as above and conclude that $\mathcal{B}$ is nontrivial by 7.3.13(iv). We construct, by 7.3.6 and 4.2.2, the canonical $\mathcal{B}$-valued models N_1 and N_2 of T on the basis of the consistency properties Γ_1 and Γ_2. We also construct the morphism F in the category of $\mathcal{B}$-valued $\mathcal{R}$ models $F : I^*(N_1) \to I^*(N_2)$. 7.3.9, 7.3.11 and 7.3.12 tell us that F is indeed a natural transformation. Finally, 7.3.13 (i), (ii) and (iii) say that $\underset{\sim}{S}(a_0)$ is not preserved by F that clearly contradicts the assumption of the theorem that F can be lifted to a natural transformation $N_1 \to N_2$. This proves the theorem under the first assumption (i). With the other assumptions (ii) and (iii) of 7.3.5, now the result is an easy corollary. Consider first (ii). Under the indirect hypothesis, we have the nontrivial complete $\mathcal{B}$, N_1 and N_2 and F as above; notice that the language $L_{\mathcal{S}}$ is countable, the logic $(L_{\mathcal{S}})_{\mathcal{A}}$ is countable as well as the set of all elements in the domains of N_1 and N_2, the latter being terms of $L_{\mathcal{S}}$. Apply the Rasiowa-Sikorski theorem (4.3.1) to obtain a 2-valued homomorphism h of $\mathcal{B}$ that preserves all sups

$$\bigvee\nolimits^{\mathcal{B}}_{s\in|N_j|_s} \|\phi(s)\| = \|\exists x\phi(x)\|_{N_j}$$

$$\bigvee\nolimits^{\mathcal{B}}_{i} \|\phi_i\|_{N_j} = \|(\bigvee\nolimits_i \phi_i)\|_{N_j}$$

for $\bigvee\phi_i \in (L_{\mathcal{S}})_{\mathcal{A}}$, for $j = 1$ and 2 (the sups that come up in evaluating $(L_{\mathcal{S}})_{\mathcal{A}}$-formulas in N_1 and N_2). Define

$$\begin{gathered} M_1 = N_1/h \\ M_2 = N_2/h \\ \text{(c.f. Chapter 4)} \\ \text{and } \; G : M_1 \to M_2 \\ \text{by } \; G_R = \{(s,t) : h(\|t = F_R(s)\|) = 1\}. \end{gathered}$$

It is easy to see that we have obtained a counterexample to assumption (ii) of the theorem.

Part (iii) of 7.3.5 is identical to 7.1.4. To obtain it from part (i) of 7.3.5, one first has to replace N_1, N_2 and $F : I^*(N_1) \to I^*(N_2)$ by N_1', N_2' and $F' : I^*(N_1') \to I^*(N_2')$ where N_1', N_2' are full int he sense of 4.3.3; the precise statement and the proof of what we need should be obvious for those familiar with the proof of 4.3.4. Having done this replacement, we can use *any* 2-valued homomorphism $h : \mathcal{B} \to 2$ to define the appropriate two-valued models and homomorphisms.

This completes the proof of 7.3.5.

§4 Infinitary generalizations; continued

Next we turn to the infinitary generalization of Theorem 7.1.6. Besides Assumption 7.3.4, we have to make further assumptions. Let $\mathcal{R}$ and $\mathcal{S}$ be weakly $\mathcal{A}$-logical categories and let $I : \mathcal{R} \to \mathcal{S}$ be an $\mathcal{A}$-logical functor as it was assumed at the outset of the last section.

Assumption 7.4.1 (i) $\mathcal{A}$-finite disjoint sums exist in $\mathcal{R}$, i.e. if $\langle R_i : i \in J \rangle$ is an $\mathcal{A}$-finite family of objects in $|\mathcal{R}|_{\mathrm{ob}}$ then there exists an object $R \in |\mathcal{R}|_{\mathrm{ob}}$ and an $\mathcal{A}$-finite family $\langle R_i \xrightarrow{j_i} R : i \in J \rangle$ of morphisms such that (a) each j_i is a monomorphism, (b) R is the sup of the R_i as subobjects, and finally (c) for distinct $i_1, i_2 \in J$, the subobjects $R_{i_1} \overset{j_{i_1}}{\hookrightarrow} R$, $R_{i_2} \overset{j_{i_2}}{\hookrightarrow} R$ of R are disjoint.

(ii) The following form of *the axiom of choice* holds. For any $\mathcal{A}$-finite family $\langle R_i : i \in J \rangle$ of objects in $|\mathcal{R}|_{\mathrm{ob}}$ there are an object R in $|\mathcal{R}|_{\mathrm{ob}}$ and an $\mathcal{A}$-finte family of morphisms $\langle R \xrightarrow{\pi'_i} R_i : i \in J \rangle$ satisfying the following condition $(*)$ (we call R together with the π'_i a pseudo-$\mathcal{A}$-product of R_i, for reasons explained below).

$(*)$ For any object S of $\mathcal{S}$ and any $\mathcal{A}$-finite system $S_i \overset{\sigma_i}{\hookrightarrow} I(R_i) \times S$, if we denote by $I(R_i) \times S \xrightarrow{\rho_i} S$ and $I(R) \times S \xrightarrow{\rho} S$ the canonical projections and by $I(R) \times S \xrightarrow{\pi_i} I(R_i) \times S$ the morphism induced by $I(R) \xrightarrow{I(\pi'_i)} I(R_i)$, then the following equality holds:

$$\bigwedge_{i \in J} \exists_{\rho_i}(S_i) = \exists_\rho \bigwedge_{i \in J} \pi_i^{-1}(S_i)$$

$$\begin{tikzcd}
 & I(R_i) & I(R) \arrow[l, "I(\pi'_i)"'] & \\
S_i \arrow[r, hook, "\sigma_i"] \arrow[d] & I(R_i) \times S \arrow[u] \arrow[d, "\rho_i"] & I(R) \times S \arrow[l, "\pi_i"'] \arrow[u] \arrow[dl, "\rho"'] & \\
\exists_{\rho_i}(S_i) \arrow[r, hook] & S & \bigwedge_{i \in J} \pi_i^{-1}(S_i) \arrow[u, hook] \arrow[r] \arrow[d] & \pi_i^{-1}(S_i) \arrow[ul, hook, "\text{p.b.}"'] \\
\bigwedge_{i \in J} \exists_{\rho_i}(S_i) \arrow[ur, hook] & \simeq & \exists_\rho \bigwedge_{i \in J} \pi_i^{-1}(S_i) \arrow[ul, hook] & \\
 & \text{requirement} \arrow[u] & &
\end{tikzcd}$$

Remark 1. *It is easy to see that a disjoint sum as described in* 7.4.1(i) *is always a coproduct.*

2. *The condition* (ii) *takes place entirely in the category $\mathcal{S}$ except that we do not need more than the existence of such a pseudo-$\mathcal{A}$-products for objects $I(R_i)$ coming from $\mathcal{R}$.*

3. *Condition* (ii) *is a form of the axiom of choice. Consider the category of sets as $\mathcal{S}$. Let A_i $(= I(R_i))$ be given sets. Now A $(= I(R))$ can be taken to be the usual Cartesian product $\prod A_i$, with canonical projections p_i $(= I(\pi'_i)) : A \to A_i$. Now $(*)$ will hold. S_i is a subset of $A_i \times S$ and we write $S_i(a, s)$ for $\langle a, s \rangle \in S_i$. The condition now becomes, in usual logical notation,*

$$\bigwedge_{i \in I} \exists a_i \in A_i \ S_i(a_i, s) \ \Leftrightarrow \ \exists \langle a_i : i \in J \rangle \bigwedge_{i \in J} S_i(a_i, s).$$

The left-to-right implication says that if each of the sets $\{a_i \in A_i : S_i(a_i, s)\} = X_s^{(i)}$ is non-empty, so is their cartesian product $\prod_{i \in J} X_s^{(i)}$.

More generally, assume that $\mathcal{S}$ is at least weakly-$\mathcal{A}$-logical, it has $\mathcal{A}$-products of $\mathcal{A}$-finite families (meaning that $\mathcal{A}$-finite families have a product with the universal property *of product required only for $\mathcal{A}$-finite families of maps) and finally, that it satisfies the axiom of choice in the form that every epimorphism has a section. Then, as it is easy to see, condition* (ii) *is satisfied.*

4. *For* finite *index-sets J, the condition of* (ii) *is automatically satisfied in any ordinary logical category. So we see that for the finitary logical case, when the admissible set $\mathcal{A}$ is taken to consist of hereditarily finite sets over certain urelements, Assumption* 7.4.1 *is automatically true once $\mathcal{R}$ is a pretopos (this is needed for the first condition* (i)*). Thus, Theorem* 7.1.6 *will be a special case of our next result, at least when $\mathcal{R}$ is a pretopos.*

Theorem 7.4.2 *Assume that $\mathcal{R}$, $\mathcal{S}$ and $I : \mathcal{R} \to \mathcal{S}$ are as before and the satisfy both Assumptions* 7.3.4 *and* 7.4.1*. Then, if $I^*_{\mathcal{B}} : \mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{S}) \to \mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{R})$ is faithful for every non-trivial complete Boolean algebra $\mathcal{B}$, then every object S in $\mathcal{S}$ is* covered by $\mathcal{R}$ via *I, i.e. there is an object R in $\mathcal{R}$, a subobject $S_1 \to I(R)$ of $I(R)$ in $\mathcal{S}$ and a morphism $S_1 \xrightarrow{f} S$ such that $S = \exists_f(S_i)$. Also, if $\mathcal{A}$ is countable* or *each set in $\mathcal{A}$ is* finite *(i.e., the logic $L_{\mathcal{A}}$ is finitary), then it is sufficient to require faithfulness for the two-element Boolean algebra $\mathcal{B}$.*

Remark *In the presence of disjoint sums, the notion of being covered as given now is equivalent to being finitely covered as used in the statement of* 7.1.6*. This point will become clear below in or use of disjoint sums.*

Proof of 7.4.2 Many of the computations in this proof are similar to those of 7.3.5 as well as previous uses of Boolean-valued models with regular open sets as Boolean values. Most of these computations will now be omitted. Although the general framework will be quite similar to the of the last section, the "mathematical content" is sufficiently distinct to deserve attention.

We start by describing a partially ordered set $\mathcal{P} = (P, \leq)$ that will be the "consistency machine" doing the work for us, just as in the last section.

Let P be the set of all triples

$$\langle \Delta_{N_1}, \Delta_{N_2}, \Delta_{F,G} \rangle$$

such that the following conditions (i), (ii) and (iii) are satisfied.

(i) Δ_{N_1}, Δ_{N_2} are finite sets of formulas of the fragment $F_{\mathcal{S}}$ (defined at the beginning of §3).

(ii) $\Delta_{F,G}$ consists of two sequences, $\langle \langle a_i, \underline{a}_i \rangle : i < k \rangle$ and $\langle \langle b_j, \underline{b}_j, \underline{\underline{b}}_j \rangle : j < \ell \rangle$ such that, for each i, a_i and $\underline{a}_i$ are terms of the same sort that is of the form $I(R)$ for some $R \in |\mathcal{R}|_{\mathrm{ob}}$; for each j, b_j, $\underline{b}_j$, $\underline{\underline{b}}_j$ are terms of the same sort that is *not* of the form $I(R)$ for some $R \in |\mathcal{R}|_{\mathrm{ob}}$. Also, each of the two sequences should be a *function*, i.e. $\underline{a}_i$ is uniquely determined by a_i, and $\underline{b}_j$, $\underline{\underline{b}}_j$ by b_j. Δ_F denotes the sequence (extracted from $\Delta_{F,G}$)

$$\langle \langle a_i, \underline{a}_i \rangle \langle b_j, \underline{b}_j \rangle : i < k, j < \ell \rangle$$

and Δ_G:

$$\langle \langle a_i, \underline{a}_i \rangle \langle b_j, \underline{\underline{b}}_j \rangle : i < k, j < \ell \rangle.$$

We will write $\underline{a} = F(a)$, or $(a, \underline{a}) \in \Delta_{F,G}$, to denote that the pair $\langle a, \underline{a} \rangle$ is among the $\langle a_i, \underline{a}_i \rangle$. Similarly for $\langle b_j, \underline{b}_j, \underline{\underline{b}}_j \rangle \in \Delta_{F,G}$ we write $\underline{b} = F(b)$ and $\underline{\underline{b}} = G(b)$. We will make a systematic distinction between terms of sorts $I(R)$ "coming form $\mathcal{R}$" and other terms by the use of letters a (with possible indices) for the former and b for the latter.

We reserve the distinct variables c, $\underline{c}$ and $\underline{\underline{c}}$ of a fixed sort S_0 (whose being cover by $\mathcal{R}$ we are proving) and we put the requirement on $\Delta_{F,G}$ that always $\langle c, \underline{c}\rangle, \langle c, \underline{\underline{c}}\rangle \in \Delta_{F,G}$. (e.g. by $b_0 = c$, $\underline{b}_0 = \underline{c}$, $\underline{\underline{b}}_0 = \underline{\underline{c}}$).

We finally make the requirement that each (free) variable in terms a (or b) such that $\langle a, \underline{a}\rangle \in \Delta_F$ (or $\langle b, \underline{b}\rangle \in \Delta_F$) for some $\underline{a}$ (or $\underline{b}$) should be either c or else it should occur free in Δ_{N_1}.

(iii) There does not exist a formula $\theta = \theta(x, c, \vec{u}, \vec{v})$ such that

($*$) θ is in the finitary logic $(L_{\mathcal{S}})^g_{\omega\omega}$, x is a variable of some sort of the form $I(R)$,

$$T_{\mathcal{S}} \vdash \Delta_{N_1} \Rightarrow \exists x \theta(x, c, \vec{a}, \vec{b}\,)$$

and

$$T_{\mathcal{S}} \vdash \Delta_{N_2}, \exists x[\theta(x, \underline{c}, \vec{\underline{a}}, \vec{\underline{b}}\,) \wedge \theta(x, \underline{\underline{c}}, \vec{\underline{a}}, \vec{\underline{\underline{b}}})] \Rightarrow \underline{c} \approx \underline{\underline{c}}$$

where $\vec{a}$, $\vec{\underline{a}}$, $\vec{b}$, etc. refer to the sequences $\langle a_i : i < k\rangle$, $\langle \underline{a}_i : i < k\rangle$, $\langle b_j : j < \ell\rangle$ etc. coming from $\Delta_{F,G}$, and, of course, the terms in $\vec{a}$ are substituted for the variables $\vec{u}$ (and thus the variables $\vec{u}$ are assumed to have sorts matching those of $\vec{a}$).

(We will say, as before, that θ *blocks* p if all the conditions after ($*$) *are* satisfied. Hence, (iii) says that there is no θ blocking p.)

This completes the definition of the set P. The partial ordering $\leq$ is defined, as before, as component-wise inclusion.

Remark *Some words about the rationale behind the consistency machine. We make the indirect assumption that the fixed object S_0 in $\mathcal{S}$ is* not *covered by $\mathcal{R}$ via I, and proceed to show the existence of Boolean valued models N_1 and N_2 of $\mathcal{S}$ (i.e., of $T_{\mathcal{S}}$) together with homomorphisms F and G*

$$N_1 \underset{G}{\overset{F}{\rightrightarrows}} N_2$$

such that the restrictions $I^*(F)$ and $I^*(G)$ are the same but F and G themselves are not, in direct contradiction to the faithfulness assumption. The universes of the models will consist of terms as before. Very roughly speaking, Δ_{N_1} and Δ_{N_2} are approximations of the full theories of the models N_1 and N_2. Δ_F and Δ_G are approximations of the homomorphisms F and G respectively, which fact is expressed in the notations $\overline{a} = F(a)$, etc., introduced above.

Some light is thrown on the crucial definition of blocking by checking the following. Assume that S_0 is not covered by $\mathcal{R}$. Then $p_0 = \langle \varnothing, \varnothing, \langle\langle \underline{c}, \underline{\underline{c}}\rangle\rangle\rangle$ belongs to P. This means the following. *If* there is a θ blocking p_0, then S_0 is covered by $\mathcal{R}$. And in fact, with $\theta(x, c)$ blocking p_0, we obtain that $S = (\exists c \theta(x, c))^{\mathcal{S}}$ is a subobject of $I(R)$ where x is of the sort $I(R)$, $P' = (\theta(x, c))^{\mathcal{S}}$ is a subobject of $I(R) \times S_0$ and also, $S \times S_0$, and (here we use the second part of the block property) that P' is a functional subset of $S \times S_0$, hence by 2.4.4 it defines a morphism $S \xrightarrow{f} S_0$; and finally, by the first part of the block property, "f is surjective", i.e. $S_0 = \exists_f(S)$, showing that S_0 is covered by $\mathcal{R}$ (through $S \to I(R)$).

The two consistency properties derived from $(P, \leq) = \mathcal{P}$, denoted by Γ_1 and Γ_2, are defined almost identically to the case in the preceding section. Put $f_1(p) = \Delta^{(p)}_{N_1}$, $\mathrm{Var}_1(p) = \mathrm{Var}(\Delta^{(p)}_{N_1}) \cup \{c\}$, $f_2(p) = \Delta^{(p)}_{N_2}$ and $\mathrm{Var}_2(p) = \mathrm{Var}(\Delta^{(p)}_{N_2}) \cup \mathrm{Var}(\mathrm{Range}\ \Delta^{(p)}_{F,G})$ where by $\mathrm{Var}(\mathrm{Range}\ \Delta^{(p)}_{F,G})$ we mean

$$\bigcup\{\mathrm{Var}(\underline{a}) : \langle a, \underline{a}\rangle \in \Delta_{F,G}\} \cup \bigcup\{\mathrm{Var}(\underline{b}) : \langle b, \underline{b}\rangle \in \Delta_{F,G}\} \cup \bigcup\{\mathrm{Var}(\underline{\underline{b}}) : \langle b, \underline{\underline{b}}\rangle \in \Delta_{F,G}\}.$$

Then we define

$$\Gamma_1 = (\Gamma, \leq, f_1, \mathrm{Var}_1)$$

and

$$\Gamma_2 = (\Gamma, \leq, f_2, \mathrm{Var}_2).$$

Lemma 7.4.3 Γ_1 *and* Γ_2 *are consistency properties.*

The proof is similar to that of the corresponding statement in §3 and the only two points of interest are the proofs of property (viii) ("for disjunctions") in the definition of consistency property for Γ_1 and Γ_2.

(i) ((viii) FOR Γ_1) Let $p = (\Delta_{N_1}, \Delta_{N_2}, \Delta_{F,G})$ and $\bigvee\Sigma \in \Delta_{N_1}$. Define p_ϕ for $\phi \in \Sigma$ by

$$p_\phi = (\Delta_{N_1} \cup \{\phi\}, \Delta_{N_2}, \Delta_{F,G}.$$

Assume (for reductio ad absurdum) that

$$p_\phi \notin P \quad \text{for all} \quad \phi \in \Sigma.$$

Hence for every $\phi \in \Sigma$ there is θ_ϕ blocking p_ϕ, i.e.

$$(1) \qquad \begin{cases} T_{\mathcal{S}} \vdash \Delta_{N_1}, \phi \Rightarrow \exists x^{(\phi)}\ \theta_\phi(x^{(\phi)}, c, \vec{a}, \vec{b}) \\ T_{\mathcal{S}} \vdash \Delta_{N_2}, \exists x^{(\phi)}(\theta_\phi(x^{(\phi)}, \underline{c}, \underline{\vec{a}}, \underline{\vec{b}}) \wedge \theta_\phi(x^{(\phi)}, \underline{\underline{c}}, \underline{\underline{\vec{a}}}, \underline{\underline{\vec{b}}})) \Rightarrow \underline{c} \approx \underline{\underline{c}}. \end{cases}$$

By the assumption that $\mathcal{S}$ is weakly $\mathcal{A}$-logical, the predicate $P(\phi, \theta)$,

$$P(\phi, \theta) \ \Leftrightarrow\ \theta \text{ is a finitary formula and (1) holds with } \theta = \theta_\phi;$$

is Σ on $\mathcal{A}$.

By the admissibility of $\mathcal{A}$, we conclude in a manner similar to the proof of 7.3.5 that there is an $\mathcal{A}$-finite family

$$\{\Theta_\phi : \phi \in \Sigma\}$$

of $\mathcal{A}$-finite sets Θ_ϕ of formulas such that each Θ_ϕ is non-empty ($\phi \in \Sigma$) and for every $\theta \in \Theta_\phi$, we have $P(\phi, \theta)$.

Consider the $\mathcal{A}$-finite set

$$J = \{\langle \phi, \theta \rangle : \theta \in \Theta_\phi\}$$

and the J-indexed family

$$\langle \theta : \langle \phi, \theta \rangle \in J \rangle$$

(i.e., the function $\langle \phi, \theta \rangle \mapsto \theta$, defined on J). Let us denote by $x^{(\phi,\theta)}$ the variable θ playing the role of $x^{(\phi)}$ in (1). $x^{(\phi,\theta)}$ is of sort $I(R^{(\phi,\theta)})$ and the function $(\phi, \theta) \mapsto R^{(\phi,\theta)}$, defined on J, is $\mathcal{A}$-finite. (Strictly speaking, this is not a consequence and should be made sure by a slightly more careful choice of Θ_ϕ as a certain set of pairs $\langle \theta, x \rangle$.)

Let us consider the disjoint sum, say R, (Assumption 7.4.1) of the objects $R^{(\phi,\theta)}$, for $\langle \phi, \theta \rangle \in J$. Because I is $\mathcal{A}$-logical, it is easy to see that $I(R)$ is the disjoint sum of the $I(R^{(\phi,\theta)})$ in $\mathcal{S}$. Let $\ell_{(\phi,\theta)} = I(j_{(\phi,\theta)})$ be the canonical injection

$$I(R^{(\phi,\theta)}) \to I(R).$$

Now, define

$$\theta'_{(\phi,\theta)}(x, c, \vec{a}, \vec{b}) := \exists x^{(\phi,\theta)}[x \approx \ell_{(\phi,\theta)}(x^{(\phi,\theta)}) \wedge \theta(x^{(\phi,\theta)}, c, \vec{a}, \vec{b})].$$

Finally, put

$$\theta(x, c, \vec{a}, \vec{b}) := \bigvee_{(\phi,\theta)\in J} \theta'_{(\phi,\theta)},$$

more precisely, a finitary formula $\mathcal{S}$-equivalent to this. It is immediate that

$$(\bigvee_{(\phi,\theta)\in J} \exists x^{(\phi,\theta)}\, \theta(x^{(\phi,\theta)}, c, \vec{a}, \vec{b}))^{\mathcal{S}} \leq (\exists x\, \theta(x, c, \vec{a}, \vec{b}))^{\mathcal{S}}$$

and hence, by $\bigvee\Sigma \in \Delta_{N_1}$, by the fact that $P(\phi, \theta)$ for $(\phi, \theta) \in J$ and by the fact that for each $\phi \in \Sigma$ there is θ such that $(\phi, \theta) \in J$, we conclude

$$T_{\mathcal{S}} \vdash \Delta_{N_1} \Rightarrow \exists x\, \theta(x, c, \vec{a}, \vec{b}).$$

By parts (a) and (c) of the definition of disjoint sum (c.f. 7.4.1), we can easily see that

$$(\exists x(\theta(x, \underline{c}, \underline{\vec{a}}, \underline{\vec{b}}) \wedge \theta(x, \underline{\underline{c}}, \underline{\underline{\vec{a}}}, \underline{\underline{\vec{b}}})))^{\mathcal{S}} \leq (\bigvee_{(\phi,\theta)\in J} \exists x^{(\phi,\theta)}(\theta(x^{(\phi,\theta)}, \underline{c}, \underline{\vec{a}}, \underline{\vec{b}}) \wedge \theta(x^{(\phi,\theta)}, \underline{\underline{c}}, \underline{\underline{\vec{a}}}, \underline{\underline{\vec{b}}})))^{\mathcal{S}}$$

(the reader should ponder the standard meaning of the formulas involved and argue by the completeness theorem). This leads to the other half of the required condition, namely

$$T_{\mathcal{S}} \vdash \Delta_{N_2}, \exists x(\theta(x, \underline{c}, \underline{\vec{a}}, \underline{\vec{b}}) \wedge \theta(x, \underline{\underline{c}}, \underline{\underline{\vec{a}}}, \underline{\underline{\vec{b}}})) \Rightarrow \underline{c} \approx \underline{\underline{c}}.$$

We have exhibited in θ a block for p, contrary to $p \in P$; hence for some $\phi \in \Sigma$, we have $p_\phi \in P$, showing "(viii)".

Remark *We did not use part* (b) *of the disjoint sum. And in fact,* (b) *is inessential in the sense that if we had a "sum" with only* (a) *and* (c)*, then taking the sup of the canonical images of the* R_i*, we obtain one satisfying* (b) *in addition.*

(i) ((viii) FOR Γ_2) Let $p = (\Delta_{N_1}, \Delta_{N_2}, \Delta_{F,G})$, $\bigvee\Sigma \in \Delta_{N_2}$; define p_ϕ for $\phi \in \Sigma$ by

$$p_\phi = (\Delta_{N_1}, \Delta_{N_2} \cup \{\phi\}, \Delta_{F,G}).$$

Assume, for reductio ad absurdum, that

$$p_\phi \notin P \quad \text{for} \quad \phi \in \Sigma.$$

Hence for every $\phi \in \Sigma$ there is θ_ϕ blocking p_ϕ, i.e.

$$(2) \qquad \begin{aligned} &T_{\mathcal{S}} \vdash \Delta_{N_1} \Rightarrow \exists x^{(\phi)}\ \theta_\phi(x^{(\phi)}, c, \vec{a}, \vec{b}) \\ &T_{\mathcal{S}} \vdash \Delta_{N_2}, \phi, \exists x^{(\phi)}(\theta_\phi(x^{(\phi)}, \underline{c}, \underline{\vec{a}}, \underline{\vec{b}}) \wedge \theta_\phi(x^{(\phi)}, \underline{\underline{c}}, \underline{\underline{\vec{a}}}, \underline{\underline{\vec{b}}})) \Rightarrow \underline{c} \approx \underline{\underline{c}}. \end{aligned}$$

Let, as before, Θ_ϕ (for $\phi \in \Sigma$) be a non-empty $\mathcal{A}$-finite set, such that $\langle \Theta_\phi : \phi \in \Sigma \rangle$ is $\mathcal{A}$-finite, and for $\theta \in \Theta_\phi$, (2) holds with $\theta \in \Theta_\phi$. Consider the family $\langle \theta : \langle \phi, \theta \rangle \in \coprod_{\phi\in\Sigma} \Theta_\phi \rangle$, and the corresponding family

$$\langle x^{(\phi,\theta)} : (\phi, \theta) \in J \rangle$$

with $x^{(\phi,\theta)}$ the variable of sort $I(R^{(\phi,\theta)})$ playing the role of $x^{(\phi)}$ in (2). We can easily arrange that $\langle R^{(\phi,\theta):(\phi,\theta)\in J} \rangle$ is $\mathcal{A}$-finite. We write i for a typical element (ϕ, θ) of J. Now we use 7.4.1(ii). Let $\langle R \xrightarrow{\pi'_i} R_i : j \in J \rangle$ be a pseudo-$\mathcal{A}$-product of the R_i, and let $\mathcal{S}$ be the product-object of the sorts c, $\vec{a}$ and $\vec{b}$ so that

$$S_i \hookrightarrow I(R_i) \times S$$

for $S_i = (\theta)^{\mathcal{S}}$, with $i = (\phi, \theta)$. Let $I(R_i) \times S \xrightarrow{\rho_i} S$, $I(R) \times S \xrightarrow{\rho} S$ and $I(R) \times S \xrightarrow{\pi_i} I(R_i) \times S$ be the morphisms as in 7.4.1(ii). Let x be a variable of sort $I(R)$. We now have, as a matter of course,

$$\bigwedge_{i\in J} \exists_{\rho_i}(S_i) = (\bigwedge_{\phi\in\Sigma,\theta\in\Theta_\phi} \exists x^{(\phi,\theta)}\, \theta(x^{(\phi,\theta)}, c, \vec{a}, \vec{b}))^{\mathcal{S}}$$

and

$$\exists_\rho \bigwedge_{i\in J} \pi_i^{-1}(S_i) = (\exists x \bigwedge_{\phi\in\Sigma,\theta\in\Theta_\phi} \theta(\pi_{(\phi,\theta)}(x), c, \vec{a}, \vec{b}))^{\mathcal{S}}.$$

We define θ_0 to be (a finitary formula $\mathcal{R}$-equivalent to) $\bigwedge_{\phi\in\Sigma,\theta\in\Theta_\phi} \theta(\pi_{(\phi,\theta)}(x), \cdot, \cdot, \cdot)$. Now, by 7.4.1(ii), the left-hand sides of the last two equalities are equal, hence so are the two right-hand sides. This implies, but the first fact listed in (2) for $\theta = \theta_\phi$, that

$$T_{\mathcal{S}} \vdash \Delta_{N_1} \Rightarrow \exists x\, \theta(x, c, \vec{a}, \vec{b}).$$

By another application of the pseudo-$\mathcal{A}$-product property of R, we obtain this:

$$(\bigwedge_{\phi\in\Sigma,\theta\in\Theta_\phi} \exists x^{(\phi,\theta)} (\theta(x^{(\phi,\theta)}, \underline{c}, \underline{\vec{a}}, \underline{\vec{b}}) \wedge \theta(x^{(\phi,\theta)}, \underline{\underline{c}}, \underline{\underline{\vec{a}}}, \underline{\underline{\vec{b}}})))^{\mathcal{S}}$$

$$= (\exists x (\theta_0(x, \underline{c}, \underline{\vec{a}}, \underline{\vec{b}}) \wedge \theta_0(x, \underline{\underline{c}}, \underline{\underline{\vec{a}}}, \underline{\underline{\vec{b}}})))^{\mathcal{S}}$$

(the S and S_i of the condition are changed now, in an easily identifiable way). This gets us, from the second part of (2) for $\theta = \theta_\phi$, $\theta \in \Theta_\phi$, $\phi \in \Sigma$ and by using that $\bigvee \Sigma \in \Delta_{N_2}$ and

$$\bigwedge_{\phi\in\Sigma,\theta\in\Theta_\phi} \phi = \bigwedge_{\phi\in\Sigma} \phi = \bigwedge \Sigma \quad \text{by} \quad \Theta_\phi \neq \varnothing,$$

the the conclusion

$$T_{\mathcal{S}} \vdash \Delta_{N_2}, \exists x (\theta_0(x, \underline{c}, \underline{\vec{a}}, \underline{\vec{b}}) \wedge \theta_0(x, \underline{\underline{c}}, \underline{\underline{\vec{a}}}, \underline{\underline{\vec{b}}})) \Rightarrow \underline{c} \approx \underline{\underline{c}}.$$

By now we know that this finishes the proof of 7.4.3.

The rest is entirely straightforward and is just like the proof of 7.3.5. We have the non-trivial Boolean algebra $\mathcal{B} = \mathcal{P}^*$, the canonical $\mathcal{B}$-valued models of $T_{\mathcal{S}}$, N_1, N_2, derived from Γ_1 and Γ_2, respectively, by 4.2.2. ($\mathcal{B}$ is non-trivial because of the indirect assumption of the proof.) Next we define the natural transformations F and $G : N_1 \to N_2$ (each N_i being considered as a functor $\mathcal{R} \to \mathrm{Sh}_{\mathcal{B}}$). We define

$$\|\underline{d} = F_S(d)\| = \bigvee_{d\in\mathcal{X}_S} \bigvee_{d'\in\mathcal{X}_S} \|d = d'\| \cdot \|\underline{d} = \underline{d}'\| \cdot U^{(F)}_{\langle d', \underline{d}'\rangle}$$

where $\mathcal{X}_S$ is the set of terms of sort S, d, $\underline{d}$ are terms of sort S. There is an identical formula for G. In a manner entirely similar to the last section, we establish that

(i) F, G are natural transformations

(ii) $I^*(F) = I^*(G)$

(iii) $\|\underline{c} = F(c)\| = \|\underline{\underline{c}} = G(c)\| = 1$ but $\|\underline{c} = \underline{\underline{c}}\| = 0$, hence $F \neq G$.

This completes the proof of the theorem.

The final result of this section and in fact, the main result of this chapter is

Theorem 7.4.4 *Let $\mathcal{R}$, $\mathcal{S}$ be weakly $\mathcal{A}$-logical categories, and I an $\mathcal{A}$-logical functor $I : \mathcal{R} \to \mathcal{S}$, with an admissible set $\mathcal{A}$. Assume that both* 7.3.4 *and* 7.4.1 *are satisfied and that $\mathcal{R}$ is a pretopos. Then if*

$$I^*_{\mathcal{B}} : \mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{S}) \to \mathrm{Mod}^{\mathcal{B}}_{\mathcal{A}}(\mathcal{R})$$

is an equivalence of categories for all non-trivial complete Boolean algebras $\mathcal{B}$, then

$$I : \mathcal{R} \to \mathcal{S}$$

is an equivalence. If $\mathcal{A}$ is countable, or *each set in $\mathcal{A}$ is finite, it is sufficient to require that $I^* : \mathrm{Mod}_{\mathcal{A}}(\mathcal{S}) \to \mathrm{Mod}_{\mathcal{A}}(\mathcal{R})$ is an equivalence (the above for $\mathcal{B}$ = the two-element algebra).*

The proof is immediate by 7.1.7, 7.3.2, 7.3.5 and 7.4.2.

As remarked above, for the finitary case, the Assumptions 7.3.4 and 7.4.1 made in the theorem are automatically satisfied. So 7.1.8 is a special case of 7.4.4.

We can regard 7.4.4 as an analysis of what exactly of the assumption of finitariness is used in 7.1.8. The answer is that we use that (i) the hereditarily finite sets form an admissible set (with arbitrary added predicates) and (ii) that the statements of 7.3.4 and 7.4.1 are true.

Chapter 8

Theories as categories

§1 Categories and algebraic logic

In the preceding, we correlated a theory $T_{\mathcal{R}}$ with a given category $\mathcal{R}$ such that $T_{\mathcal{R}}$ could replace $\mathcal{R}$ for many purposes. Here we are going to perform the opposite step to show that categories and theories are practically indistinguishable. First we present a certain general perspective in which we can place our discussion.

The basic notion is that of an interpretation of a theory in category, the one we have been working with extensively so far. To have a precise codification, a theory for us is a pair $\mathcal{T} = (F, T)$ where F is a fragment of $L_{\infty\omega}$ and T is a set of axioms, i.e. sequents, of F. An *interpretation* (or *model*) of $\mathcal{T}$, M, in a category $\mathcal{R}$ is an interpretation of the language of F in $\mathcal{R}$ such that (i) M is "adequate" for F in a sense such as: F is stable with respect to M *or:* F is distributive with respect to M (c.f. Chapter 3); and such that (ii) M makes every axiom in T hold in $\mathcal{R}$. We employ the arrow notation (and we justify this below)

$$M : \mathcal{T} \to \mathcal{R} \tag{1}$$

to signify the fact that M is a model of $\mathcal{T}$ in $\mathcal{R}$.

There are *two* related kinds of arrows that offer themselves for comparison. One is a functor

$$I : \mathcal{R} \to \mathcal{S} \tag{2}$$

between categories. Under certain conditions, we can compose the two arrow (1) and (2), to obtain a composite

$$I \circ M : \mathcal{T} \to \mathcal{S}$$

in the obvious way: first of all, $I \circ M$ will always make sense as an $\mathcal{S}$-interpretation of the language L of F, in $\mathcal{T}(F, T)$. In order to have that $I \circ M$ is actually a model of $\mathcal{S}$, we need a certain degree of "logicalness" on the part of I. For example, if F is a finitary coherent fragment and I is *logical* in the technical sense introduced above, then clearly, $I \circ M$ is a model too.

The other kind of arrow

$$I' : \mathcal{T}' \to \mathcal{T}$$

is between theories, and is what is ordinarily called a (relative) interpretation of the theory $\mathcal{T}'$ in $\mathcal{T}$. The ordinary definition is clearly not broad enough; e.g. it does not incorporate the possibility of interpreting individuals of sort s as *pairs* of individuals

$\langle a_1, a_2 \rangle$ of sorts s_1, s_2 in $\mathcal{T}$. Nevertheless, it is clear that, under any reasonable definition of this notion, we'll have a composite

$$M \circ I' : \mathcal{T}' \to \mathcal{S}$$

and this will be a *model* of $\mathcal{T}'$. We note that our actual aim in this Chapter is to *subsume* the first and third kinds of arrows under the notion of functors.

In the following discussion, we fix F to be the full finitary coherent fragment over a given language L and we call a theory of the form (F, T) (finitary) *coherent*. We assume all categories to be *logical*. In fact, all functors will be assumed *logical*.

First we claim that for a given logical category $\mathcal{R}$, the theory $\mathcal{T}_\mathcal{R} = (F_\mathcal{R}, T_\mathcal{R})$ with $T_\mathcal{R}$ introduced above and with the obvious $F_\mathcal{R}$, has, and in fact, is essentially characterized by, a universal property, as follows.

Proposition 8.1.1 *For an arbitrary $\mathcal{R}$-model $M : \mathcal{T} \to \mathcal{R}$ where $\mathcal{T} = (F, T)$ is a coherent theory (F is not necessarily $F_\mathcal{R}$) there is an essentially unique interpretation $I : \mathcal{T} \to \mathcal{T}_\mathcal{R}$ such that*

$$\begin{tikzcd} \mathcal{T} \arrow[r, "M"] \arrow[d, "I"'] & \mathcal{R} \\ \mathcal{T}_\mathcal{R} \arrow[ur, "\text{canonical}"'] & \end{tikzcd}$$

commutes.

Even with an "unsatisfactory" definition of interpretation (see above), this is a completely obvious statement.

Another fact (which has been used extensively above) is

Proposition 8.1.2 (Universal property of $\mathcal{R}$ with respect to $\mathcal{T}_\mathcal{R}$.) *For any model*

$$M : \mathcal{T}_\mathcal{R} \to \mathcal{S}$$

there is a unique logical functor $I : \mathcal{R} \to \mathcal{S}$ such that

$$\begin{tikzcd} \mathcal{T}_\mathcal{R} \arrow[r, "M"] \arrow[d, "\text{canonical}"'] & \mathcal{S} \\ \mathcal{R} \arrow[ur, "I"'] & \end{tikzcd}$$

commutes.

This suggest the question if the theories $\mathcal{T}_\mathcal{R}$ are any special with having a "universal" model $\mathcal{R}$; the answer is that any coherent theory has this property.

Theorem 8.1.3 *For any finitary coherent theory $\mathcal{T}$ there is a logical category $\mathcal{R} = \mathcal{R}_\mathcal{T}$ together with an $\mathcal{R}$ model $M_0 : \mathcal{T} \to \mathcal{R}$ such that: for any $M : \mathcal{T} \to \mathcal{S}$, an $\mathcal{S}$ model for an arbitrary logical category $\mathcal{S}$ there is a logical functor $I : \mathcal{R} \to \mathcal{S}$ such that*

$$\begin{tikzcd} \mathcal{T} \arrow[r, "M_0"] \arrow[dr, "M"'] & \mathcal{R}_\mathcal{T} \arrow[d, "I"] \\ & \mathcal{S} \end{tikzcd}$$

commutes; I is determined uniquely up to a unique isomorphism.

$\mathcal{R}_{\mathcal{T}}$ has an even stronger property that is related to categories of models. For models $M_1, M_1 : \mathcal{T} \to \mathcal{S}$, a *morphism* $F : M_1 \to M_2$ is what is ordinarily called a homomorphism. I.e., F is a family

$$\langle F_s : s \text{ is a sort of } L \rangle$$

of morphisms $F_s : M_1(s) \to M_2(s)$ in $\mathcal{S}$ such that for any operation symbol

$$f : s_1 \times \cdots \times s_n \to s$$

we have the commutative diagram

$$\begin{array}{ccc} M_1(s_1) \times \cdots \times M_1(s_n) & \xrightarrow{M_1(f)} & M_1(s) \\ {\scriptstyle F_{s_1} \times \cdots \times F_{s_n}} \downarrow & & \downarrow {\scriptstyle F_s} \\ M_2(s_1) \times \cdots \times M_2(s_n) & \xrightarrow[M_2(f)]{} & M_2(s) \end{array}$$

There is a related diagram concerning relation symbols of the language. This is of course closely related to the notion of natural transformation; in case of $\mathcal{T} = \mathcal{T}_{\mathcal{R}}$, the theory associated with a logical category $\mathcal{R}$, a homomorphism between models of $\mathcal{T}$ is exactly a natural transformation between them as functors $\mathcal{R} \to \mathcal{S}$.

The above notion of (homo-)morphism, with an obvious notion of composition, defines the *category of $\mathcal{S}$-models of $\mathcal{T}$*, $\mathrm{Mod}_{\mathcal{S}}(\mathcal{T})$. Now, it is easy to see that if

$$\mathcal{T} \xrightarrow{M} \mathcal{R}$$

and $\mathcal{S}$ are given, then

$$M^*_{\mathcal{S}} : \mathrm{Mod}_{\mathcal{S}}(\mathcal{R}) \to \mathrm{Mod}_{\mathcal{S}}(\mathcal{T}),$$

defined by composition, is actually a functor (extends in a natural way to a morphism). Now we can state

Theorem 8.1.4 *With $\mathcal{R}_{\mathcal{T}}$ and M_0 essentially uniquely determined by 8.1.3, we have that for any logical $\mathcal{S}$*

$$(M_0)^* : \mathrm{Mod}_{\mathcal{S}}(\mathcal{R}_{\mathcal{T}}) \to \mathrm{Mod}_{\mathcal{S}}(\mathcal{T}),$$

is an equivalence of categories.

The proofs of 8.1.3 and 8.1.4 are discussed in the next section. Here we make a few remarks.

First of all, $\mathcal{R}_{\mathcal{T}}$ is the result of a usual sort of 'universal construction' that is quite straightforward, and in fact, essentially determined by the property of $\mathcal{R}_{\mathcal{T}}$ itself. However, we will see that, by considering pretopoi as categories $\mathcal{R}_{\mathcal{T}}$ with a theory $\mathcal{T}$ arising in a natural way, we will be able to give a new, formal or syntactic, view of the category of coherent objects in a coherent topos. In short, it turns out to be quite useful to keep in mind the actual construction of $\mathcal{R}_{\mathcal{T}}$ as described in the next section.

Secondly, let us point out that there is a close analogy between the construction of $\mathcal{R}_{\mathcal{T}}$ and the construction of the Lindenbaum-Tarski algebra $\mathcal{B}_{\mathcal{T}}$ of a theory. In fact, in the case $\mathcal{T}$ has a negation (i.e., for any formula $\phi(\vec{x})$ there is another one, $\psi(\vec{x})$, such that $T \vdash \phi(\vec{x}), \psi(\vec{x}) \Rightarrow$ and $T \vdash \Rightarrow \phi(\vec{x}), \psi(\vec{x})$, $\mathcal{B}_{\mathcal{T}}$ will be a more or less well identifiable *part* of $\mathcal{R}_{\mathcal{T}}$. This is a point where we can observe how categories provide an algebraization of logic. Just as cylindric and polyadic algebras are richer than Boolean algebras, the category $\mathcal{R}_{\mathcal{T}}$ is a richer structure than the Lindenbaum-Tarski algebra, and this makes it able to fully reflect the content of the theory $\mathcal{T}$.

The content of 8.1.3 and 8.1.4 can be expressed by saying that for all practical purposes, $\mathcal{T}$ and $\mathcal{R}_{\mathcal{T}}$ are the same. If one draws the conclusion from this that thereby logic as conceived traditionally has been eliminated in favor of categorical notions, then one should be reminded that 8.1.2 and 8.1.3 together will equally show that categories (in the "logical" situation at least) can be eliminated in favor of logic. We think that the main point is that the present natural identification of logic with a categorical formulation is, at least potentially, useful since it makes possible relating two vast but so far rather unrelated resources, namely logical-model theoretical experience and category theoretical experience.

§2 The categorization of a coherent theory

The construction of $\mathcal{R}_{\mathcal{T}}$ given here is described in great detail in Dionne [1973]. Our exposition will be somewhat sketchy, but will contain all the essential details.

We will freely use the *completeness theorem* (Theorem 5.1.7) to conclude instances of formal consequence relationships $T \vdash \sigma$ form the fact that in all ordinary set-models of T, σ holds in the usual sense. This could be avoided at the expense of tedious formal reasonings.

Let $\mathcal{T} = (F, T)$ be a (finitary) coherent theory over a language L where F is the full finitary coherent fragment over L. The *objects* of $\mathcal{R}_{\mathcal{T}}$ are defined to be all the formulas in F. We note that we can economize on objects and still have an *equivalent* category by taking any subset $\mathcal{O}$ of formulas of T to be the set of objects of the category such that for any ϕ in F there is a ψ in $\mathcal{O}$ with exactly the same free variables such that $T \vdash \phi \Rightarrow \psi$ and $T \vdash \psi \Rightarrow \phi$ (for which we will write $T \vdash \phi \Leftrightarrow \psi$). One such choice for $\mathcal{O}$ is the following. Let $\mathcal{O}$ consist of all formulas of the form

$$\bigvee_{i=1}^{n} \exists x_{i1} \cdots \exists x_{ik_i} \bigwedge_{j=1}^{m_i} \theta_{ij}$$

where each θ_{ij} is an atomic formula, in fact, a *simple* atomic formula: a *simple* atomic formula is one of the form

$$Ry_1 \cdots y_\ell$$

for any (ℓ-ary) predicate symbol L, or

$$fy_1 \cdots y_\ell \approx y_{\ell+1}$$

for any (ℓ-ary) operation symbol in L (here the y_i are *variables*). The natural numbers n, k_i, m_i can each be equal to zero; so $\uparrow = \bigwedge\varnothing$ and $\downarrow = \bigvee\varnothing$ are represented in $\mathcal{O}$.

We also note that we consider a countable infinity of variables for each sort fixed once a for all and use only them in forming the objects of $\mathcal{R}_{\mathcal{T}}$.

For reasons of easier orientation, we will write sometimes $[\phi(\vec{x})]$ when we mean $\phi(\vec{x})$ as an object of $\mathcal{R}_{\mathcal{T}}$.

The morphisms of $\mathcal{R}_{\mathcal{T}}$ will be given by "definable mappings" between the object-formulas; here, however, the theory will be taken into account and provably equal maps will be identified. Let $\phi(\vec{x})$ and $\psi(\vec{y})$ be two objects of $\mathcal{R}_{\mathcal{T}}$, with the distinct free variables indicated. A *premorphism* $\phi \to \psi$ is any formula

$$\mu(\vec{x}\,', \vec{y}\,')$$

such that (i) the variables in the sequences $\vec{x}\,', \vec{y}\,'$ have, term for term, the same sorts as the corresponding variables $\vec{x}, \vec{y}$, respectively;

(ii) the sequences $\vec{x}\,', \vec{y}\,'$ have disjoint ranges, and each consists of distinct variables; and finally

(iii) $T \vdash$ "μ is functional"
where "μ is functional" are the three sequents

$$\begin{array}{rcl} \mu(\vec{x}\,', \vec{y}\,') & \Rightarrow & \phi(\vec{x}\,') \wedge \psi(\vec{y}\,') \\ \mu(\vec{x}\,', \vec{y}\,') \wedge \mu(\vec{x}\,', \vec{y}\,'') & \Rightarrow & \vec{y}\,' \approx \vec{y}\,'' \\ \phi(\vec{x}\,') & \Rightarrow & \exists \vec{y}\,' \mu(\vec{x}\,', \vec{y}\,') \end{array}$$

with the obvious conditions and abbreviations regarding the handling of variables. Two premorphisms $\mu_1(\vec{x}\,', \vec{y}\,')$ and $\mu_2(\vec{x}\,'', \vec{y}\,''): \phi \to \psi$ are *equivalent* if $T \vdash \mu_1(\vec{x}\,', \vec{y}\,') \Leftrightarrow \mu_2(\vec{x}\,', \vec{y}\,')$ (note the change of $\vec{x}\,'', \vec{y}\,''$ to $\vec{x}\,', \vec{y}\,'$). It is clear (strictly speaking, by the completeness theorem, since one argues with $T \models \quad$ instead of $T \vdash \quad$) that this defines an equivalence relation $\sim = \sim_{(\phi,\psi)}$ between premorphisms $\phi \to \psi$. Finally, a *morphism* $\phi \to \psi$ is a triple

$$(\mu/\sim, \phi, \psi)$$

with $mu/\sim$ being the equivalence class of a premorphism $\mu: \phi \to \psi$. We will write $[\mu]$ for the morphism exhibited. Sometimes we use a suffix such as in $[\mu(\vec{x}, \vec{y}\,')]_{(\vec{x} \to \vec{y}\,')}$ to point out the roles of the variables.

The *composition* is of no surprise either. First, we verify that given premorphisms

$$\mu(\vec{x}\,', \vec{y}\,'): \phi(\vec{x}) \to \psi(\vec{y})$$

$$\nu(\vec{y}\,'', \vec{z}\,''): \psi(\vec{y}) \to \theta(\vec{z}),$$

the formula $\lambda(\vec{x}\,', \vec{z}\,')$, with $\vec{z}\,'$ chosen so that $\vec{x}\,' \cap \vec{z}\,' = \varnothing$ and $\vec{y}\,' \cap \vec{z}\,' = \varnothing$, defined as

$$\exists \vec{y}\,'(\mu(\vec{x}\,', \vec{y}\,') \wedge \nu(\vec{y}\,', \vec{z}\,'))$$

is a premorphism $\phi \to \theta$. One does this, of course, by checking that λ is again functional with can be easily seen on the lever of (set-)models, and then inferred on the "formal level" ($T \vdash \quad$) by the completeness theorem. Next one verifies that changing μ and ν to equivalent premorphisms, the composition premorphism λ changes to an equivalent one too. This defines, finally, the composition

$$[\nu] \circ [\mu] = [\nu \circ \mu] = [\lambda].$$

Next we verify that the composition is associative. The identity morphism $\phi(\vec{x}) \to \phi(\vec{x})$ is represented by the premorphism $\mu(\vec{x}, \vec{y}) := \vec{x} \approx \vec{y}$. ($\vec{x} \approx \vec{y}$ denotes $x_1 \approx y_1, \ldots, x_n \approx y_n$, $\vec{x} = \langle x_i \rangle_{i=1,\ldots,n}, \vec{y} = \langle y_i \rangle_{i=1,\ldots,n}$.)

Having defined the category $\mathcal{R}_{\mathcal{T}}$, we next name M_0 the obvious interpretation of the language L in $\mathcal{R}_{\mathcal{T}}$. M_0 will correlate the object $Rx_1 \cdots x_\ell$ with the (ℓ-ary) predicate symbol R and the morphism represented by the premorphism

$$fx_1 \cdots x_\ell \approx x_{\ell+1}$$

with the (ℓ-ary) operation symbol f (with the x_i having matching sorts).

Proposition 8.2.1 *$\mathcal{R}_{\mathcal{T}}$ is a logical category and $M_0: \mathcal{T} \to \mathcal{R}_{\mathcal{T}}$ is an $\mathcal{R}_{\mathcal{T}}$-model of $\mathcal{T}$.*

Lemma 8.2.2 (i) *The category $\mathcal{R}_{\mathcal{T}}$ has finite left limits.*

(ii) *Every formula of F is interpretable by M_0, and in fact, the interpretation function $M_0(\cdot)$ is essentially the identity. More precisely, if $\phi(\vec{x})$ is a formula of F, $X = X_1 \times$*

$\cdots \times X_n$ is the product of the objects $[x_i \approx x_i]$ ($\vec{x} = \langle x_1, \ldots, x_n\rangle$, $X = [\bigwedge_{i=1}^n x_i \approx x_i]$, then $M_0(\phi(\vec{x})) = (M_0)_{\vec{x}}(\phi(\vec{x}))$ exists and it is the subobject

$$[\phi(\vec{x})] \hookrightarrow X$$

with the injection represented by the premorphism $\mu(\vec{x}, \vec{x}\,') = \phi(\vec{x}) \wedge \bigwedge_{i=1}^n x_i \approx x_i'$. (If we have chosen a proper subset of the set of formulas for objects, we will have to take a formula equivalent to ϕ for forming the subobject required.)

For the more general case

$$(M_0)_{\vec{y}}(\phi(\vec{x}))$$

with $\vec{x}$ included in $\vec{y}$, there is a corresponding natural formula. Now we will have that $(M_0)_{\vec{y}}(\phi(\vec{x}))$ is

$$[\phi(\vec{x}) \wedge \bigwedge_{i=1}^n y_i \approx y_i] \xhookrightarrow{[\phi(\vec{x}) \wedge \bigwedge_{i=1}^n y_i \approx y_i']} \bigwedge_{i=1}^n y_i \approx y_i.$$

(iii) *An arbitrary sequent σ is true in M_0, $M_0 \models \sigma$ iff $T \vdash \sigma$. In particular, M_0 is a model of T.*

Proof (AD (i)) Given two objects $A = [\phi(\vec{x})]$ and $B = [\psi(\vec{y})]$, the product of A and B will be $C = [\phi(\vec{x}) \wedge \psi(\vec{y}\,')]$ where the $\vec{y}\,'$ are chosen so that $\vec{x}$ and $\vec{y}\,'$ are disjoint and $\vec{y}\,'$ is termwise of the same sort as $\vec{y}$. The projection $C \xrightarrow{\pi_A} A$ is represented by the premorphism $\mu(\vec{x}, \vec{y}\,'; \vec{x}\,') := \phi(\vec{x}) \wedge \psi(\vec{y}\,') \wedge \vec{x} \approx \vec{x}\,'$. There is a similar expression for the projection $C \xrightarrow{\pi_B} B$.

Next we directly verify that $\begin{tikzcd}[row sep=tiny] & A \\ C \arrow[ur, "\pi_A"] \arrow[dr, "\pi_B"'] & \\ & B \end{tikzcd}$ has the required universal property. Given any $\begin{tikzcd}[row sep=tiny] & A \\ D \arrow[ur, "f_A"] \arrow[dr, "f_B"'] & \\ & B \end{tikzcd}$, $D = [\phi_D(\vec{z})]$ and f_A, f_B are represented by premorphisms, say $\mu'_A(\vec{z}, \vec{x}\,'')$, $\mu'_B(\vec{z}, \vec{y}\,'')$. Then the obvious candidate for the required morphism $D \xrightarrow{g} C$ will be represented by

$$\mu_{D,C}(\vec{z}, \vec{x}\,'', \vec{y}\,'') := \mu'_A(\vec{z}, \vec{x}\,'') \wedge \mu'_B(\vec{z}, \vec{y}\,'').$$

The necessary facts $f_A = \pi_A \circ g$, $f_B = \pi_B \circ g$ are then seen to be equivalent to the fact that certain sequents constructed form the above formulas are consequences of T which fact in turn will then be seen by inspection.

The uniqueness of g requires a similar argument.

Similarly, we can define and verify equalizers.

This completes our sketch of showing (i).

(AD (ii)) By induction on the complexity of the formula ϕ. We will say a few words on the induction step concerning the passage from $\phi(\vec{x}, y)$ to $\exists y\ \phi(\vec{x}, y)$ and leave the rest to the reader.

We verify that

$$[\phi(\vec{x}, y)] \xrightarrow{[\phi(\vec{x},y) \wedge \vec{x} \approx \vec{x}\,']} [\exists y\ \phi(\vec{x}, y)]$$

is an image-diagram in $\mathcal{R}_{\mathcal{T}}$. This, together with the induction hypothesis

$$M_0(\phi(\vec{x}, y)) = [\phi(\vec{x}, y)]$$

will show the analogous claim for $\exists y\ \phi$. (Strictly speaking one should consider $(M_0)_{\vec{y}}(\exists y\ \phi(\vec{x}, y))$, etc.)

The required verification is done in a direct fashion, similarly to that in the case of the proposed product diagram handled above. The minimality of $B = [\exists y\ \phi(\vec{x}, y)]$ among all subobjects C of B is proved by exhibiting a formula that defines, as a premorphism, an isomorphism between C and B under the assumption that $A = [\phi] \to B$ factors through $C \to B$ and then by proving that the exhibited formula indeed works.

(AD (iii)) Assume that $T \vdash \phi(\vec{x}) \Rightarrow \psi(\vec{y})$ and let $\vec{z}$ be the union of $\vec{x}$ and $\vec{y}$. The subobject $A \hookrightarrow Z$:

$$[\phi(\vec{x}) \wedge \vec{z} \approx \vec{z}] \xhookrightarrow{[\phi(\vec{x})\wedge\vec{z}\approx\vec{z}\,']_{(\vec{z}\to\vec{z}')}} [\vec{z} \approx \vec{z}]$$

is *smaller than* the corresponding subobject $B \hookrightarrow Z$ served from ψ, by the morphism

$$[\phi(\vec{x}) \wedge \vec{z} \approx \vec{z} \wedge \vec{z} \approx \vec{z}\,']_{(\vec{z}\to\vec{z}')} : A \to B.$$

This is a direct consequence of the definition of composition and the fact $T \vdash \phi \Rightarrow \psi$.

By 8.2.2(ii), $(M_0)_{\vec{z}}(\phi)$ is A and $(M_0)_{\vec{z}}(\psi)$ is B, hence we have

$$(M_0)_{\vec{z}}(\phi) \leq (M_0)_{\vec{z}}(\psi), \qquad i.e.$$

$$M_0 \models \phi \Rightarrow \psi.$$

The argument is completely reversible, showing the other direction.

Proof of 8.2.1 We make the general remark that the task of doing the remaining verifications is made somewhat easier by the use of 8.2.2 and the earlier Theorem 2.4.5. In particular, if we want to verify that a given diagram composed of objects and morphisms in $\mathcal{R}_{\mathcal{T}}$ has a certain property, e.g. it is an image diagram, by 2.4.5 and 8.2.1(i), it suffices to show that certain sequents (namely, in the example, the single image-axiom (item 9 in 2.4.5)) are valid in $\mathcal{R}_{\mathcal{T}}$. By 8.2.2(iii), this follows from (and in fact, is equivalent to) the fact that these sequents are consequences of T. So, e.g. in the case of a proposed image-diagram the image property will be verified once a sequent is shown to be a consequence of T.

Now, to show that $\mathcal{R}_{\mathcal{T}}$ has images, let us take a diagram

$$A = [\phi_A(\vec{x})] \xrightarrow{f=[\mu(\vec{x},\vec{y})]_{(\vec{x}\to\vec{y})}} B = [\psi_B(\vec{y})].$$

Define $A \xrightarrow{g} C \xrightarrow{h} B$ by

$$\begin{aligned} C &= [\exists \vec{x}\ \mu(\vec{x}, \vec{y})] \\ g &= [\mu(\vec{x}, \vec{y}\,')] \\ h &= [\exists \vec{x}\ \mu(\vec{x}, \vec{y}) \wedge \vec{y} \approx \vec{y}\,']. \end{aligned}$$

Then by the method indicated above we can verify that the subobject $C \hookrightarrow B$ is in fact the image $\exists_f(A)$.

When we want to verify that a given image $B = \exists_f(A)$ is stable under pullbacks, we proceed similarly. Let us start with the pullback diagram

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "g'"'] \arrow[dr, phantom, "\text{p.b.}"] & B \arrow[d, "g"] \\ A' \arrow[r, "f'"'] & B' \end{tikzcd}$$

and assume $B = \exists_f(A)$. We want to show that $B' = \exists_{f'}(A')$.

We take formulas representing each of the objects and morphisms involved. By 2.4.5 (and 3.5.2), the hypotheses (of the pullback character and $B = \exists_f(A)$) are expressed equivalently by saying that certain sequents, built up from the given formulas, are true in M_0. By 8.2.2(iii), these sequents are consequences of T. Now, the required conclusion also is equivalent to a sequent being a consequence of T. We now use plain common sense to conclude that the latter sequent is a consequence of the earlier ones, interpreted in ordinary set-models. By the completeness theorem, the required conclusion follows.

Existence and satiability of finite sups are shown similarly.

This completes our sketch of the proof of 8.2.1.

Remark *It would be more satisfactory to have a set of inference rules for defining $T \vdash (\cdot)$ so that each particular fact needed to verify* 8.2.1 *would be a direct consequence of a rule present, rather than having to appeal to a deep completeness theorem concerning the system $T \vdash (\cdot)$. Of course, such a system could be st up automatically by examining the needs of the proof of* 8.2.1*. Probably, it would be hard to make this proof-system attractive. – There is a similar discrepancy between a natural (Gentzen-type) proof-system (with non-logical axioms and cut rule) for ordinary one-sorted logic on the one hand and the Hilbert-type system that is "inherent" in the notion of Lindenbaum-Tarski algebra.*

Proposition 8.2.3 (The universal property of $\mathcal{R}_\mathcal{T}$). *Given a model of a theory $\mathcal{T}$*

$$M : \mathcal{T} \to \mathcal{S}$$

in a logical category $\mathcal{S}$, there is a logical functor $I : \mathcal{R}_\mathcal{T} \to \mathcal{S}$ such that the diagram

$$\begin{tikzcd}
\mathcal{T} \arrow[r, "M"] \arrow[dr, "M_0"'] & \mathcal{S} \\
& \mathcal{R}_\mathcal{T} \arrow[u, "I"']
\end{tikzcd}$$

is commutative. I is uniquely determined up to a unique isomorphism, i.e., if I_1, I_2 both satisfy the requirements, then there is a unique natural transformation $\nu : I_1 \to I_2$ that is an isomorphism in the category of all functors $\mathcal{R}_\mathcal{T} \to \mathcal{S}$.

Proof. We first exhibit a suitable I. Given a formula $\phi(\vec{x})$, the interpretation $M_{\vec{x}}(\phi(\vec{x}))$ is determined only up to isomorphism among diagrams of the following sort

$$\begin{tikzcd}
& \vdots \\
M(\vec{x}) \arrow[r, "\pi_i"] & M(x_i) \\
& \vdots \\
M_{\vec{x}}(\phi) \arrow[uu, hook]
\end{tikzcd} \tag{1}$$

where (i) $M(\vec{x}) \xrightarrow{\pi_i} \begin{array}{c}\vdots\\ M(x_i)\\ \vdots\end{array}$ is a product of the *given* $M(x_i)$ (but otherwise is undetermined) and (ii) $M_{\vec{x}}(\phi) \to M(\vec{x})$ is a monomorphism, (iii) some additional properties (related to the meaning of ϕ) are satisfied.

To define the functor I, we have to use the axiom of choice. Given an object $A = [\phi(\vec{x})]$ of $\mathcal{R}_\mathcal{T}$, let $I'(\phi)$ be any one of the diagrams (1). Let $I(A)$ be the object $M_{\vec{x}}(\phi)$ in the diagram $I'(\phi)$. To define the action of I on morphisms, let

$$A = [\phi(\vec{x})] \xrightarrow{f = [\mu(\vec{x}, \vec{y}\,')]_{(\vec{x} \to \vec{y}\,')}} B = [\psi(\vec{y})]$$

be a morphism in $\mathcal{R}_{\mathcal{T}}$. The subobject

$$M_{\vec{x},\vec{y}\,'}(\mu) \hookrightarrow M(\vec{x}) \times M(\vec{y}) = X$$

is smaller than the subobject $I(A) \times I(B) \to X$ derived from the diagrams $I'(\phi)$ and $I'(\psi)$ and in fact, $M_{\vec{x},\vec{y}\,'}(\mu)$ will be a functional subobject of $I(A) \times I(B)$; all these facts are consequences of μ being a premorphism. By 2.4.4, there is a unique morphism

$$I(A) \xrightarrow{f'} I(B)$$

whose graph is $M_{\vec{x},\vec{y}\,'}(\mu)$. We put $I(f) = f'$.

Having defined I, we have to verify that I is logical. Let D be a diagram in $\mathcal{R}_{\mathcal{T}}$ with a given property to be shown to be preserved by I. We will apply 2.4.5 and 3.5.2 again.

By these results and 8.2.2(iii), the fact that D has the given property is translated to saying that certain sequents are consequences of T. Using the definition of I, we realize that the fact that the I-image of D in $\mathcal{S}$ also has the given property, is equivalent that the same sequents are *true in* M. Since M was supposed to be a model of T, by the soundness theorem 3.5.4, we conclude that these sequents are indeed true in M.

The uniqueness part will be proved as a consequence of

Proposition 8.2.4 *For any logical category $\mathcal{S}$, the natural functor*

$$(M_0)^*_{\mathcal{S}} \colon \mathrm{Mod}(\mathcal{R}_{\mathcal{T}}) \to \mathrm{Mod}_{\mathcal{S}}(\mathcal{T})$$

is an equivalence of categories.

First, let us remark that the fullness of $(M_0)^*_{\mathcal{S}}$ implies that if $I_1 \circ M_0 = I_2 \circ M_0 = M$ in the notation of 8.2.3, then $I_1 \simeq I_2$ in $\mathrm{Mod}_{\mathcal{S}}(\mathcal{R}_{\mathcal{T}})$, hence the uniqueness statement in 8.2.3.

Secondly, let us note that 8.2.4 generalizes 8.2.3 and in fact, the (essential) surjectivity of $(M_0)^*$ is equivalent to the existence statement of 8.2.3. Hence what is left to show is the fullness and faithfulness of $(M_0)^*_{\mathcal{S}}$. Let I_1, I_2 be logical functors $\mathcal{R}_{\mathcal{T}} \to \mathcal{S}$ and let F be a homomorphism $M_1 = I_1 \circ M_0 \to M_2 = I_2 \circ M_0$. We want to define a natural transformation $G \colon I_1 \to I_2$ such that $(M_0)^*_{\mathcal{S}} G = F$. Let $[\phi(\vec{x})]$ be any object of $\mathcal{R}_{\mathcal{T}}$. Consider the following diagram in $\mathcal{R}_{\mathcal{T}}$:

$$\begin{tikzcd}
 & & {[x_1 \approx x_1]} \\
{[\phi(\vec{x})]} \arrow[r, hook, "i"] & X \arrow[ur, "\pi_1"] \arrow[dr, "\pi_n"'] & \vdots \\
 & & {[x_n \approx x_n]}
\end{tikzcd}$$

with the obvious i, $\pi_1, \ldots, \pi_n$ ($\vec{x} = \langle x_1, \ldots, x_n \rangle$). This diagram is carried over to $\mathcal{S}$ both by I_1 and I_2; the resulting two diagrams are related by F. All this is depicted in the following.

$$\begin{tikzcd}
 & M_1(x_1) \arrow[r, "F_{x_1}"] & M_2(x_1) & \\
I_1(X) \arrow[ur, "I_1(\pi_1)"] \arrow[dr, "I_1(\pi_n)"'] \arrow[rrr, dashed, "{(F_{x_1},\cdots,F_{x_n})}"] & \vdots & \vdots & I_2(X) \arrow[ul, "I_2(\pi_1)"'] \arrow[dl, "I_2(\pi_n)"] \\
 & M_1(x_n) \arrow[r, "F_{x_n}"'] & M_2(x_n) & \\
(M_1)_{\vec{x}}(\phi(\vec{x})) = I_1([\phi(\vec{x})]) \arrow[uu, "I_1(i)"] \arrow[rrr, dashed, "G_{[\phi(\vec{x})]}"'] & & & I_2([\phi(\vec{x})]) = (M_2)_{\vec{x}}(\phi(\vec{x})) \arrow[uu, "I_2(i)"']
\end{tikzcd}$$

Notice that with s_i the sort of x_i, $M_1(x_i) \underset{\text{df}}{=} M_1(s_i) = I_1[x_i \approx x_i]$ and $F_{x_i} \underset{\text{df}}{=} F_{s_i} : M_1(x_i) \to M_2(x_i)$. We claim that there is a unique morphism in $\mathcal{S}$, $G_{[\phi(\vec{x})]}$, making the diagram commutative. The uniqueness is clear since $I_2(i)$ is a monomorphism. The existence of such a morphism can be proved by an induction on the complexity of the formula ϕ, by parallelling the proof that existential positive formulas are preserve under homomorphisms. This latter fact is the special case of the above statement for the case when $\mathcal{S}$ is the category of sets.

This completes the definition of the proposed natural transformation $G : I_1 \to I_2$. It is easy to see that $(M_0)^*(G) = F$, i.e., that G "extends" F. Next, this is how we can verify that G is in fact a natural transformation. Recall the following elementary fact: a diagram

$$\begin{tikzcd} A \arrow[r, "\nu"] \arrow[d, "f"'] & A' \arrow[d, "f'"] \\ B \arrow[r, "\mu"'] & B' \end{tikzcd}$$

is commutative in any category with finite $\varprojlim$ iff there is a morphism h making the following commute:

$$\begin{tikzcd} \mathrm{graph}(f) \arrow[r, hook] \arrow[d, "h"'] & A \times B \arrow[d, "\nu\times\mu"] \\ \mathrm{graph}(f') \arrow[r, hook] & A' \times B' \end{tikzcd}$$

Let the morphism $f : [\phi(\vec{x})] \to [\psi(\vec{y})]$ in $\mathcal{R}_{\mathcal{T}}$ be defined by the premorphism $\mu(\vec{x}, \vec{y}\,')$. We have that

$$\mathrm{graph}(f) \hookrightarrow [\phi] \times [\psi] \hookrightarrow X \times Y$$

is defined as the subobject $M_0(\mu(\vec{x}, \vec{y}\,'))$. So, by taking the images by I_1 and I_2, we obtain the commutative diagram

$$\begin{tikzcd} \mathrm{graph}(I_1(f)) = M_1(\mu) \arrow[r, hook] \arrow[d] \arrow[d, shift left=8ex, "G_{[\mu]}"] & I_1([\phi] \times [\psi]) \arrow[r, hook] \arrow[d, "G_{[\phi]\times[\psi]} = G_\phi \times G_\psi"] & I_1(X \times Y) \arrow[d, "(F_{x_1}, \ldots, F_{x_n})"] \\ \mathrm{graph}(I_2(f)) = M_2(\mu) \arrow[r, hook] & I_2([\phi] \times [\psi]) \arrow[r, hook] & I_2(X \times Y). \end{tikzcd}$$

By the above fact, this establishes that

$$\begin{tikzcd} I_1([\phi]) \arrow[r, "I_1(f)"] \arrow[d, "G_{[\phi]}"'] & I_1([\psi]) \arrow[d, "G_{[\psi]}"] \\ I_2([\phi]) \arrow[r, "I_2(f)"'] & I_2([\psi]) \end{tikzcd}$$

is commutative, showing that G is indeed a natural transformation.

We have shown that $(M_0)^*_{\mathcal{S}}$ is full. In fact, we have seen that in the above proof G was uniquely determined that shows the faithfulness of $(M_0)^*_{\mathcal{S}}$.

We have completed the proof of 8.2.4.

To finish this section, notice that in 8.2.1, 8.2.3 and 8.2.4 we established 8.1.3 and 8.1.4.

§3 Infinitary generalizations

Let κ be a regular cardinal number. We denote by $L_{\kappa\omega}$ the fragment of $L_{\infty\omega}$ consisting of the formulas that only use conjunction and disjunction of sets Σ such that $\mathrm{card}(\Sigma) < \kappa$.

Let $L^g_{\kappa\omega}$ denote the coherent part of $L_{\kappa\omega}$, $L^g_{\kappa\omega} = L^g_{\infty\omega} \cap L_{\kappa\omega}$, and let $\mathcal{T}$ be a theory in $L^g_{\kappa\omega}$. With such theories, we can associate a *κ-logical* category $\mathcal{R}_\kappa(\mathcal{T})$, in much the same way we associated $\mathcal{R}(\mathcal{T})$ with $\mathcal{T}$ in the previous section.

The category $\mathcal{R}$ is called *κ-logical* if it is logical and in addition, it has stable sups of families $\langle X_i \to A : i \in I \rangle$ of subobjects of cardinality $< \kappa$, (i.e., $\operatorname{card} I < \kappa$) (c.f. Chapter 3, Section 4).

A functor $I : \mathcal{R} \to \mathcal{S}$ between *κ-logical* categories is *κ-logical* if it is logical and preserves all κ-sups. Such an I will also be called an $\mathcal{S}, \kappa$-model of $\mathcal{R}$.

If $\kappa = \aleph_0$, we obtain the ordinary "logical" notions we discussed in the last two sections. On the other hand, we can altogether drop the restriction on the size of the disjunctions and the sups of subobjects. In this way we obtain ∞-logical categories and functors. In the sequel, we allow κ to be either an infinite regular cardinal or ∞. For either kind of κ, form Chapter 3, Section 5, we know that for every κ-logical $\mathcal{R}$, there is a theory $T = T_{\mathcal{R}}$ in $L^g_{\kappa\omega}$, for $L = L_{\mathcal{R}}$ the standard language associated with $\mathcal{R}$, such that the $\mathcal{S}$-models of T are exactly the κ-logical functors $\mathcal{R} \to \mathcal{S}$, for any κ-logical $\mathcal{S}$.

We have the following theorem with is a converse to the last fact.

Theorem 8.3.1 (i) *Let κ be an infinite regular cardinal, or $\kappa = \infty$. For any theory $\mathcal{T}$ in $L_{\kappa\omega}$, there is a κ-logical category $\mathcal{R} = \mathcal{R}_\kappa(\mathcal{T})$ together with an $\mathcal{R}$-model $M_0 : \mathcal{T} \to \mathcal{R}$ of $\mathcal{T}$ such that: for any $M : \mathcal{T} \to \mathcal{S}$, an $\mathcal{S}$-model for an arbitrary κ-logical category $\mathcal{S}$ there is a κ-logical functor $I : \mathcal{R} \to \mathcal{S}$ such that*

$$\begin{tikzcd} \mathcal{T} \arrow[r, "M_0"] \arrow[dr, "M"'] & \mathcal{R}_\kappa(\mathcal{T}) \arrow[d, "I"] \\ & \mathcal{S} \end{tikzcd}$$

commutes; I is determined uniquely up to a unique isomorphism. Also, $\mathcal{R}$ can be taken to be a small category, even for $\kappa = \infty$, whenever $\mathcal{T}$ is a set (as opposed to being a proper class).

(ii) *In fact, we have that for any κ-logical $\mathcal{S}$*

$$(M_0)^*_{\mathcal{S}} : \operatorname{Mod}_{\mathcal{S},\kappa}(\mathcal{R}) \to \operatorname{Mod}_{\mathcal{S}}(\mathcal{T})$$

is an equivalence of categories.

Remark *Here, of course, $\operatorname{Mod}_{\mathcal{S},\kappa}(\mathcal{R})$ is the full subcategory of the functor category $\mathcal{S}^{\mathcal{R}}$ whose objects are the $\mathcal{S}, \kappa$-models of $\mathcal{R}$. $\operatorname{Mod}_{\mathcal{S}}(\mathcal{T})$ is the full subcategory of all $\mathcal{S}$-interpretations of the language of $\mathcal{T}$ whose objects are models of $\mathcal{T}$. $(M_0)^*_{\mathcal{S}}$ is defined from $M_0 : \mathcal{T} \to \mathcal{R}$ by composition in a natural way.*

The proof is a natural extension of the proofs outlined in the last section (which establish 8.3.1 for $\kappa = \aleph_0$). We only indicate a few important points.

First, let us define the class $\mathcal{O}_\kappa$ of *simple κ-formulas* of $L_{\infty\omega}$ as consisting of disjunctions

$$\bigvee_{i \in I} \gamma_i = \bigvee\{\gamma_i : i \in I\}$$

such that $\operatorname{card}(I) < \kappa$ and such that each γ_i is a primitive formula, i.e. it is of the form $\exists x_1 \cdots x_k \bigwedge_{j=1}^m \theta_j$ with *finite* $k, m < \omega$ and with each θ_j a simple atomic formula of L (c.f. §2). (For $\kappa = \infty$, there is no restriction on $\operatorname{card}(I)$.) Since the cardinality of all primitive formulas based on the language L is $\lambda = \max(\aleph_0, \operatorname{card}(L))$, the set of all simple ∞-formulas of $L_{\infty\omega}$ is 2^λ. Also, for $\kappa > \lambda$, "κ-simple" is the same as "∞-simple".

The point of simple κ-formulas is that every formula in $L^\gamma_{\kappa\omega}$ is logically equivalent to a simple κ-formula.

Proposition 8.3.2 *For every ϕ in $L^g_{\kappa\omega}$ there is $\psi \in \mathcal{O}_\kappa$ with the same free variables as ϕ such that*

$$\vdash \phi \Rightarrow \psi \qquad \text{and} \qquad \vdash \psi \Rightarrow \phi$$

$\vdash$ indicates here derivability pin one of the systems of Chapter 5, with T there taken to be the empty set of axioms. (We write $\phi \vdash\dashv \psi$, and we say that ϕ and ψ are logically equivalent.)

Proof. The proof is by induction on the complexity of ϕ. To handle atomic formulas other that simple ones, we note that e.g.

$$f(g(x)) \approx y \vdash\dashv \exists z[g(x) \approx z \wedge f(z) \approx y].$$

This logical equivalence is inferred most easily by the completeness theorem for $\vdash$ (c.f. 5.1.7). Clearly, for any ordinary set-structure, the meanings of the formulas at hand are the same. With a natural extension of the above example, every atomic ϕ can be "turned into" a simple formula.

It remains to handle the three inductive cases $\phi = \exists x\phi_1$, $\phi = \phi_1 \wedge \phi_2$ and $\phi = \bigwedge_{i\in I}\phi_i$.

If $$\phi_1 \vdash\dashv \psi_1$$

then $$\exists x\phi_1 \vdash\dashv \exists x\psi_1$$

as is easily seen; on the other hand, if

$$\psi_1 = \bigvee_{j\in J}\gamma_j \qquad \text{then}$$

$$\exists x\psi_1 \vdash\dashv \bigvee_{j\in J}\exists x\gamma_j.$$

This last fact can be inferred from the Boolean completeness theorem 5.1.2 (and from the two-valued one, 5.1.7, in case $\kappa \le \aleph_1$). This calls for showing that in case

$$\|\phi_1[a,\vec{a}]\|_M = \|\bigvee_{j\in J}\gamma_j[a,\vec{a}]\|_M$$

for any $\mathcal{B}$-valued M, and any a, $\vec{a}$ in $|M|$, then $\|\exists x\ \phi_1[\vec{a}]\|_M = \|(\bigvee_{j\in J}\exists x\ \gamma_j)[\vec{a}]\|_M$; this last fact is an easy consequence of the definition of $\|\cdots\|_M$. The required logical equivalences below can be seen similarly. The equivalence (1) takes care of the inductive case $\phi = \exists x\phi_1$, since the formula on the right is simple.

For $\phi = \phi_1 \wedge \phi_2$

$$\phi_1 \vdash\dashv \psi_1 = \bigvee_{j\in J_1}\gamma^1_j$$

$$\phi_2 \vdash\dashv \psi_2 = \bigvee_{j\in J_2}\gamma^2_j$$

by induction hypothesis. It follows that

$$\phi_1 \wedge \phi_2 \vdash\dashv \psi_1 \wedge \psi_2 \vdash\dashv \bigvee_{(j_1,j_2)\in J_1\times J_2}(\gamma^1_{j_1} \wedge \gamma^2_{j_2}) \tag{2}$$

and if

$$\gamma_1 = \exists\vec{x}\bigwedge_{i=1}^n\theta_i \qquad \gamma_2 = \exists\vec{y}\bigwedge_{j=1}^m\theta'_j$$

then $\gamma_2 \vdash\dashv \exists\vec{z}\bigwedge_{j=1}^m\theta''_j$ for some $\vec{z}$ such that $\vec{x}$ and $\vec{z}$ are disjoint and then

$$\gamma_1 \wedge \gamma_2 \vdash\dashv \exists\vec{x}\ \exists\vec{z}\ \bigwedge\{\theta_1,\ldots\theta_n,\theta''_1,\ldots,\theta''_m\}.$$

Using this fact, in (2) we can replace the conjunctions $\gamma^1_{j_1} \wedge \gamma^2_{j_2}$ by primitive formulas, showing that $\phi_1 \wedge \phi_2$ is logically equivalent to a simple κ-formula.

Finally, for $\phi = \bigvee\Sigma$, $\operatorname{card}\Sigma < \kappa$, by induction hypothesis $\sigma \vdash\dashv \bigvee\Sigma_\sigma$ for every $\sigma \in \Sigma$ and thus $\phi \vdash\dashv \bigvee(\bigcup_{\sigma\in\Sigma}\Sigma_\sigma)$. Since $\bigcup_{\sigma\in\Sigma}\Sigma_\sigma$ is the union of $< \kappa$ many sets of cardinality $< \kappa$, it has cardinality $< \kappa$ since κ is regular.

Observe that the proof of 8.3.2 actually gives the construction of a well-determined simple formula ψ such that $\phi \vdash\dashv \psi$. Let us denote this ψ by $[\phi]$. □

Next we turn to the definition of $\mathcal{R} = \mathcal{R}_\kappa(\mathcal{T})$. The objects of $\mathcal{R}$ are defined to be the simple κ-formulas of $L_{\kappa\omega}$, $\mathcal{T}$ being a theory in $L^\gamma_{\kappa\omega}$. The rest of the definition is an exact replica of the in the last section. Again, we have premorphisms $\mu(\vec{x}\,', \vec{y}\,')$ with μ being a simple κ-formula satisfying the conditions stated before, and we identify premorphisms if the "define the same morphism", provably in T (we use either of the provability relations $T \vdash (\cdot)$ introduced in Chapter 5). Let us denote by $[\phi(\vec{x})] \hookrightarrow X$ the subobject $A \overset{i}{\hookrightarrow} X$ where

$$X = \textstyle\bigwedge_{i=1}^n x_i \approx x_i$$

$$A = [\phi(\vec{x})] = \psi(\vec{x})$$

for the simple κ-formula $\psi(\vec{x})$ constructed in 8.3.2 such that $\phi \vdash\dashv \psi$, i is defined by the premorphism $[\phi(\vec{x})] \wedge \bigwedge_{i=1}^n x_i \approx x_i$. Then we can prove the assertions of 8.2.2 just as before, also using the fact that $\phi \vdash\dashv [\phi]$.

This completes our remarks on the proof of 8.3.1. □

We make two obvious observations on the construction of $\mathcal{R}_\kappa(\mathcal{T})$ that will be useful later. Let $\vec{x} = \langle x_1, \ldots, x_n\rangle$. With $M_0 : \mathcal{T} \to \mathcal{R} = \mathcal{R}_\kappa(\mathcal{T})$ as in 8.3.1, let us write $[\phi]_{\vec{x}} \hookrightarrow X$ $(X = M_0(\vec{x}))$ for $(M_0)_{\vec{x}}(\phi) \overset{i}{\hookrightarrow} X$ with i defined by the premorphism $\bigwedge_{i=1}^n x_i \approx x_i'$.

Proposition 8.3.3 (i) *Every subobject of X in $\mathcal{R}$ is (isomorphic to one) of the form $[\phi]_{\vec{x}} \hookrightarrow X$.*

(ii) *Every object in $\mathcal{R}$ is a subobject $[\phi]_{\vec{x}}$ (with an appropriate monomorphism, actually, with the one in* (i)*) of an object of the form $X = M_0(\vec{x})$, i.e., a finite product of sorts of the language of $\mathcal{T}$.*

§4 The κ-pretopos correlated to a theory

Recall the definition of a κ pretopos (c.f. Chapter 3, Section 4) for κ an infinite regular cardinal, or $\kappa = \infty$.

Theorem 8.4.1 *For any small κ-logical category $\mathcal{R}$ there is a κ-pretopos $\mathcal{P} = \mathcal{P}_\kappa(\mathcal{R})$ and a κ-logical functor $I_0 : \mathcal{R} \to \mathcal{P}$ such that for any κ-logical $I : \mathcal{R} \to \mathcal{P}'$ with an arbitrary κ-pretopos $\mathcal{P}'$ there is κ-logical functor $J : \mathcal{P} \to \mathcal{P}'$ such that the following commutes*

$$\begin{tikzcd} \mathcal{R} \arrow[r, "I_0"] \arrow[dr, "I"'] & \mathcal{P} \arrow[d, "J"] \\ & \mathcal{P}' \end{tikzcd}$$

If $\kappa < \infty$, and $\mathcal{R}$ is small, then $\mathcal{P}_\kappa(\mathcal{R})$ is a small category. In case $\kappa = \infty$, and $\mathcal{R}$ is small, $\mathcal{P}_\kappa(\mathcal{R})$ is a Grothendieck topos, denoted by $\mathcal{E}(\mathcal{R})$.

Theorem 8.4.1′ *For any coherent theory $\mathcal{T}$ in $L^g_{\kappa\omega}$ there is a κ-pretopos $\mathcal{P} = \mathcal{P}_\kappa(\mathcal{T})$ and a model $M_0 : \mathcal{T} \to \mathcal{P}$ such that for any model $M : \mathcal{T} \to \mathcal{P}'$ in a κ-pretopos $\mathcal{P}'$ there* *is a logical functor $J : \mathcal{P} \to \mathcal{P}'$ such that the following commutes*

$$\begin{tikzcd} \mathcal{T} \arrow[r, "M_0"] \arrow[dr, "M"'] & \mathcal{P} \arrow[d, "J"] \\ & \mathcal{P}' \end{tikzcd}$$

J is determined up to isomorphism.

If $\kappa < \infty$, and $\mathcal{T}$ is a set, then $\mathcal{P}_\kappa(\mathcal{T})$ is a small category. In case $\kappa = \infty$, and $\mathcal{T}$ is a set, $\mathcal{P}_\infty(\mathcal{T})$ is a Grothendieck topos, denoted by $\mathcal{E}(\mathcal{T})$, and also it is called the classifying topos of $\mathcal{T}$.

Proofs. By the preceding sections, the two versions are equivalent. Our proof will relate to both versions at the same time. We will prove the first version directly but we will make use of the theory $\mathcal{T}_\mathcal{R}$. Here $\mathcal{T}_\mathcal{R}$ is the theory in $L^g_{\kappa\omega}$ whose $\mathcal{S}$-models are exactly the κ-logical functors $\mathcal{R} \to \mathcal{S}$, for any κ-logical $\mathcal{S}$. This offers a slight notational simplification over treating an arbitrary theory $\mathcal{T}$.

We will define an *extension* of the theory $\mathcal{T}_\mathcal{R}$, $\mathcal{T}_\mathcal{R} \to \mathcal{T}'$, and we will put $\mathcal{P}_\kappa(\mathcal{R}) = \mathcal{P}_\kappa(\mathcal{T}_\mathcal{R}) = \mathcal{R}_\kappa(\mathcal{T}')$.

The construction of $\mathcal{T}'$ takes place in two steps. In the first, we extend $\mathcal{T} = \mathcal{T}_\mathcal{R}$ to $\mathcal{T}_1$ by *formally adjoining sums* and in the second, we extend $\mathcal{T}_1$ to $\mathcal{T}_2$ by *formally adjoining quotients of equivalence relations.*

(A) The coproduct completion of a category

Theorem 8.4.2 *Let $\mathcal{R}$ be a κ-logical category. There is a κ-logical category $\mathcal{R}_1 = \mathcal{R}^\sharp$ together with a κ-logical functor $I_1 : \mathcal{R} \to \mathcal{R}_1$ such that*

(i) *$\mathcal{R}_1$ has disjoint κ-sums, i.e., for any family $\{R_i : i \in J\}$ of $< \kappa$ many objects in $\mathcal{R}_1$, their disjoint sum $\coprod_{i \in J} R_i$ exists in $\mathcal{R}_1$.*

(ii) *$\mathcal{R}_1$ is the solution to the universal problem of finding a κ-logical extension of $\mathcal{R}$ with disjoint κ-sums, i.e., if*

$$\mathcal{R} \xrightarrow{I'} \mathcal{R}'$$

is a κ-logical functor with $\mathcal{R}'$ having disjoint κ-sums then there is κ-logical $F : \mathcal{R}^\sharp \to \mathcal{R}'$ such that

$$\begin{tikzcd} \mathcal{R} \arrow[r, "I'"] \arrow[dr, "I_1"'] & \mathcal{R}' \\ & \mathcal{R}^\sharp \arrow[u, "F"'] \end{tikzcd}$$

commutes; F is determined up to a unique isomorphism.

(iii) *I_1 is conservative, i.e. if $I_1(R_1) \le I_1(R_2)$ for subobjects R_1, R_2 of R in $\mathcal{R}$, then $R_1 \le R_2$.*

(iv) *I_1 is full with respect to subobjects, i.e., for any subobject S of $I_1(R)$ in $\mathcal{R}_1$, there is a subobject $R' \hookrightarrow R$ such that $I(R') \hookrightarrow I(R)$ is isomorphic to $S \hookrightarrow R$. Hence, by* (iii) *too, I_1 is full.*

(v) *Every object in $\mathcal{R}_1$ is isomorphic to a disjoint sum $\coprod_{i \in J} I_1(S_i)$, for objects S_i in $\mathcal{R}$, and with* card$(I) < \kappa$.

Proof. We extend both the language of the theory $\mathcal{T} = \mathcal{T}_\mathcal{R}$ and its set of axioms. Let S_i ($i \in J$, and $J < \kappa$) be arbitrary $< \kappa$ many objects in $\mathcal{R}$. Let us introduce a new sort

S, also denoted by $\coprod_{i\in J} S_i$, and new operation symbols $j_i : S_i \to S$ together with the following axioms

$$\begin{aligned} j_i(x_i) \approx j_i(x'_i) &\Rightarrow x_i \approx x'_i \\ j_i(x_i) \approx j_{i'}(x'_{i'}) &\Rightarrow \qquad\qquad \text{for } i, i' \in J \text{ such that } i \neq i'; \end{aligned}$$

and

$$\Rightarrow \bigvee_{i\in J} \exists x_i \ (x \approx j_i(x_i));$$

here x_i is of sort S_i, x of sort S. These axioms express that S is the disjoint sum of the S_i, with canonical injections j_i. Performing these additions to $\mathcal{T}$ simultaneously for all sets $\{S_i : i \in J\}$ of cardinality $< \kappa$ of objects in $\mathcal{R}$, we obtain a larger language $L_1 \supset L$ (= language of $\mathcal{T}$) and a larger set of axioms $T_1 \supset T$ (= set of axioms of $\mathcal{T}$). Consider $\mathcal{R}_1 = \mathcal{R}(\mathcal{T}_1)$ as constructed in the previous sections. We have the canonical κ-model

$$\mathcal{T}_1 \xrightarrow{M_1} \mathcal{R}_1.$$

Let us denote by $[S]$ the interpretation of the sort S by M_1, $[S] = M_1(S)$ (recall: this in nothing but the formula $S(x)$), and similarly $[j] = M_1(j)$ for any operation symbol j in L_1. We have the inclusion as an interpretation:

$$\mathcal{T}_{\mathcal{R}} \xrightarrow{\text{incl.}} \mathcal{T}_1$$

hence we have the composite

$$\mathcal{T}_{\mathcal{R}} \longrightarrow \mathcal{T}_1 \xrightarrow{M_1 = \text{canonical}} \mathcal{R}_1$$

denoted by

$$\mathcal{T}_{\mathcal{R}} \xrightarrow{M'_1} \mathcal{R}_1.$$

Hence we also have the κ-logical functor I_1 such that the following commutes:

$$\begin{tikzcd} \mathcal{R} \arrow[r, "I_1"] & \mathcal{R}_1 \\ \mathcal{T}_{\mathcal{R}} \arrow[u, "\text{can.}"] \arrow[r, "\text{incl.}"'] \arrow[ur, "M'_1"'] & \mathcal{T}_1 \arrow[u, "\text{can.}"'] \end{tikzcd} \tag{1}$$

Keeping in mind the Lindenbaum-Tarski type construction of $\mathcal{R}_1$, we see that $\mathcal{R}_1$ is obtained by formally adjoining to $\mathcal{R}$ disjoint sums of $< \kappa$ objects of $\mathcal{R}$.

Since $\mathcal{T}_1$ contains the axioms expressing that $S = \coprod_{i\in I} S_i$ is the disjoint sum of the S_i with canonical injections j_i, it follows that in $\mathcal{R}_1$, $[S]$ is the disjoint sum $\coprod_{i\in I}[S_i] = \coprod_{i\in I} I_1(S_i)$ with canonical injections $[j_i]$. In particular, any family of cardinality $< \kappa$ of objects "coming form $\mathcal{R}$", i.e., of the form $I_1(S_i)$, $S_i \in \operatorname{Ob}\mathcal{R}$, has a disjoint sum in $\mathcal{R}_1$.

Proof of (iii) We will apply 7.3.2. The assertion will follow from this theorem once we know that $(I_1)^*_{\mathcal{B}} : \operatorname{Mod}^{\mathcal{B}}_{\kappa} \mathcal{R}_1 \to \operatorname{Mod}^{\mathcal{B}}_{\kappa} \mathcal{R}$ is surjective on objects, for any complete Boolean algebra $\mathcal{B}$. Now, a $\mathrm{Sh}_{\mathcal{B}}$, κ-model of $\mathcal{R}_1$ is essentially the same as a $\mathrm{Sh}_{\mathcal{B}}$-model of $\mathcal{T}_1$ and a model of $\mathcal{R}$ is the same as one of $\mathcal{T}_{\mathcal{R}} = \mathcal{T}$. Moreover, for $N : \mathcal{T}_1 \to \mathrm{Sh}_{\mathcal{B}}$ of $\mathcal{T}_1$, $(I_1)^*_{\mathcal{B}}(N)$ is nothing but the reduct of N to the language $L_{\mathcal{R}}$ (i.e., the result of forgetting all but the structure denoted by symbols in $L_{\mathcal{R}}$. Hence, the required fact can be expressed by saying that every model M of $\mathcal{T}$ has an expansion N (whose reduct is M) which is a model of $\mathcal{T}_1$. This latter fact is seen easily as follows. Let M be a $\mathrm{Sh}_{\mathcal{B}}$-model of $\mathcal{T}$, $M:\mathcal{T}_{\mathcal{R}} \to \mathrm{Sh}_{\mathcal{B}}$. To define the expansion N of M, let $S = \coprod_{i\in I} S_i$ be a new sort and let j_i be the corresponding new operation symbols $j_i : S_i \to S$. We have to specify $N(S)$ and the $N(j_i)$; of course, we are given that $N(S_i) = M(S_i)$. Since $\mathrm{Sh}_{\mathcal{B}}$ has disjoint sums, we can define $N(S)$ to be a disjoint sum of the $M(S_i)$ in $\mathrm{Sh}_{\mathcal{B}}$, with the $N(j_i)$ the canonical injections $N(j_i) : M(j_i) \to N(S) = \coprod_{i\in I} M(S_i)$. By these choices, we have made sure that N will satisfy the additional axioms expressing the disjoint-sum property of S. This completes the description of the expansion N of M that is a model of $\mathcal{T}_1$.

Proof of (iv) For the purposes of san induction, we have to formulate a more elaborate statement to prove. To save on notation for this proof we will regard I_1 as an inclusion, i.e., we write S for $I_1(S)$ ($S \in \mathrm{Ob}(\mathcal{R})$) and f for $I_1(f)$ ($f : S_1 \to S_2$ in $\mathcal{R}$). Similarly, we identify symbols in the language L_1 (giving rise to $\mathcal{R}_1$) with their canonical interpretations in $\mathcal{R}_1$. Any *sort* of the language L_1 is of the form $\coprod_{i\in I} S_i$ for $S_i \in \mathrm{Ob}(\mathcal{R})$; namely, those coming from L, the objects of $\mathcal{R}$, can be regarded as one elements sums. Let $X = S^1 \times\cdots\times S^n$ be a finite product of sorts, and $S^k = \coprod_{i\in J_k} S^k_i$, $S^k_i \in \mathrm{Ob}(\mathcal{R})$. We have the following familiar identity (which is a consequence of the definition of disjoint sums and that of products)

$$X = S^1 \times \cdots \times S^n = \textstyle\coprod_{\varepsilon\in \times_{k=1}^n J_k} \times_{k=1}^n S^k_{\varepsilon(k)}. \tag{2}$$

More precisely, we have the canonical maps

$$X_\varepsilon \underset{\mathrm{df}}{=} \times_{k=1}^n S^k_{\varepsilon(k)} \xrightarrow[j_\varepsilon \underset{\mathrm{df}}{=} \times_{k=1}^n j^k_{\varepsilon(k)}]{} \times_{k=1}^n S^k \tag{3}$$

where $j^k_i : S^k_i \to S^k$ is the canonical injection; (2) is understood to mean that X is a disjoint sum of the X_ε, $\varepsilon \in J \underset{\mathrm{df}}{=} \times_{k=1}^n J_k$, with canonical injections in (3). Now, let $[\phi]_{\vec{x}} \hookrightarrow X$ be a typical subobject of X, using the notation of 8.3.3; hence ϕ is a formula of $(L_1)^g_{\kappa\omega}$ with free variables at most $\vec{x}$; $\vec{x} = \langle x_1,\ldots,x_k\rangle$ have the respective sorts $S_1,\ldots,S^k$. Using the above notations, consider the meet of the two subobjects

$$X_\varepsilon \xhookrightarrow{j_\varepsilon} X \qquad \text{and} \qquad [\phi]_{\vec{x}} \hookrightarrow X$$

and denote it by $X_\varepsilon \cap [\phi]_{\vec{x}} \hookrightarrow X$. Of course, this factors through $X_\varepsilon \hookrightarrow X$ and we obtain the subobject $X_\varepsilon \cap [\phi]_{\vec{x}} \hookrightarrow X_\varepsilon$. We are going to prove by an induction on the complexity of ϕ that this last subobject *is in* $\mathcal{R}$, for every $\varepsilon \in J$, i.e., it is isomorphic to a monomorphism $A \hookrightarrow X_\varepsilon$ (notice that $X_\varepsilon \in \mathrm{Ob}(\mathcal{R})$).

We establish the claim by first showing if for a *simple* atomic formula of L_1, and then by performing three inductive steps corresponding to $\wedge$, $\bigvee$ and $\exists$. This will suffice since every formula is logically equivalent to one built up from simple atomic formulas using $\wedge$, $\bigvee$ and $\exists$; (namely, to a simple κ-formula).

Every simple atomic formula of L_1 is of the form $f(x) \approx y$. Either f is a morphism in $\mathcal{R}$, or $f = j$ for a canonical injection $j : S \to \coprod_{i\in J} S_i = S'$ introduced into L_1 such that $S = S_{i_0}$ for some $i_0 \in J$.

We leave the case when f is in $\mathcal{R}$ to the reader. If $f = j$, x is of sort S, y is of sort S', $\phi = f(x) \approx y$ and $\vec{x} = \langle x, y\rangle$, then $X = S \times \coprod_{i\in J} S_i$ where S is treated as a one element sum, $S = \coprod_{j\in\{i_0\}} S_j$. If $\varepsilon \in \{i_0\}\times J$ is such that $\varepsilon(1) = i_0$ and $\varepsilon_2 = i \neq i_0$, then (as it is easy to see) $X_\varepsilon \cap [\phi]_{\vec{x}} \hookrightarrow X_\varepsilon$ is the empty (zero) subobject, and if $\varepsilon(2) = i_0$, then $X_\varepsilon \cap [\phi]_{\vec{x}} \hookrightarrow X_\varepsilon = S_{i_0} \times S_{i_0}$ is isomorphic to the diagonal $\Delta_{S_{i_0}} \hookrightarrow S_{i_0} \times S_{i_0}$. The slightly more general case when $\vec{x}$ includes, but is more than, $\langle x, y\rangle$ is left to the reader.

Next we turn to the inductive cases. If $\phi = \phi_1 \wedge \phi_2$ then (as it is easy to see) the subobject $X_\varepsilon \cap [\phi]_{\vec{x}} \hookrightarrow X_\varepsilon$ is the meet of the corresponding ones made of ϕ_1 and ϕ_2, so the required induction inference can be made. If $\phi = \bigvee_{i \in K} \phi_i$, then as one can verify, we have

$$X_\varepsilon \cap [\phi]_{\vec{x}} \hookrightarrow X_\varepsilon = \bigvee_{i \in K} (X_\varepsilon \cap [\phi_i]_{\vec{x}} \hookrightarrow X_\varepsilon).$$

Since $\mathcal{R}$ has κ-sups, and I_1 (= inclusion) preserves κ-sups, we are again done. Finally, for $\phi = \exists x \phi'(x, \vec{x})$, we have the following situation. The induction hypothesis refers to

$$[\phi']_{x,\vec{x}} \hookrightarrow S^0 \times X$$

where S^0 is the sort of x, $S^0 = \coprod_{i \in J_0} S_i^0$. For a given $\varepsilon in J = \times_{k=1}^n J_k$ (referring to $[\phi]_{\vec{x}} \hookrightarrow X$ as before) and for an index $i \in J_0$, let $i^\frown \varepsilon$ denote the function $\delta \in \times_{k=0}^n J_k = J_0 \times J$ such that $\delta(0) = i$ and $\delta(k) = \varepsilon(k)$ for $k = 1, \dots, n$. So, the induction hypothesis says that the subobject

$$(S_i^0 \times X_\varepsilon) \cap [\phi']_{x,\vec{x}} \hookrightarrow S_i^0 \times X_\varepsilon$$

is always in $\mathcal{R}$, using the above notations. Let π_i be the canonical projection $S_i^0 \times X_\varepsilon \to X_\varepsilon$. Now, we can check the following identity:

$$X_\varepsilon \cap [\phi]_{\vec{x}} \hookrightarrow X_\varepsilon = \bigvee_{i \in J_0} \exists_{\pi_i} ((S_i^0 \times X_\varepsilon) \cap [\phi']_{x,\vec{x}} \hookrightarrow S_i^0 \times X_\varepsilon).$$

Using the induction hypothesis, the identity and the properties of $\mathcal{R}$ and I_1, the assertion to be proved follows.

This completes the proof of our claim on the special kinds of subobjects. Notice that if we take X to be a single object in $\mathcal{R}$, the claim simply says that every subobject $[\phi]_{\vec{x}} \hookrightarrow X$ is actually *in* $\mathcal{R}$. Taking into account 8.3.3 too, this shows that I_1 (treated as the inclusion) is full with respect to subobjects, proving (iv). The fullness of I_1 is a consequence of (iii) and (iv).

Remark *In this proof, we have used server identities involving sups, disjoint sums, products and images. They are easy to verify for an arbitrary "sufficiently" logical category $\mathcal{R}$, but as a matter of fact, it is sufficient to verify them in the category of $\mathcal{B}$-valued sets,* $\mathrm{Sh}_{\mathcal{B}}$*, for an arbitrary Boolean algebra $\mathcal{B}$ (and in case the sups involved are at most countable, in the category of sets,* **Set***). The reason is that then the identities will follow in $\mathcal{R}$, by using the completeness and soundness theorems. Namely, the fact that a particular such identity holds in $\mathcal{R}$ is equivalent to the truth of one or more Gentzen sequents σ_i in the language $L_{\mathcal{R}}$, in the canonical interpretation $\mathcal{R}$ of $L_{\mathcal{R}}$, or by the soundness theorem, it is equivalent to the fact that $T_{\mathcal{R}} \vdash \sigma_i$ $(i = 1, 2, \dots)$ for a theory $T_{\mathcal{R}}$ all of whose axioms are true in $\mathcal{R}$. To infer $T_{\mathcal{R}} \vdash \sigma_i$, we deduce $T_{\mathcal{R}} \models \sigma_i$, for ordinary set models, say, in the case when all sups involved are at most countable. But inspection shows that $T_{\mathcal{R}} \models \sigma_i$ $(i = 1, 2, \dots)$ will be a consequence that the identity in question holds in* **Set***. This remark justifies the customary procedure of verifying certain facts in* **Set***, and then generalizing it to arbitrary toposes, say. Actually, there is a more involved general principle (based on the so-called Levi absoluteness theorem) that justifies considering just* **Set** *in certain cases even if uncountable sups are involved.*

Proof of (v) The proof is essentially contained in the computations made for (iv) above. First of all, every object is of the form $[\phi]_{\vec{x}}$ (c.f. 8.3.3). Using the notations of the previous proof, it is easy to see that the object $[\phi]_{\vec{x}}$ is isomorphic to

$$\coprod_{\varepsilon \in J} X_\varepsilon \cap [\phi]_{\vec{x}}.$$

Since the object $X_\varepsilon \cap [\phi]_{\vec{x}}$ is isomorphic to one in $\mathcal{R}$ as we have shown above, we obtain what we want.

Using (v) and the fact that $< \kappa$ many objects coming from $\mathcal{R}$ do have disjoint sums, we obtain that $\mathcal{R}_1$ has disjoint κ-sums, i.e. the assertion in (i).

The proof of (ii) is easy an it is left to the reader. The assertion of (ii) is intuitively clear since when constructing the extension $\mathcal{R}_1$ we did not do more that absolutely necessary to have disjoint $< \kappa$-sums in $\mathcal{R}_1$. □

(B) The exact completion of a category

Theorem 8.4.3 *Let $\mathcal{R}_1$ be a logical category (no infantry hypothesis is used). There is a logical category $\mathcal{R}_2 = (\mathcal{R}_1)^{\mathrm{ex}}$ together with a logical functor $I_2 : \mathcal{R}_1 \to \mathcal{R}_2$ such that*

(i) *$\mathcal{R}_2$ has quotients by equivalence relations, c.f. Definition 3.3.6.*

(ii) *$\mathcal{R}_2$ is the solution to the universal problem of finding a logical extension of $\mathcal{R}_1$ such that quotients by equivalence relations exist (a detail statement would look like that in* 8.4.2(ii)*),*

(iii) *I_2 is conservative,*

(iv) *I_2 is full with respect to subobjects and full.*

(v) *Every object in $\mathcal{R}_2$ is isomorphic to a quotient $I_2(A)/I_2(R)$, for some $A \in \mathrm{Ob}(\mathcal{R}_1)$ and some equivalence relation $R \hookrightarrow A \times A$ in $\mathcal{R}_1$.*

(vi) *If $\mathcal{R}_1$ is κ-logical, then $\mathcal{R}_2$ is κ-logical and I_2 is a κ-logical functor.*

(vii) *If $\mathcal{R}_1$ is κ-logical and has disjoint κ-sums, then $\mathcal{R}_2$ is a κ-pretopos.*

Proof. Let us denote by L_1 the canonical language associated with $\mathcal{R}_1$, and let $\mathcal{T}_1 = (T_1, (L_1)^{\gamma}_{\omega\omega})$ be the "internal theory" of $\mathcal{R}_1$ such that models of $\mathcal{T}_1$ are exactly the logical functors from $\mathcal{R}_1$. We are going to extend L_1 to L_2 and T_1 to T_2, as follows. Let $R \overset{f}{\hookrightarrow} A \times A$ be an arbitrary equivalence relation on A, all in $\mathcal{R}_1$. We associate with $R \to A \times A$ a new sort, denoted by A/R, and a new operation symbol $p : A \to A/R$ (this notation indicates the sorting of p). We also add the following axioms to T_1:

$$\Rightarrow \exists a[x \approx p(a)] \qquad (x \text{ is a variable of sort } A/R)$$

$$\underset{\sim}{R}(a, a') \Rightarrow p(a) \approx p(a')$$

$$p(a) \approx p(a') \Rightarrow \underset{\sim}{R}(a, a')$$

(as usual, $\underset{\sim}{R}(a, a')$ stands for

$$\exists r[a \approx \pi_1 \rho(r) \wedge a' \approx \pi_2 \rho(r)]$$

where $A \times A \underset{\pi_2}{\overset{\pi_1}{\rightrightarrows}} A$ are the canonical projections). Performing simultaneously these additions to the language and to the set of axioms, we arrive at the new theory $\mathcal{T}_2 = (T_2, L_2)$. (It goes without saying that we take care that for different pairs (A, R) of date, the derived symbols A/R, p should also be always distinct.)

We put $\mathcal{R}_2 = \mathcal{R}(\mathcal{T}_2)$. As before, we have the commutative diagram

$$\begin{tikzcd}
\mathcal{R}_1 \arrow[r, "I_2"] & \mathcal{R}_2 \\
\mathcal{T}_1 \arrow[u, "\text{can.}"] \arrow[ur, "M_2'"] \arrow[r, "\text{incl.}"'] & \mathcal{T}_2 \arrow[u, "\text{can.}=M_2"']
\end{tikzcd}$$

with I_2 a logical functor.

Since I_2 is logical, it preserves equivalence relations. Therefore, if R is an equivalence relation on A (in $\mathcal{R}$), then $[R] = M_2(R) = I_2(R)$ is an equivalence relation on $[A] = I_2(A)$ and, because of the axioms put into $\mathcal{T}_2$, $[A/R] = M_2(A/R)$ is the quotient: $[A/R] = I_2(A)/I_2(R)$, with canonical surjections $[p] = M_2(p) : I_2(A) \to [A/R]$. In particular, quotients of equivalence relations coming from $\mathcal{R}$ exist.

The proof of (iii) is just like that of 8.4.2(iii) and uses only the fact that **Set** has quotients by equivalence relations.

Proof of (iv) We will use 7.1.4 and thereby simplify the arguments considerably (a similar simplification in 8.4.2(iv) is available in the finitary case, $\kappa = \aleph_0$). The required conclusion will follow once we have shown the following. Let $N_1, N_2 : \mathcal{R}_2 \to \mathbf{Set}$ be models (logical functors) and let $F : I_2^*(N_1) = M_1 \to I_2^*(N_2) = M_2$ be a natural transformation (= homomorphism); then there is $G : N_1 \to N_2$ such that $I_2(F) = G$. Instead of talking about models of $\mathcal{R}_2$, we can equivalently talk about models of the theory $\mathcal{T}_2$. Also, M_i becomes the reduct of N_i to the sublanguage $L_1 \subset L_2$.

Even more importantly, by 8.2.4 it is sufficient to extend F to a homomorphism $G : N_1 \to N_2$ between N_1 and N_2 as structures of the similarity type L_2. We extend $F : M_1 \to M_2$ to $G : N_1 \to N_2$ as follows. Let $S = A/R$ be a typical new sort and $p : A \to A/R$ the corresponding "formal surjection" in the language L_2. We have to define $G_S : N_1(S) \to N_2(S)$ such that

$$\begin{tikzcd}
M_1(A) = N_1(A) \arrow[r, "N_1(p)"] \arrow[d, "F_A"'] & N_1(A/R) \arrow[d, "G_S = G_{A/R}"] \\
M_2(A) = N_2(A) \arrow[r, "N_2(p)"'] & N_2(A/R)
\end{tikzcd}$$

commutes, for all A and R as above. (Once we have done this, we have defined G as an extension of F and we have verified that G has the homomorphism property with respect to the new operation symbols p, hence, with respect to all symbols in L_2.) Now, use the fact that since N_1, N_2 are models of T_2 in **Set**, $N_k(S)$ is the quotient by the equivalence relation $N_k(R) \hookrightarrow N_k(A) \times N_k(A)$, with the canonical surjection $N_k(p)$. To show that the map G_S exists making the above diagram commutative is straightforward but here are the details. Let $A_k = N_k(A)$, $p_k = N_k(p)$, $R_k = N_k(R)$ and $f = F_A$; remember that we are in the category of sets. Let us make the simplifying assumption that for $R \xrightarrow{\rho} A \times A$, $N_k(\rho)$ is the inclusion, $R_k \subset A_k \times A_k$. Using now the part F_R of the homomorphism F, we see that $\langle a, a' \rangle \in A_1 \Rightarrow \langle f(a), f(a') \rangle \in A_2$. If $p_1(a) = p_1(a')$, then $\langle a, a' \rangle \in A_1$ by the definition of the quotient, hence $\langle f(a), f(a') \rangle \in A_2$, and thus $p_2(f(a)) = p_2(f(a'))$ in A_2/R_2, again by the definition of quotient. Since p_1 is surjective, this says that the function $p_1(a) \mapsto p_2(f(a))$ is well defined on the whole of $A_1/R_1 = N_1(S)$. This function can be taken to be G_S and the above diagram will commute.

This shows that I_2^* is full, hence by 7.1.4 I_2 is full with respect to subobjects.

Note that parts (iii) and (iv) of the theorem say that I_2 induces an isomorphism between the subobject lattices of A and $I_2(A)$, for any object A of $\mathcal{R}_1$.

Proof of (iv) Every object in $\mathcal{R}_2$ that is the M_2-interpretation of a sort in L_2 can be represented as a quotient $I_2(A)/I_2(R)$ where $R \hookrightarrow A \times A$ is an equivalence relation in $\mathcal{R}_1$. Hence, by the observation 8.3.3, every object B in $\mathcal{R}_2$ is a subobject of a product of such quotients

$$B \hookrightarrow \bigtimes_{i=1}^{n} I_2(A_i)/I_2(R_i).$$

Use now the formula

$$\times_{i=1}^{n}(\overline{A_i}/\overline{R_i}) = \times_{i=1}^{n}\overline{A_i}/\times_{i=1}^{n}\overline{R_i}$$

(c.f. also the Remark after the proof of 8.4.2(iv)) and the fact that

$$\times_{i=1}^{n}\overline{R_i} \hookrightarrow (\times_{i=1}^{n}\overline{A_i}) \times (\times_{i=1}^{n}\overline{A_i})$$

is an equivalence relation (we put $\overline{A} = I_2(A)$, etc.). This reduces the question to the case when B is a subobject of a quotient

$$B \hookrightarrow \overline{A}/\overline{R}.$$

Consider the pullback diagrams:

$$\begin{tikzcd}
C \arrow[r, hook] \arrow[d] \arrow[dr, phantom, "\text{p.b.}"] & \overline{A} \arrow[d, "p=\text{canonical}"] \\
B \arrow[r, hook] & \overline{A}/\overline{R}
\end{tikzcd}$$

and

$$\begin{tikzcd}
R' \arrow[r, hook] \arrow[d] & C \times C \arrow[d, "\gamma\times\gamma"] \\
\overline{R} \arrow[r, hook] & \overline{A} \times \overline{A}
\end{tikzcd}$$

It is easy to see that $R' \hookrightarrow C \times C$ is an equivalence relation on C, and $B \simeq C/R$. Using part (iv), we have $C \simeq I_2(C_0)$ for some $C_0 \in \mathrm{Ob}(\mathcal{R}_1)$, and without loss of generality, we can assume $C = I_2(C_0)$. Using (iv) again, $R' \hookrightarrow I_2(C_0) \times I_2(C_0)$ is of the form $R' = I_2(R'')$. By conservativeness (iii), R'' is an equivalence relation on C_0 and we have $B \simeq I_2(C_0)/I_2(R'')$, completing the proof of (v).

Proof of (i) This is an easy consequence of (v). Using (v), it is enough to show the following. If $R' \hookrightarrow (I_2(A)/I_2(R))^2$ is an equivalence relation on $I_2(A)/I_2(R)$, then $(I_2(A)/I_2(R))/R'$ exists in $\mathcal{R}_2$. But it is easy to see that $(I_2(A)/I_2(R))/R' = I_2(A)/I_2(R'')$ where

$$\begin{tikzcd}
I_2(A \times A) \arrow[r, "\text{can. surj.}"] \arrow[dr, phantom, "\text{p.b.}"] & (I_2(A)/I_2(R))^2 \\
\overline{R} \arrow[u, hook] \arrow[r] & R' \arrow[u, hook]
\end{tikzcd}$$

and the equivalence relation R'' is such that $\overline{R} = I_2(R'')$ (by (iv)).

Proof of (vii) The required fact is that now $\mathcal{R}_2$ has disjoint κ-sums. This is a consequence of part (v) and the identity

$$\textstyle\coprod_{i\in J} A_i/R_i = \coprod_{i\in J} A_i / \coprod_{i\in J} R_i$$

(with the obvious injection $\coprod_{i\in J} R_i \to (\coprod_{i\in J} A_i)^2$).

We have completed the proof of 8.4.3. □

Theorems 8.4.2 and 3 not only prove 8.4.1, but give considerable additional information on the pretopos generated by $\mathcal{R}$.

Theorem 8.4.4 *For any κ-logical category $\mathcal{R}$, there is a κ-pretopos $\mathcal{P} = \mathcal{P}_\kappa(\mathcal{R})$ and a κ-logical functor $I_0 : \mathcal{R} \to \mathcal{P}$ such that*

(i) $\mathcal{P}_\kappa(\mathcal{R})$ *satisfies* 8.4.1.

(ii) I_0 *induces an isomorphism of the subobject lattices of* R *and* $I_0(R)$, *for every object* R *of* $\mathcal{R}$,

(iii) *every object of* $\mathcal{P}$ *is isomorphic to one of the form* $(\coprod_{i\in J} I_0(S_i))/R$, *for some objects* S_i *in* $\mathcal{R}$ *and for some equivalence relation* $R \hookrightarrow (\coprod_{i\in J} I_0(S_i))^2$; *here* card $J < \kappa$.

(iv) *The equivalence relations* R *on an object of the form* $\coprod_{i\in J} I_0(S_i)$ *are exactly the ones that can be obtained as follows: consider subobjects* $R_{ij} \hookrightarrow S_i \times S_j$ *in* $\mathcal{R}$; *assume the following are true in* $\mathcal{R}$:

$$\text{“each } R_{ij} \text{ is reflexive an symmetric”}$$

$$\underset{\sim}{R}_{ij}(s_i, s_j) \wedge \underset{\sim}{R}_{jk}(s_j, s_k) \Rightarrow \underset{\sim}{R}_{ik}(s_i, s_k);$$

construct R *as*

$$R = \coprod_{\langle i,j\rangle\in J^2} I_0(R_{ij}) \hookrightarrow \coprod_{\langle i,j\rangle\in J^2} I_0(S_i) \times I_0(S_j) \simeq (\coprod_{\langle i,j\rangle\in J^2} I_0(S_i))^2.$$

The $R_{ij} \hookrightarrow S_i \times S_j$ *are uniquely determined by* $R \hookrightarrow (\coprod I_0(S_i))^2$.

(v) *The morphisms* f *in* $\mathcal{P}$ *of the form:*

$$(\coprod_{i\in J_1} I_0(S_i^1))/R^1 \xrightarrow{f} (\coprod_{i\in J_2} I_0(S_i^2))/R^2$$

are exactly the ones that are obtained as follows:

consider subobjects $F_{ij} \hookrightarrow S_i \times S_j$ in $\mathcal{R}$ for $i \in J_1, j \in J_2$

and consider $R^1_{ii'} \hookrightarrow S^1_i \times S^1_{i'}$ $\quad R^2_{jj'} \hookrightarrow S^2_j \times S^2_{j'}$

giving rise to R^1 *and* R^2, *respectively, as in (iv); assume the following are true in* $\mathcal{R}$:

$$\underset{\sim}{R}^1_{ii'}(s^1_i, s^1_{i'}) \wedge \underset{\sim}{F}_{ij}(s^1_i, s^2_j) \wedge \underset{\sim}{F}_{i'j'}(s^1_{i'}, s^2_{j'}) \Rightarrow \underset{\sim}{R}_{jj'}(s^2_j, s^2_{j'})$$

for $i, i' \in J_1$ *and* $j, j' \in J_2$,

$$\Rightarrow \bigvee_{j\in J_2} \exists s^2_j F_{ij}(s^1_i, s^2_j)$$

for any $i \in J_1$;

construct F *as the image*

$$\exists_{p\times q}(\coprod_{\langle i,j\rangle\in J_1\times J_2} I_0(F_{ij}))$$

with

$$A = \coprod_i I_0(S^1_i)$$

$$B = \coprod_i I_0(S^2_i)$$

$$A \xrightarrow{p=\text{ can.}} A/R^1$$

$$B \xrightarrow{q=\text{ can.}} B/R^2$$

$$\begin{tikzcd} \coprod_{i,j} I_0(F_{ij}) \arrow[r, hook] \arrow[d] & A \times B \arrow[d, "p\times q"] \\ \exists_{p\times q}(\coprod_{i,j} I_0(F_{ij})) \arrow[r, hook] & A/R^1 \times B/R^2 \end{tikzcd}$$

and finally, construct f as the morphism whose graph is F. The F_{ij} are uniquely determined by f, R^1, R^2.

Remarks 8.4.4 *give a complete description of $\mathcal{P}(\mathcal{R})$ as obtained from $\mathcal{R}$. The mere existence of $\mathcal{P}(\mathcal{R})$ satisfying* 8.4.1 *is more or less a consequence of general principles, but the specific properties listed in* 8.4.4 *are not. E.g., a priori it might have happened that the addition of disjoint sums and quotients had to be repeated infinitely many times to get a pretopos. (iii) tells us that two such steps are sufficient.*

Proof of 8.4.4 Start with an arbitrary κ-logical category $\mathcal{R}$. Construct $\mathcal{R}_1$ as in (A), Theorem 8.4.2 and the construct $\mathcal{R}_2$ from $\mathcal{R}_1$ as in (B), Theorem 8.4.3. Define $\mathcal{R} \xrightarrow{I_0} \mathcal{P} = \mathcal{R}_2$ as the composite

$$\mathcal{R} \xrightarrow{I_1} \mathcal{R}_1 \xrightarrow{I_2} \mathcal{R}_2 = \mathcal{P}.$$

The fact that $\mathcal{P}$ is a κ-pretopos, moreover (i), (ii) and (iii) are obvious consequences of 8.4.2, and 8.4.3. Finally, (iv) and (v) are consequences of analysis based entirely on (ii) and (iii); we omit the details. $\square$

We would like to emphasize the description of certain categories appearing in algebraic geometry that results from 8.4.4. We first take the finitary case, $\kappa = \aleph_0$. Let $\mathcal{C}$ be a category with finite $\varprojlim$ and let $\mathcal{C}$ be turned into an algebraic site by imposing the generating collections $\mathrm{Cov}_0(\mathcal{C})$ of finite covering families for any $C \in \mathrm{Ob}(\mathcal{C})$. Let $\mathcal{T}$ be the theory associated with $\mathcal{C}$ as a site (c.f. Chapter 6, Section 1). Let $\mathcal{R} = \mathcal{R}(\mathcal{T})$. Finally, let $\mathcal{P}$ be the pretopos $\mathcal{P}(\mathcal{R}) = \mathcal{P}(\mathcal{T})$. As we are going to point out in the next chapter, $\mathcal{P}$ is nothing but the *category of coherent objects and morphisms in the coherent topos* $\widetilde{\mathcal{C}}$ (= category of sheaves over $\mathcal{C}$). To recapitulate the description of $\mathcal{P}$ that we have obtained, first recall the class $\mathcal{O}$ of *simple* formulas (in the language having the objects of $\mathcal{C}$ as sorts, and the morphisms of $\mathcal{C}$ as operation symbols) that we singled out as sufficient to define the objects of $\mathcal{R}(\mathcal{T})$. Then a syntactic representation of $\mathcal{P}$ is the following:

(i) consider simple formulas S_i $(i = 1, \dots, n)$ and $\underset{\sim}{R}_{ij}$ such that the sequents in 8.4.4(iv) are consequences of T in the formal sense "$\vdash$". From the formal entity

$$(\textstyle\coprod_{i=1}^n S_i)/(\coprod_{i=1}^n R_i) = A/R;$$

all objects of $\mathcal{P}$ are of this form.

(ii) continuing with the notation of (i), add a new morphism (operation symbol) p:

$$A \xrightarrow{p} A/R,$$

add new morphisms (operation symbols):

$$S_i \xrightarrow{j_i} \textstyle\coprod_{i=1}^n S_i$$

$$R_{ij} \xrightarrow{h_{ij}} \textstyle\coprod_{i,j} R_{ij}$$

and add the 'axioms' related to disjoint sums and quotients as *defining relations* (the precise meaning of this is incorporated in the definition of $\mathcal{R}(\cdot)$; in our case, the pretopos $\mathcal{P}$ is obtained as $\mathcal{R}(\mathcal{T}')$ for a suitable $\mathcal{T}'$ as described above)

(iii) the representation of a general morphism of $\mathcal{P}$ is obtained following 8.4.4(v) in particular, it uses simple formulas F_{ij}, and finally,

(iv) two representations of morphisms, with the same domain and codomain presented as in (i) are *identified* to define the same morphism if and only if the corresponding simple formulas F_{ij}, F'_{ij} are provably equivalent in $\mathcal{T}'$ (for each pair of indices i, j); here $\mathcal{T}'$ is $\mathcal{T}$ together with the new axioms mentioned in (iii).

This description, supplemented with the description when it is true that $T \vdash \sigma$ for a sequent σ, provides a full presentation of $\mathcal{P}(\mathcal{C}) = \mathcal{P}$. Having in mind a suitable notion "primitive recursive presentation", we have

Metatheorem 8.4.5 For an algebraic site $\mathcal{C}$, the pretopos $\mathcal{P}$ of coherent objects and morphisms of $\widetilde{\mathcal{C}}$ is presented primitive recursively *in* $\mathcal{C}$. In particular, if $\mathcal{C}$ is recursively presented, so is $\mathcal{P}$.

At another extreme, we can take $\kappa = \infty$, and we obtain a syntactical description of the classifying topos $\mathcal{E}(\mathcal{T})$ of an arbitrary theory $\mathcal{T}$ in $L^g_{\infty\omega}$. Applying this to an arbitrary site $\mathcal{C}$ with finite left limits, through the theory $\mathcal{T} = \mathcal{T}_{\mathcal{C}}$, we obtain a syntactical description of $\mathcal{E}(\mathcal{T}_{\mathcal{C}})$. In the next chapter, we will point out that $\mathcal{E}(\mathcal{T}_{\mathcal{C}})$ is nothing but $\widetilde{\mathcal{C}}$, the category of (set-valued) sheaves on $\mathcal{C}$ (up to equivalence).

Chapter 9

Classifying topoi

§1. Classifying topoi

This section summarizes and reformulates some of the work of earlier chapters around the notion of 'classifying topos'.

Let $\mathcal{T}$ be a 'theory', in a purposely unspecified sense, and let us assume that we have the notion of "$\mathcal{E}$-model of T", for an arbitrary (Grothendieck) topos $\mathcal{E}$. We write $M : T \to \mathcal{E}$ to denote that M is an $\mathcal{E}$-model of T. For simplicity, assume that we have a category, also denoted by $\mathcal{T}$, such that models $M : \mathcal{T} \to \mathcal{E}$ can be identified with *certain* functors $\mathcal{T} \to \mathcal{E}$. E.g., if $\mathcal{T}$ is a theory in $L^g_{\omega\omega}$ and $\mathcal{R}(\mathcal{T})$ is its associated logical category (c.f. Chapter 8, Section 1), then $\mathcal{T}$ can be identified with $\mathcal{R}(\mathcal{T})$ and models $M : \mathcal{T} \to \mathcal{E}$ can be identified with *logical* functors $\mathcal{R}(\mathcal{T}) \to \mathcal{E}$.

Given a topos $\mathcal{E}_0$ and an $\mathcal{E}_0$-model $M_0 : T \to \mathcal{E}_0$, we call $\mathcal{E}_0$ a (the) *classifying topos of* T *with canonical* $\mathcal{E}_0$*-model* $M_0 : T \to \mathcal{E}_0$ (or: classifying topos of T via $M_0 : T \to \mathcal{E}_0$) if for every $\mathcal{E}$-model $M : \mathcal{T} \to \mathcal{E}$, for any topos $\mathcal{E}$, there is an $\mathcal{E}$-model of $\mathcal{E}_0$, $u^* : \mathcal{E}_0 \to \mathcal{E}$ making the following commutative:

$$\begin{tikzcd} T \arrow[r, "M_0"] \arrow[dr, "M"'] & \mathcal{E}_0 \arrow[d, "u^*"] \\ & \mathcal{E} \end{tikzcd}$$

such that, moreover, u^* is unique up to an isomorphism in the functor category $\mathcal{E}^{\mathcal{E}_0}$. Note that, according to our terminology, here 'an $\mathcal{E}$-model of $\mathcal{E}_0$" means one that is a continuous functor between the sites $\mathcal{E}_0$ and $\mathcal{E}$ with the canonical topologies (c.f. Chapter 1); also, equivalently, u^* can be required to be u^* form a geometric morphism $U = (u_*, u^*, \phi) : \mathcal{E} \to \mathcal{E}_0$, c.f. 1.3.9 and 1.3.12.

We sometimes say that $M_0 : \mathcal{T} \to \mathcal{E}_0$ is a (the) *generic model* of $\mathcal{T}$, for $M_0 : \mathcal{T} \to \mathcal{E}_0$ the canonical model into a classifying topos $\mathcal{E}_0$.

Notice that a classifying topos $\mathcal{E}_0$ with canonical model $M_0 : \mathcal{T} \to \mathcal{E}_0$ is uniquely determined up to equivalence *over* $\mathcal{T}$ in the sense that in case $M_1 : \mathcal{T} \to \mathcal{E}_1$ is another generic model in a topos $\mathcal{E}_1$, then there are equivalence functors $u_1^* : \mathcal{E}_0 \to \mathcal{E}_1$, $u_0^* : \mathcal{E}_1 \to \mathcal{E}_0$ such that $u_1^* \circ u_0^* \simeq \mathrm{id}_{\mathcal{E}_0}$, $u_0^* \circ u_1^* \simeq \mathrm{id}_{\mathcal{E}_1}$ and $M_0 = u_1^* \circ M_1$, $M_1 = u_0^* \circ M_0$; here '$\simeq$' means isomorphism in the respective functor categories. This fact allows us to talk about *the* classifying topos which we denote by $\mathcal{E}(\mathcal{T})$ and *the* generic model $M_0 : \mathcal{T} \to \mathcal{E}(\mathcal{T})$.

There is a slightly stronger notion of classifying topos. Consider the category of all $\mathcal{E}$-models of $\mathcal{E}_0$; i.e., the full subcategory of the functor category $\mathcal{E}^{\mathcal{E}_0}$ whose objects are the

$\mathcal{E}$-models of $\mathcal{E}_0$. Denote this category by $\mathrm{Mod}(\mathcal{E}_0, \mathcal{E})$. (We note that $\mathrm{Mod}(\mathcal{E}_0, \mathcal{E})$ is to the category $\mathrm{Hom}_{\mathrm{top}}(\mathcal{E}, \mathcal{E}_0)^{\mathrm{op}}$, the opposite of the category of geometric morphisms $\mathcal{E} \to \mathcal{E}_0$ as the latter is introduced in Section 3.2 of SGA4, Exposé IV.) Also, $\mathrm{Mod}(\mathcal{T}, \mathcal{E})$ denotes the full subcategory of the functor category $\mathcal{E}^{\mathcal{T}}$ whose objects are the $\mathcal{E}$-models of $\mathcal{T}$. Now clearly, given $M_0 : \mathcal{T} \to \mathcal{E}_0$, M_0 induces (by composition) $\widehat{M_0} : \mathrm{Mod}(\mathcal{E}_0, \mathcal{E}) \to \mathcal{E}^{\mathcal{T}}$. The stronger definition of 'classifying topos" is that $\mathcal{E}_0$ is a classifying topos of $\mathcal{T}$ with canonical model $M_0 : \mathcal{T} \to \mathcal{E}_0$ if the induced functor $\widehat{M_0} : \mathrm{Mod}(\mathcal{E}_0, \mathcal{E}) \to \mathcal{E}^{\mathcal{T}}$ maps a model $\mathcal{E}_0 \to \mathcal{E}$ into a model $\mathcal{T} \to \mathcal{E}$ and actually establishes an equivalence, also denoted by $\widehat{M_0}$, $\widehat{M_0} : \mathrm{Mod}(\mathcal{E}_0, \mathcal{E}) \xrightarrow{\sim} \mathrm{Mod}(\mathcal{T}, \mathcal{E})$. It is easy to see that this is indeed at least as strong a condition as the previous one.

Now, Chapter 8 explicitly constructs classifying topoi for finitary or infinitary coherent theories; we think of Theorems 8.4.1 and 8.4.4 in particular. More precisely, we can put the matter as follows.

Recall (c.f. Chapter 8, Section 1) that a *coherent theory* $\mathcal{T}$ is a pair $\mathcal{T} = (F, T)$ given by a fragment F of $L^g_{\infty\omega}$ and a *set* T of axioms in F. Also, for a topos $\mathcal{E}$ an $\mathcal{E}$-model $M : \mathcal{T} \to \mathcal{E}$ is any interpretation of the language L of $\mathcal{T}$ in $\mathcal{E}$ that satisfies the axioms in T; namely, since $\mathcal{E}$ is a topos, every fragment of $L^g_{\infty\omega}$ is automatically adequately (i.e., stably) interpreted by any interpretation $M : L \to \mathcal{E}$. In other words, since we are interested in 'target categories' with are topoi, the fragment F in specifying $\mathcal{T}$ is irrelevant. Enlarge F to some (any) $L_{\mu\omega}$ with μ an infinite regular cardinal or $\mu = \infty$, (clearly, we can assume that F is a set). Then by 8.3.1(i), the μ-logical category $\mathcal{R} = \mathcal{R}_\mu(\mathcal{T})$ will, if we wish, serve as a category that can replace $\mathcal{T}$ in diagrams involving models $M : \mathcal{T} \to \mathcal{E}$ for $\mathcal{E}$ a topos; a model M now corresponds to a μ-logical functor $\mathcal{R} \to \mathcal{E}$.

Theorem 9.1.1 *Any coherent theory $\mathcal{T}$ has a classifying topos, in fact in the stronger sense indicated above.*

PROOF. Consider jus the *set* T of axioms of $\mathcal{T}$; forget about the fragment F. Now consider T axiomatizing a theory in the full logic $L^g_{\infty\omega}$. By 8.3.1(i), there is a small ∞-logical category $\mathcal{R} = \mathcal{R}_\infty(\mathcal{T})$ that can replace $\mathcal{T}$ whenever $\mathcal{T}$ is models in a topos. Now, apply 8.4.1 (or 8.4.4) for $\kappa = \infty$. We obtain an ∞-pretopos $\mathcal{E}_0 = \mathcal{P}_\infty$ and a model $M_0 : \mathcal{T} \to \mathcal{E}_0$ ($M_0 : \mathcal{T} \xrightarrow{\text{can.}} \mathcal{R} \xrightarrow{I_0} \mathcal{E}_0$) with exactly the universal property defining "$\mathcal{E}_0$ is a classifying topos of $\mathcal{T}$ with canonical model M_0", with 'topoi' replaced by ∞-pretopoi. It is left to show that $\mathcal{E}_0$ has a set of generators (to make it into a topos, c.f. 3.4.8) and that $M_0 : \mathcal{T} \to \mathcal{E}$ "works in the stronger sense" as well.

The first claim follows from the fact that $\mathcal{R}$ is small as well as the representation of the objects of $\mathcal{E}_0$ as quotients of disjoint sums of objects coming from $\mathcal{R}$, c.f. 8.4.4(iii). Since the family of canonical injections for any disjoint sum is obviously an effective epimorphic family, and the canonical morphism form an object to a quotient of that object by an equivalence relation is an effective epimorphism, it clearly follows that the set of objects $I_0(S)$ for $S \in \mathrm{Ob}(\mathcal{R})$ form a set of generators for $\mathcal{E}$.

The second claim will not be verified in detail; it would easily follow using 8.3.1(ii) as wee as consequences of it in the situations of 8.4.2 and 8.4.3. □

Note that our construction is independent of the general topos theory of Chapter 1 and is in some sense syntactical, i.e., a "*presentation relative*" to the logical operations of a topos.

Here is a way to construct the classifying topos as the category of sheaves over a site. A "better" way will be described in Section 4 (Coste's construction).

Given a theory $\mathcal{T} = (F, T)$, we first of all assume that F is $L^g_{\mu\infty}$ for an infinite regular cardinal μ and we pass to $\mathcal{R} = \mathcal{R}_\mu(\mathcal{T})$ (c.f. Chapter 8, Section 3; actually, we could define $\mathcal{R}$ more sparingly on the basis of *any* fragment F containing T). Now, we regard the μ-logical category $\mathcal{R}$ as a site with the *μ-precanonical topology*, i.e., the topology generated by those coverings in the canonical topology (c.f. 1.1.9) which have $< \mu$ members. E.g., if $\mu = \aleph_0$, the $\aleph_0$-precanonical, or simply precanonical, topology is one that is generated by finite coverings which are stable effective epimorphic families. Denote by $\mathcal{E}(\mathcal{R})$ of $\mathcal{E}(\mathcal{T})$ the topos which is the category of sheaves $\widetilde{\mathcal{R}}$ over the site $\mathcal{R}$ with the μ-precanonical topology. Let $\varepsilon : \mathcal{R} \to \mathcal{E}(\mathcal{R})$ be the 'representable sheaf' functor (c.f. Chapter 1, Section 3) and M_0 the composite $\mathcal{T} \xrightarrow{\text{can.}} \mathcal{R} \xrightarrow{\varepsilon} \mathcal{E}(\mathcal{R}) = \mathcal{E}(\mathcal{T})$.

Proposition 9.1.2 *$M_0 : \mathcal{T} \to \mathcal{E}(\mathcal{T})$ is a generic model of $\mathcal{T}$.*

This proposition becomes obvious once one realizes

Proposition 9.1.3 *The μ-logical functors from any μ-logical category $\mathcal{R}$ are exactly the functors from $\mathcal{R}$ which are continuous with respect to the μ-precanonical topology.*

This last proposition is the '$< \infty$' version of 3.4.10 and it is proved in the same way.

Now, 9.1.2 immediately follows since the requisite universal property for $\mathcal{E}(\mathcal{T})$ will be identical to the universal property of $\widetilde{\mathcal{R}}$ (c.f. 1.3.15) with respect to continuous functors from the site $\mathcal{R}$ with the μ-precanonical topology.

Next we turn to a consideration which is a converse of the just preceding one, namely, we will regard $\widetilde{\mathcal{C}}$ for an arbitrary site $\mathcal{C}$, as a classifying topos for a theory. From this view point, in the next section, we will reconsider Grothendieck's theory of coherent objects of a coherent topos, c.f. Exposé VI of SGA 4.

First, recall the theory $\mathcal{T}_\mathcal{C}$ associated with a site $\mathcal{C}$ form Chapter 6, Section 1. $\mathcal{T}_\mathcal{C}$ is formulated in the language $L_\mathcal{C}$ canonically associated with the underlying category of $\mathcal{C}$. The set of axioms $T_\mathcal{C}$ of $\mathcal{T}_\mathcal{C}$ contains the following

(i) all axioms of category, items 1 and 2 before 2.4.5, for true identity morphisms and true commutative triangles of $\mathcal{C}$;

(ii) all the axioms related to finite left limit diagrams in $\mathcal{C}$ (c.f. 2.4.5);

(iii) all axioms of the form

$$a \approx a \Rightarrow \bigvee_{i \in I} \exists a_i (f(a_i) \approx a)$$

for any (basic) covering family $(A_i \xrightarrow{f_i} A)_{i \in I}$ of $\mathcal{C}$. We have (c.f. 6.1.1)

Proposition 9.1.4 *$M : \mathcal{C} \to \mathcal{E}$ is continuous (an $\mathcal{E}$-model of $\mathcal{C}$) iff M is an $\mathcal{E}$-model of the theory $T_\mathcal{C}$, for any topos $\mathcal{E}$.*

Observe that any functor from $\mathcal{C}$ *is*, at the same time, an interpretation of the language $L_\mathcal{C}$. Conversely, any interpretation of $L_\mathcal{C}$ satisfying (i) above *is* a functor from $\mathcal{C}$.

Proposition 9.1.5 *$\widetilde{\mathcal{C}}$ is equivalent to the classifying topos $\mathcal{E}(\mathcal{T}_\mathcal{C})$ of $\mathcal{T}_\mathcal{C}$* over *$\mathcal{C}$; i.e., with $\mathcal{C} \xrightarrow{\varepsilon} \widetilde{\mathcal{C}}$ and the generic model $\mathcal{T}_\mathcal{C} \xrightarrow{M_0} \mathcal{E}(\mathcal{T}_\mathcal{C})$, there is an equivalence $\widetilde{\mathcal{C}} \underset{e_2}{\overset{e_1}{\rightleftarrows}} \mathcal{E}(\mathcal{T}_\mathcal{C})$ that carries ε into M_0: $\varepsilon = e_2 \circ M_0$, $M_0 = e_1 \circ \varepsilon$ (here we consider M_0 a functor $\mathcal{C} \to \mathcal{E}(\mathcal{T}_\mathcal{C})$).*

The proof is immediate on the basis of 9.1.4, the definition of the classifying topos and the universal property 1.3.14 of $\widetilde{\mathcal{C}}$.

Notice that, via the description provided by 8.4.4, this gives an explicit presentation (with respect to logical operations) of the category of sheaves $\widetilde{\mathcal{C}}$, base on a presentation, namely $\mathcal{T}_\mathcal{C}$, of the site $\mathcal{C}$ itself.

The main conclusion of this section (c.f. especially 9.1.2 and 9.1.5) is that the construction of the classifying topos $\mathcal{E}(\mathcal{T})$ of a theory has precisely the same scope as the Grothendieck construction of the category of sheaves $\widetilde{\mathcal{C}}$. Each $\mathcal{E}(\mathcal{T})$ can be thought of as a $\widetilde{\mathcal{C}}$, and each $\widetilde{\mathcal{C}}$ as an $\mathcal{E}(\mathcal{T})$; both, in fact, in a natural way.

§2 Coherent objects

Next we turn to *coherent objects* of a topos (c.f. Exposé VI, SGA 4).

Definition 9.2.1 (c.f. loc. cit.) *Let $\mathcal{E}$ be a topos.*

(i) *An object X of $\mathcal{E}$ is called* quasi-compact *(q.c.) if every covering $(X_i \to X)_{i\in I}$ (in the canonical topology on $\mathcal{E}$) has a finite subcovering $(X_i \to X)_{i\in I'}$, $I' \subset I$ finite.*

(ii) *An object $X \in \mathrm{Ob}(\mathcal{E})$ is* coherent *if it is q.c. and if the following is true: whenever Q_1, Q_2 are q.c. objects in $\mathcal{E}$, $Q_1 \to X$, $Q_2 \to X$ are arbitrary morphisms in $\mathcal{E}$, then $Q_1 \times_X Q_2$ in the pullback diagram*

$$\begin{tikzcd} Q_1 \arrow[r] & X \\ Q_1 \times Q_2 \arrow[u] \arrow[r] & Q_2 \arrow[u] \end{tikzcd}$$

is q.c.

(iii) $\mathrm{Coh}(\mathcal{E})$*, the category of coherent objects of $\mathcal{E}$, is the full subcategory of $\mathcal{E}$ whose objects are the coherent objects of $\mathcal{E}$.*

Recall (c.f. Chapter 6) that an *algebraic site* is one (which has finite left limits and) whose topology is generated by finite covering families and a *coherent topos* is one that is equivalent to the category of sheaves $\widetilde{\mathcal{C}}$ over an algebraic site $\mathcal{C}$.

Theorem 9.2.2 (Grothendieck; loc. cit. Exercise 3.11, p. 232)

(i) *For every coherent topos $\mathcal{E}$, $\mathrm{Coh}(\mathcal{E})$ is a pretopos, it is equivalent to a small category, and the inclusion functor $\mathrm{Coh}(\mathcal{E}) \hookrightarrow \mathcal{E}$ is (finitely) logical, i.e. it preserves finite left limits, finite sups and images.*

(ii) *For $\mathcal{P}$ a small pretopos, for $\widetilde{\mathcal{P}}$ the category of sheaves over $\mathcal{P}$ as a site with the precanonical topology, $\widetilde{\mathcal{P}}$ is a coherent tops and $\mathcal{P} \xrightarrow{\varepsilon} \widetilde{\mathcal{P}}$ factors through the inclusion $\mathrm{Coh}(\widetilde{\mathcal{P}}) \to \widetilde{\mathcal{P}}$ in $\varepsilon' : \mathcal{P} \to \mathrm{Coh}(\widetilde{\mathcal{P}})$*

$$\begin{tikzcd} \mathcal{P} \arrow[r, "\varepsilon'"] \arrow[rr, bend right, "\varepsilon"'] & \mathrm{Coh}(\widetilde{\mathcal{P}}) \arrow[r] & \widetilde{\mathcal{P}} \end{tikzcd}$$

such that ε' is an equivalence

$$\varepsilon' : \mathcal{P} \xrightarrow{\sim} \mathrm{Coh}(\mathcal{P}).$$

(iii) *With $\mathcal{P} = \mathrm{Coh}(\mathcal{E})$, $\mathcal{E}$ a coherent topos, the inclusion $\mathrm{Coh}(\mathcal{E}) \to \mathcal{E}$ satisfies the universal property of $\mathcal{P} \xrightarrow{\varepsilon_{\mathcal{P}}} \widetilde{\mathcal{P}}$, where $\mathcal{P}$ is regarded as a site with the precanonical topology.*

Before proceeding to the proof, we mention two immediate corollaries.

Corollary 9.2.3 *A small category is a pretopos if and only if it is equivalent to* $\mathrm{Coh}(\mathcal{E})$ *for a coherent topos* $\mathcal{E}$*.*

Corollary 9.2.3 *For pretopoi* $\mathcal{P}_1$*,* $\mathcal{P}_2$*, if* $\widetilde{\mathcal{P}}_1$ *and* $\widetilde{\mathcal{P}}_2$ *are equivalent categories (where the sites are meant with the precanonical topologies) then* $\mathcal{P}_1$ *and* $\mathcal{P}_2$ *are equivalent as well.*

Recall the theory $\mathcal{T}_{\mathcal{C}}$ associated to the site $\mathcal{C}$ and the pretopos $\mathcal{P}(\mathcal{T}_{\mathcal{C}}) = \mathcal{P}_{\aleph_0}(\mathcal{T}_{\mathcal{C}})$ associated to the theory $\mathcal{T}_{\mathcal{C}}$, c.f. 8.4.1′ for $\kappa = \aleph_0$. During our proof of 9.2.2 we will also establish

Theorem 9.2.5 *For an algebraic site* $\mathcal{C}$*, we have canonical equivalences*

$$\mathrm{Coh}(\mathcal{E}(\mathcal{T}_{\mathcal{C}})) \simeq \mathrm{Coh}(\widetilde{\mathcal{C}}) \simeq \mathcal{P}(\mathcal{T}_{\mathcal{C}}).$$

Via 8.4.4*, this gives an explicit presentation of the category of coherent objects of a coherent topos, in terms of a presentation of an algebraic site giving rise to the topos.*

For the proof of 9.2.2, we need the following lemma coming form SGA 4.

Lemma 9.2.6 *Let* $\mathcal{C}$ *be an algebraic site,* $\mathcal{E} = \widetilde{\mathcal{C}}$ *the category of sheaves over* $\mathcal{C}$*,* $\varepsilon : \mathcal{C} \to \mathcal{E}$ *the canonical functor*

(i) *Every covering in* $\mathcal{C}$ *contains a finite subcovering (i.e., every object in* $\mathcal{C}$ *is "q.c." in the site* $\mathcal{C}$*).*

(ii) *For every* $A \in \mathrm{Ob}(\mathcal{C})$*,* $\varepsilon(A)$ *is q.c. in* $\mathcal{E}$*.*

(iii) *Suppose* $(X_i \to X)_{i\in I}$ *is a* finite *covering in* $\mathcal{E}$ *of* X *by q.c. objects* X_i *(*I *finite). Then* X *is q.c.*

(iv) *Every object of the form* $\varepsilon(A)$*, for* $A \in \mathrm{Ob}(\mathcal{C})$*, is coherent in* $\mathcal{E}$*.*

PROOFS. (AD (i)) The coverings in $\mathcal{C}$ form the smallest system containing some given finite families and closed under the closure condition 1.1.1(i)-(iv). "By induction" corresponding to these closure conditions, it is straightforward to show that every covering in $\mathcal{C}$ has the required property.

(AD (ii)) We will use, among others, the fact (1.3.7) that the objects εA ($A \in \mathrm{Ob}(\mathcal{C})$) form a set of generators for $\mathcal{E}$, as well as the technical Lemma 1.3.8(i). Let $(X_i \to \varepsilon A)_{i\in I}$ be a covering in $\mathcal{E}$. By 1.3.7, for each $i \in I$ there are coverings of the form $(\varepsilon B_{ij} \to X_i)_{j\in J_i}$. Furthermore, by 1.3.8(i) we can assume (by further refining these coverings) that here each of the composites $\varepsilon B_{ij} \to X_i \to \varepsilon A$ is of the form εg_{ij} for a morphism $g_{ij} : B_{ij} \to A$ in $\mathcal{C}$. By composing coverings, we get that $(\varepsilon B_{ij} \to \varepsilon A)_{i\in I, j\in J_i}$ is a covering in $\mathcal{E}$. By 1.3.3(ii), $(B_{ij} \xrightarrow{g_{ij}} A)_{i,j}$ is a covering in $\mathcal{C}$. By part of (i) of the present lemma, there is a finite subset $I' \subset I$ (and finite subsets $J'_i \subset J_i$) such that $(B_{ij} \to A)_{i\in I', j\in J'_i}$ is a covering in $\mathcal{C}$. Then so is $(\varepsilon B_{ij} \xrightarrow{\varepsilon g_{ij}} \varepsilon A)_{i\in I', j\in J'_i}$ in $\mathcal{E}$, by 1.3.3(ii) again. A fortiori, $(X_i \to \varepsilon A)_{i\in I'}$ is a covering, proving assertion (ii).

(AD (iii)) Let $(Y_j \to X)_{j\in J}$ be a covering. We have the coverings $(Y_j \times_X X_i \to X_i)_{j\in J}$ 'by pullback', for each $i \in I$. Each of the latter has a finite subcovering, say with index sets J'_i. Put $J' = \bigcup_{i\in I} J'_i$; J' is finite. By composition, $(Y_j \times_X X_i \to X)_{i\in I, j\in J'_i}$ is a covering. A fortiori, $(Y_j \to X)_{j\in J'}$ is a covering, proving (iii).

(AD (iv)) Let Q_1, Q_2 be q.c. objects of $\mathcal{E}$, and consider a pullback diagram

$$\begin{tikzcd} Q_1 \arrow[r] & \varepsilon A \\ Y = Q_1 \times Q_2 \arrow[u] \arrow[r] & Q_2 \arrow[u] \end{tikzcd}$$

By 1.3.7, Q_1 and Q_2 can be covered by morphisms with domains which are objects of the form εB; since they are q.c., finitely many such morphisms suffice. Let $(\varepsilon B_i \to Q_1)_{i\in I}$, $(\varepsilon C_j \to Q_2)_{j\in J}$ be finite coverings. By 1.3.8(i), we can further assume that each of the composites $\varepsilon B_i \to Q_1 \to \varepsilon A$, $\varepsilon C_j \to Q_2 \to \varepsilon A$ is of the form εg_i, εh_j; respectively, for some $g_i : B_i \to A$, $h_j : C_j \to A$ in $\mathcal{C}$. By applying pullback to these coverings twice (as well as using 'composition'), we clear have that

$$(\varepsilon B_i \times_{\varepsilon A} \varepsilon C_j \to Y)_{i\in I, j\in J}$$

is a covering family in $\mathcal{E}$.

Now, since ε is left exact and the morphisms $\varepsilon B_i \to \varepsilon A$, $\varepsilon C_j \to \varepsilon A$ 'come from' the morphisms $g_i : B_i \to A$, $h_j : C_j \to A$, respectively, each of the morphisms $Y_{ij} = \varepsilon B_i \times_{\varepsilon A} \varepsilon C_j \to \varepsilon A$ in the pullback diagram

$$\begin{tikzcd} \varepsilon B_i \arrow[r, "\varepsilon g_i"] & \varepsilon A \\ Y_{ij} = \varepsilon B_i \times_{\varepsilon A} \varepsilon C_j \arrow[u] \arrow[r] & \varepsilon C_j \arrow[u, "\varepsilon h_j"'] \end{tikzcd}$$

is (can be taken to be) the ε-image of the pullback diagram

$$\begin{tikzcd} B_i \arrow[r, "g_i"] & A \\ B_i \times_A C_j \arrow[u] \arrow[r] & C_j \arrow[u, "h_j"'] \end{tikzcd}$$

in $\mathcal{C}$. But then $\varepsilon(B_i \times_A C_j) = Y_{ij}$ is quasi compact by (ii) of the present lemma. Since $(Y_{ij} \to Y)_{i,j}$ is a finite covering, by (iii) it follows that $Y = Q_1 \times_{\varepsilon A} Q_2$ is q.c., showing the second condition for coherence of εA. Since εA is q.c. by (ii), εA is coherent as required. □

Now we begin the proof of 9.2.2 and 9.2.5. First we prove 9.2.2(ii). Let $\mathcal{P}$ be a pretopos, consider $\mathcal{P}$ a site with the precanonical topology an let $\widetilde{\mathcal{P}}$ be the category of sheaves over this site. Consider the representable sheaf functor $\varepsilon : \mathcal{P} \to \widetilde{\mathcal{P}}$. Note that since every covering in the site $\mathcal{P}$ belongs to the canonical topology on $\mathcal{P}$ and hence every presheaf over $\mathcal{P}$ is a sheaf, ε (being the composition of the Yoneda embedding $\mathcal{C} \to \widehat{\mathcal{C}}$ and the associated sheaf functor $a : \widehat{\mathcal{C}} \to \widetilde{\mathcal{C}}$) is full and faithful. ε is conservative as well: the proof of 1.4.8 applies! Also, $\mathcal{P}$ has quotients of its equivalence relations, since $\mathcal{P}$ is a pretopos. These facts mean that $\varepsilon : \mathcal{P} \to \widetilde{\mathcal{P}}$ satisfies conditions (i) to (iii) of 1.4.11. Using 1.4.11, we show that the essential image of ε in $\widetilde{\mathcal{P}}$ is exactly the collection of coherent objects of $\widetilde{\mathcal{P}}$, i.e., that (i) εA is coherent for every $A \in \mathrm{Ob}(\mathcal{P})$ and (ii) whenever X is coherent in $\widetilde{\mathcal{P}}$, there is $A \in \mathrm{Ob}(\mathcal{P})$ such that $X \simeq \varepsilon(A)$. Assertion (i) follows from 9.2.6(iv) since the site $\mathcal{P}$ is obviously algebraic. Conversely, let X be a coherent object in $\widetilde{\mathcal{P}}$; we verify the condition 1.4.11 for X as S there. by 1.3.7 and the fact that X is quasi compact, there are finitely many objects A_i ($i \in I$) in $\mathcal{P}$ and morphisms f_i such that $(\varepsilon A_i \xrightarrow{f_i} X)_{i\in I}$ is an (effective) epimorphic family (in $\widetilde{\mathcal{P}}$). Then, with p induced by the universal property of the disjoint sum (coproduct) $\coprod_{i\in I} \varepsilon A_i$, we have an (effective) epimorphism $p : \coprod_{i\in I} \varepsilon A_i \to X$ (c.f. also 1.4.7). Being a pretopos, $\mathcal{P}$ has finite disjoint sums. Let $A = \coprod_{i\in I} A_i$ be the disjoint sum of the A_i in $\mathcal{P}$. ε preserves finite disjoint sums by 3.4.13 since $\varepsilon : \mathcal{P} \to \widetilde{\mathcal{P}}$ is logical by 9.1.3. Hence $\coprod_{i\in I} \varepsilon A_i$ is (can be take to be) εA and we have the effective epimorphism $p:\varepsilon A \to X$. εA is quasi compact in $\widetilde{\mathcal{P}}$ by 9.2.6(i). Since X is assumed to be coherent, the fibered product $Y = \varepsilon A \times_X \varepsilon A$ (using this morphism p) is quasi compact. Hence, as before, there is an effective epimorphism of the form $\varepsilon(B) \to Y$ with $B \in \mathrm{Ob}(\mathcal{P})$. This shows precisely the condition 1.4.11 concerning the object X, thus assertion (ii) follows by 1.4.11.

We have shown that the full and faithful functor $\varepsilon : \mathcal{P} \to \widetilde{\mathcal{P}}$ has essential image the coherent objects in $\widetilde{\mathcal{P}}$. It follows that ε establishes an equivalence $\mathcal{P} \xrightarrow{\sim} \mathrm{Coh}(\widetilde{\mathcal{P}})$ in the precise sense stated in 9.2.2(ii).

Turning to part (i) of 9.2.2, let $\mathcal{C}$ be an algebraic site. We use the associated theory $\mathcal{T}_{\mathcal{C}}$ and the pretopos $\mathcal{P} = \mathcal{P}(\mathcal{T}_{\mathcal{C}})$ together with the canonical model $M_0 : \mathcal{T}_{\mathcal{C}} \to \mathcal{P}$ given by 8.4.1$'$ as also in the statement of 9.1.5.

We now show the following

Claim 9.2.7 *The categories $\widetilde{\mathcal{C}}$ and $\widetilde{\mathcal{P}}$ are equivalent, $\widetilde{\mathcal{C}} \simeq \widetilde{\mathcal{P}}$.*

By the fundamental property of $\mathcal{T}_{\mathcal{C}}$ formulated in 9.1.4 as well as the construction 8.4.1$'$ of $\mathcal{P}(\mathcal{T}_{\mathcal{C}})$ $M_0 : \mathcal{C} \to \mathcal{P}$ is a continuous functor and we have the following universal property of $\mathcal{P}$: for any topos $\mathcal{E}$ and continuous functor $F : \mathcal{C} \to \mathcal{E}$ we have the logical functor $\mathcal{P} \dashrightarrow \mathcal{E}$, unique up to isomorphism in $\mathcal{E}^{\mathcal{P}}$, making

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "M_0"] \arrow[d] & \mathcal{P} \arrow[dl, dashed] \\
\mathcal{E} &
\end{tikzcd}$$

commutative.

We build the diagram

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "M_0"] \arrow[d, "\varepsilon_1"'] & \mathcal{P} \arrow[dl, "\eta"'] \arrow[d, "\varepsilon_2"] \\
\widetilde{\mathcal{C}} \arrow[r, shift left, "\mu_1"] & \widetilde{\mathcal{P}} \arrow[l, shift left, "\mu_2"]
\end{tikzcd}$$

as follows. $\varepsilon_1 = \varepsilon_{\mathcal{C}} : \mathcal{C} \to \widetilde{\mathcal{C}}$ is the representable sheaf functor. $\widetilde{\mathcal{P}}$ is the category of sheaves over $\mathcal{P}$ as a site with the pre-canonical topology, $\varepsilon_2 = \varepsilon_{\mathcal{P}} : \mathcal{P} \to \widetilde{\mathcal{P}}$ is the representable sheaf functor for this site. Next, we use the universal property of $\mathcal{P}$ described above, with $\mathcal{E} = \widetilde{\mathcal{C}}$, to obtain the logical $\eta : \mathcal{P} \to \widetilde{\mathcal{C}}$, such that the triangle commutes. By 9.1.3, η is continuous as a functor between sites, with the respective topologies considered here ($\mathcal{P}$: precanonical, $\widetilde{\mathcal{C}}$: canonical).

Having the continuous $\eta : \mathcal{P} \to \widetilde{\mathcal{C}}$, by the universal property of $\widetilde{\mathcal{P}}$ over $\mathcal{P}$ (1.3.15) we obtain the continuous $\mu_2 \widetilde{\mathcal{P}} \to \widetilde{\mathcal{C}}$ such that the lower triangle having μ_2 as a side commutes. Finally, the universal property of $\widetilde{\mathcal{C}}$ with respect to $\mathcal{C}$ implies the existence of μ_1 such that the rectangle with lower side μ_1 commutes.

We claim that $\mu_2 \circ \mu_1 \simeq \mathrm{id}_{\widetilde{\mathcal{C}}}$, $\mu_1 \circ \mu_2 \simeq \mathrm{id}_{\widetilde{\mathcal{P}}}$ (the isomorphisms taking place in the respective functor categories), which will establish Claim 9.2.7.

Denoting $\mu_2 \circ \mu_1$ by f, we have $f \circ \varepsilon_1 = \varepsilon_1$ but the commutativity properties of the above diagram. From the uniqueness part of the universal property of $\widetilde{\mathcal{C}}$ over $\mathcal{C}$ ("for $\mathcal{E} = \widetilde{\mathcal{C}}$"), it follows that $f \simeq \mathrm{id}_{\widetilde{\mathcal{C}}}$ as required for the first isomorphism.

Let $g \underset{\mathrm{df}}{=} \mu_1 \circ \mu_2$. By the construction of our diagram, $g \circ \varepsilon_2 \circ M_0 = \varepsilon_2 \circ M_0$. By the uniqueness part of the universal property of $\mathcal{P}$ over $\mathcal{C}$ (c.f. 8.4.1$'$; $\mathcal{P} = \mathcal{P}(\mathcal{T}_{\mathcal{C}})$; note that models $\mathcal{T}_{\mathcal{C}} \cdots$ are models $\mathcal{C} \cdots$), $g \circ \varepsilon_2 \simeq \varepsilon_2$ in the functor category $(\widetilde{\mathcal{P}})^{\mathcal{P}}$. To conclude that $g \simeq \mathrm{id}_{\widetilde{\mathcal{P}}}$, we will use something slightly stronger that the statement of the universal property in 1.3.15 of $\widetilde{\mathcal{P}}$ over $\mathcal{P}$. Namely, we use that the functor

$$\widehat{\varepsilon_2} : \mathrm{Mod}(\widetilde{\mathcal{P}}, \mathcal{E}) \to \mathrm{Mod}(\mathcal{P}, \mathcal{E}) \tag{1}$$

(defined by composition with ε_2) is an equivalence of categories, for any topos $\mathcal{E}$. This could have been verified in Chapter 1 at the appropriate place but it also follows from the identification of $\widetilde{\mathcal{P}}$ as the classifying topos of (the internal theory of) $\mathcal{P}$ (c.f. 9.1.2) and the fact that classifying topoi do have the stronger universal property, c.f. 9.1.1.

Accepting that (1) is an equivalence, we can say the following. Apply (1) for $\mathcal{E} = \widetilde{\mathcal{P}}$. By the functor $\widehat{\varepsilon_2}$, $\mathrm{id}_{\widetilde{\mathcal{P}}}$ is mapped onto ε_2, g is mapped onto $g \circ \varepsilon_2$. Since $\widehat{\varepsilon_2}$ is full and faithful, $g \circ \varepsilon_2 \simeq \varepsilon_2$ implies that $g \simeq \mathrm{id}_{\widetilde{\mathcal{P}}}$ as required.

This completes showing Claim 9.2.7.

Finally, to see 9.2.2(i) it is sufficient to invoke 9.2.7 and 9.2.2(ii) for $\mathcal{P} = \mathcal{P}(\mathcal{T}_{\mathcal{C}})$. In particular, since $\widetilde{\mathcal{C}}$ and $\widetilde{\mathcal{P}}$ are equivalent, any "purely categorical" property of $\widetilde{\mathcal{P}}$ is transferred to $\widetilde{\mathcal{C}}$. The definitions of coherence are "purely categorical". Since $\mathrm{Coh}(\widetilde{\mathcal{P}}) \simeq \mathcal{P}$ (c.f. 9.1.2) is a pretopos, $\mathrm{Coh}(\widetilde{\mathcal{C}})$ is a pretopos. Moreover, the inclusion $\mathrm{Coh}(\widetilde{\mathcal{P}}) \hookrightarrow \widetilde{\mathcal{P}}$ is logical since $\mathcal{P} \xrightarrow{\varepsilon_2} \widetilde{\mathcal{P}}$ is, and since we have the precise statement of 9.2.2(ii). Hence, by the above "transference principle", the inclusion $\mathrm{Coh}(\widetilde{\mathcal{C}}) \to \widetilde{\mathcal{C}}$ is logical. This completes the proof of 9.2.2(i); notice that we have shown 9.2.2(iii) and 9.2.5 as well. □

We note that our proof of Grothendieck' theorem 9.2.2 is essentially different form the proof which is suggested in Exercise 3.11, p. 232, loc. cit. The difference lies in our use of a pretopos $\mathcal{P}$ constructed before we knew that $\mathrm{Coh}(\mathcal{E})$ was a pretopos itself.

Using the language of 'coherent objects', we now reformulate one of our main results, 7.1.8. First, a definition from SGA 4.

Definition 9.2.8 *Given topoi $\mathcal{E}_1$, $\mathcal{E}_2$ and a continuous $u^* : \mathcal{E}_1 \to \mathcal{E}_2$ (or, a geometric morphism $U = (u_*, u^*, \phi) : \mathcal{E}_2 \to \mathcal{E}_1$), we say that u^* is* coherent *(or, the geometric morphism U is* coherent*), if u^* maps coherent objects of $\mathcal{E}_1$ into coherent objects of $\mathcal{E}_2$. Equivalently, u^* is coherent if we have a commutative diagram as follows:*

$$\begin{tikzcd}
\mathcal{E}_1 \arrow[r, "u^*"] & \mathcal{E}_2 \\
\mathrm{Coh}(\mathcal{E}_1) \arrow[r, "I"'] \arrow[u, hook, "\text{incl.}", "i_1"'] & \mathrm{Coh}(\mathcal{E}_2) \arrow[u, hook, "i_2", "\text{incl.}"']
\end{tikzcd}$$

Theorem 9.2.9 *For a coherent continuous functor $u^* : \mathcal{E}_1 \to \mathcal{E}_2$ between coherent topoi $\mathcal{E}_1$, $\mathcal{E}_2$, if $\widehat{u^*} : \mathrm{Mod}(\mathcal{E}_2, \mathbf{Set}) \to \mathrm{Mod}(\mathcal{E}_1, \mathbf{Set})$ (defined by composition) is an equivalence, then u^* is an equivalence as well.*

Proof. We will refer to the commutative diagram in 9.2.8. By 9.2.2(i), $\mathcal{P}_1 = \mathrm{Coh}(\mathcal{E}_1)$, $\mathcal{P}_2 = \mathrm{Coh}(\mathcal{E}_2)$ are pretopoi. Since i_1 and i_2 are logical (c.f. 9.2.2(i)), u^* is ∞-logical and u_2 is full and faithful, it easily follows that I is logical. Consider $\widehat{I} : \mathrm{Mod}(\mathcal{P}_2) \to \mathrm{Mod}(\mathcal{P}_1)$ (with $\mathrm{Mod}(\mathcal{P}) = \mathrm{Mod}(\mathcal{P}, \mathbf{Set})$ of course) defined by composition (and denoted I^* in 7.1.8). Using 9.2.2(iii), we have $\mathrm{Mod}(\mathcal{E}_i, \mathbf{Set}) \simeq \mathrm{Mod}(\mathcal{P}_i)$ $(i = 1, 2)$ and in fact, the fact that the *diagram* $\mathrm{Mod}(\mathcal{E}_2, \mathbf{Set}) \xrightarrow{\widehat{u^*}} \mathrm{Mod}(\mathcal{E}_1, \mathbf{Set})$ is equivalent to

$\mathrm{Mod}(\mathcal{P}_2) \xrightarrow{\widehat{I}} \mathrm{Mod}(\mathcal{P}_1)$. It follows that $\widehat{I}$ is an equivalence of categories. By 7.1.8, I is an equivalence. Hence, by 9.2.2(iii) again, u^* is an equivalence □

9.2.9 is 'equivalent' to Theorem 7.1.8 because the argument above is essentially reversible. On the other hand, 9.2.9 by its general form resembles Deligne's theorem on coherent topoi (6.2.2) (tit is a "Deligne theorem on a pair of coherent topoi") and it ought to be a reasonable basic an useful theorem on coherent topoi just as Deligne's theorem is. It remains to be seen if it really is. We will illustrate the effect of the theorem on a familiar special case in the next section.

Also observe that 9.2.9 is concerned with *categories* of models whereas in Deligne's theorem there is no reference to the category of models.

We can generalize the notion of coherent object to that of κ-coherent object, for any infinite regular cardinal κ. For this purpose, we talk about κ-quasi compact objects by requiring a sub covering of power $< \kappa$ instead of a finite subcovering in 9.2.1.

Then everything in this section except 9.2.9 automatically generalizes. A 'coherent topos' is replaced by a "κ-coherent topos' which can be defined as the category of sheaves $\widetilde{\mathcal{C}}$ over a κ-algebraic site $\mathcal{C}$, the latter one having a topology generated by covering families of cardinality $< \kappa$. 'Pretopos' should be replaced by 'κ-pretopos', the precanonical topology by the κ-precanonical one.

Partly in view of the example in the next section, we give two more reformulations of Theorem 9.2.9. Observe that the second reformulation does not mention pretopoi or coherent objects.

Theorem 9.2.10 ("Points are enough for classifying".) *Let $\mathcal{T}$ be a finitary coherent theory with language L.*

(i) *Let M be an interpretation of the language L in* $\mathrm{Coh}(\mathcal{E})$*, with a coherent topos $\mathcal{E}$. Suppose M induces, by composition, an equivalence*

$$\mathrm{Mod}(\mathcal{E}, \mathbf{Set}) \xrightarrow{\sim} \mathrm{Mod}(\mathcal{T}, \mathbf{Set})$$

(in particular, whenever $N : \mathcal{E} \to \mathbf{Set}$ is a model of $\mathcal{E}$, the composition $L \overset{M}{\to} \mathrm{Coh}(\mathcal{E}) \overset{\text{incl.}}{\to} \mathcal{E} \xrightarrow{N} \mathbf{Set}$ is a model of $\mathcal{T}$). Then the composite $L \xrightarrow{M} \mathrm{Coh}(\mathcal{E}) \overset{\text{incl.}}{\longrightarrow} \mathcal{E}$ is a generic model of $\mathcal{T}$ in $\mathcal{E}$.

(ii) *Let $\mathcal{C}$ be an algebraic site and M and interpretation of the language L in the (underlying) category $\mathcal{C}$. Suppose that M induces, by composition, an equivalence*

$$\widehat{M} : \mathrm{Mod}(\mathcal{C}, \mathbf{Set}) \xrightarrow{\sim} \mathrm{Mod}(\mathcal{T}, \mathbf{Set})$$

(in particular, whenever $\mathcal{C} \xrightarrow{N} \mathbf{Set}$ is a model of $\mathcal{C}$, $L \xrightarrow{M} \mathcal{C} \xrightarrow{N} \mathbf{Set}$ is a model of $\mathcal{T}$). Then the composite $L \overset{M}{\to} \mathcal{C} \xrightarrow{\varepsilon} \widetilde{\mathcal{C}}$ is a generic model of $\mathcal{T}$ (or, $\widetilde{\mathcal{C}}$ is the classifying topos of $\mathcal{T}$, with canonical model $L \xrightarrow{M} \mathcal{C} \xrightarrow{\varepsilon} \widetilde{\mathcal{C}}$).

Proof. We treat (ii) only; (i) is similar. We will build the following diagram:

$$\begin{tikzcd}
 & & L \arrow[r, "M"] \arrow[lld] \arrow[d] & \mathcal{C} \arrow[rrd] \arrow[d, "\varepsilon=\varepsilon_{\mathcal{C}}"'] & & \\
\mathcal{P}(\mathcal{T}) \arrow[r, shift left] \arrow[rrd, bend right=15] \arrow[rrrrr, bend right=40] & \mathrm{Coh}(\mathcal{E}(\mathcal{T})) \arrow[l, shift left, "\sim"] \arrow[rd, "\text{incl.}"'] & & & \mathrm{Coh}(\widetilde{\mathcal{C}}) \arrow[r, shift left] \arrow[ld, "\text{incl.}"] & \mathcal{P}(\mathcal{T}_{\mathcal{C}}) \arrow[l, shift left, "\sim"] \arrow[lld, bend left=15] \\
 & & \mathcal{E}(\mathcal{T}) \arrow[r, "u^*"'] & \widetilde{\mathcal{C}} & &
\end{tikzcd}$$

Here $\mathcal{T}_\mathcal{C}$ is the theory associated to the site $\mathcal{C}$, $\mathcal{P}(\mathcal{T}_\mathcal{C})$ is the pretopos completion of $\mathcal{T}_\mathcal{C}$, $\mathcal{C} \to \mathcal{P}(\mathcal{T}_\mathcal{C})$ is the canonical model of $\mathcal{T}$ in $\mathcal{P}(\mathcal{T}_\mathcal{C})$. First of all, we claim that the interpretation $M' : L \xrightarrow{M} \mathcal{C} \longrightarrow \mathcal{P}(\mathcal{T}_\mathcal{C})$ is a model of $\mathcal{T}$ in $\mathcal{P}(\mathcal{T}_\mathcal{C})$, i.e. it satisfies all axioms of T. Let $\phi \Rightarrow \psi$ be such an axiom; we want to see that $M'_{\vec{x}}(\phi) \leq M'_{\vec{x}}(\psi)$. Let $M'_{\vec{x}}(\phi) := A \hookrightarrow X$, $M'_{\vec{x}}(\psi) := B \hookrightarrow X$ in $\mathcal{P}(\mathcal{T}_\mathcal{C})$. Suppose $A \not\leq B$. By the completeness theorem applied to the logical category $\mathcal{P}(\mathcal{T}_\mathcal{C})$, there is a **Set** model $F : \mathcal{P}(\mathcal{T}_\mathcal{C}) \to \mathbf{Set}$ such that $F(A) \not\leq F(B)$. But then for the model $N : \mathcal{C} \longrightarrow \mathcal{P}(\mathcal{T}_\mathcal{C}) \xrightarrow{F} \mathbf{Set}$ of $\mathcal{T}$, $N \circ M$ does not satisfy the axiom $\phi \Rightarrow \psi$, contrary to the assumption (c.f. parenthetical phrase in (ii)). Thus $M' : L \to \mathcal{P}(\mathcal{T}_\mathcal{C})$ must indeed be a model of $\mathcal{T}$.

With $\varepsilon : \mathcal{C} \to \widetilde{\mathcal{C}}$ having the standard meaning, we obtain by (the proof of) 9.2.5 the equivalence $\mathrm{Coh}(\widetilde{\mathcal{C}}) \rightleftarrows \mathcal{P}(\mathcal{T}_\mathcal{C})$ and the morphism $\mathcal{P}(\mathcal{T}_\mathcal{C}) \to \widetilde{\mathcal{C}}$ such that all the following commute

$$\begin{tikzcd} \mathcal{C} \arrow[r] \arrow[rr, bend right] & \mathcal{P}(\mathcal{T}_\mathcal{C}) \arrow[r] & \widetilde{\mathcal{C}}, \end{tikzcd} \quad \begin{tikzcd} \mathrm{Coh}(\widetilde{\mathcal{C}}) \arrow[r] \arrow[rr, bend right, "\text{incl.}"'] & \mathcal{P}(\mathcal{T}_\mathcal{C}) \arrow[r] & \widetilde{\mathcal{C}}, \end{tikzcd} \quad \mathcal{P}(\mathcal{T}_\mathcal{C}) \longrightarrow \mathrm{Coh}(\widetilde{\mathcal{C}}) \xrightarrow{\text{incl.}} \widetilde{\mathcal{C}},$$

Similarly, we get the left-hand side of the diagram, with symmetric commutation properties. Here $\mathcal{E}(\mathcal{T})$ is the classifying topos of $\mathcal{T}$; it is a coherent topos. The logical functor $\mathcal{P}(\mathcal{T}) \to \mathcal{P}(\mathcal{T}_\mathcal{C})$ is induced by the universal property of $\mathcal{P}(\mathcal{T})$ over $\mathcal{T}$ by the fact that the composite $L \xrightarrow{M} \mathcal{C} \longrightarrow \mathcal{P}(\mathcal{T}_\mathcal{C})$ is a model of T; the rectangle

$$\begin{tikzcd} & L \arrow[r] \arrow[dl] & \mathcal{C} \arrow[dr] & \\ \mathcal{P}(\mathcal{T}) \arrow[rrr, bend right] & & & \mathcal{P}(\mathcal{T}_\mathcal{C}) \end{tikzcd}$$

will commute.

Finally, the continuous functor $u^* : \mathcal{E}(\mathcal{T}) \to \widetilde{\mathcal{C}}$ is derived from the universal property of $\mathcal{E}(\mathcal{T})$ over $\mathrm{Coh}(\mathcal{E}(\mathcal{T}))$ (c.f. 9.2.2(iii)) using $\mathcal{P}(\mathcal{T}) \to \mathcal{P}(\mathcal{T}_\mathcal{C})$; we will have that the diagram

$$\begin{tikzcd} \mathrm{Coh}(\mathcal{E}(\mathcal{T})) \arrow[r] \arrow[d, hook, "\text{incl.}"'] & \mathcal{P}(\mathcal{T}) \arrow[r] & \mathcal{P}(\mathcal{T}_\mathcal{C}) \arrow[r] & \mathrm{Coh}(\widetilde{\mathcal{C}}) \arrow[d, hook, "\text{incl.}"'] \\ \mathcal{E}(\mathcal{T}) \arrow[rrr, "u^*"] & & & \widetilde{\mathcal{C}} \end{tikzcd}$$

is commutative. Now, we can apply 9.2.9 to u^*. The assumption of the equivalence

$$\widehat{M} : \mathrm{Mod}(\mathcal{C}, \mathbf{Set}) \xrightarrow{\sim} \mathrm{Mod}(\mathcal{T}, \mathbf{Set})$$

immediately lifts to an equivalence

$$\widehat{u^*} : \mathrm{Mod}(\widetilde{\mathcal{C}}, \mathbf{Set}) \xrightarrow{\sim} \mathrm{Mod}(\mathcal{E}(\mathcal{T}), \mathbf{Set})$$

hence by 9.2.9, u^* is an equivalence as well. This proved theorem.

§3 The Zariski topos

We consider an example for classifying topos.

In the following discussion 'ring' means 'commutative ring with 1'.

Let $\mathcal{T}$ be the (coherent) theory of nontrivial local rings. $\mathcal{T}$ is formulated in the language L whose nonlogical symbols are the following operation symbols.

$$\begin{array}{ll} 0, 1 & \text{(0-ary)} \\ +, -, \cdot & \text{(binary)} \end{array}$$

L has only one sort, "the underlying set of a ring". The axioms of $\mathcal{T}$ are as follows:

$$\begin{array}{l} \text{axioms for a commutative ring with 1} \\ 0 = 1 \Rightarrow \quad , \\ \qquad\quad \Rightarrow \exists y(x \cdot y \approx 1) \vee \exists y((1-x) \cdot y \approx 1). \end{array}$$

An interpretation of the language L in a topos $\mathcal{E}$, $M : L \to \mathcal{E}$, consists of an object $M(s)$ of the topos, together with appropriate morphisms $M(0), M(1) : 1_{\mathcal{E}} \rightrightarrows M(s)$, $M(+), M(-), M(\cdot) : M(s) \times M(s) \rightrightarrows M(s)$. If the interpretation satisfies the axioms of rings, we briefly say that we have a 'ring object in $\mathcal{E}$'. Similarly, we can talk about a local ring object, etc. We clearly have

9.3.1 *The models (in* **Set***) of $\mathcal{T}$ are exactly the nontrivial local rings.*

We now identify the classifying topos $\mathcal{E}(\mathcal{T})$ of $\mathcal{T}$ as the well-known Zariski topos, showing a result due to Hakim [1972]. The application of 9.2.10(ii) for this purpose was suggested to us by Chris Mulvey.

Let $\mathcal{R}_f$ be the category of finitely presented rings (= quotients of polynomial rings $\mathbb{Z}[x_1, \ldots, x_n]$ (with $\mathbb{Z}$ the ring of integers) by ideals generated by finitely many polynomials $f(x_1, \ldots, x_n)$ in $\mathbb{Z}[x_1, \ldots, x_n]$).

The category $\mathcal{C}$ is defined as the opposite of $\mathcal{R}_f$, $\mathcal{C} = \mathcal{R}_f^{\text{opp}}$. We make $\mathcal{C}$ into a site by introducing the *Zariski topology* on $\mathcal{C}$ as follows. Note that a covering family $(A_i \to A)_i$ in $\mathcal{C}$ becomes a 'cocovering' family $(A \to A_i)_i$ in $\mathcal{R}_f$. Now, the Zariski topology on $\mathcal{C}$ is generated by the following cocovering families (i) and (ii):

(i) the empty family 'cocovering' the zero ring;

(ii) for any $A \in \mathrm{Ob}(\mathcal{R}_f)$,

$$\begin{tikzcd} & A[\frac{1}{a}] \\ A \arrow[ur] \arrow[dr] & \\ & A[\frac{1}{b}] \end{tikzcd}$$

whenever $a, b \in A$, $a + b = 1$.

In (ii) $A[\frac{1}{a}]$ is obtained by localization, or by introducing an inverse $\frac{1}{a}$ generically, i.e., $A \to A[\frac{1}{a}]$ has the following universal property:

$$\begin{tikzcd} A \arrow[r] \arrow[dr, "f"'] & A[\frac{1}{a}] \arrow[d, dashed] \\ & B \end{tikzcd}$$

in the diagram, whenever in B b is an inverse of $f(a)$, there is a unique $A[\frac{1}{a}] \to B$ mapping $\frac{1}{a} \in A[\frac{1}{a}]$ into b, and making the diagram commute.

Note that the Zariski topology on $\mathcal{C}$ can be generated by precisely two covering families as follows:

(i) as above

(ii)′
$$\begin{tikzcd} & \mathbb{Z}[x,\frac{1}{x}] \\ \mathbb{Z}[x] \arrow[ur, "x\mapsto x"] \arrow[dr, "x\mapsto x"'] & \\ & \mathbb{Z}\left[x,\frac{1}{1-x}\right] \end{tikzcd}$$

where (ii)′ is self-explanatory. The reason is that any cocovering in (ii) can be obtained by 'push out' in $\mathcal{R}_f$, i.e., by pullback in $\mathcal{C}$, from (ii)′, hence the topology generated by (i) and (ii)′ contains each covering (ii).

The Zariski topos $\mathcal{Z}$ is $\widetilde{\mathcal{C}}$, with the site $\mathcal{C}$ specified above. Let $\varepsilon:\mathcal{C}\to\widetilde{\mathcal{C}}$ be the canonical functor.

We can interpret the language in the category $\mathcal{C}$ by $M:L\to\mathcal{C}$ as follows. Again, remember that a morphism $A\to B$ in $\mathcal{C}$ is a morphism $B\to A$ in $\mathcal{R}_f$. $M(s)$ (s the unique sort of L) is defined as $\mathbb{Z}[x]$. The terminal object 1 of $\mathcal{C}$ is the initial object $\mathbb{Z}$ of $\mathcal{R}_f$. Accordingly, $M(0)$ is defined as the unique morphism $\mathbb{Z}[x]\xrightarrow{\alpha}\mathbb{Z}$ such that $\alpha(x)=0$. Similarly for $M(1)$. The product $\mathbb{Z}[x]\times\mathbb{Z}[x]$ in $\mathcal{C}$ is the coproduct $\mathbb{Z}[x]\cup\mathbb{Z}[x]$ in $\mathcal{R}_f$; and the latter is $\mathbb{Z}[x_1,x_2]$. $M(+)$ is defined as the unique morphism

$$\mathbb{Z}[x]\xrightarrow{\alpha}\mathbb{Z}[x_1,x_2]$$

such that $\alpha(x)=x_1+x_2$. Similarly for '$-$' and '$\cdot$'. The reader is invited to check that the interpretation $M:L\to\mathcal{C}$ just defined satisfies all axioms of commutative rings with 1. (A related, more general fact is stated in Section 4.)

The 'generic model' of $\mathcal{T}$ in the topos $\widetilde{\mathcal{C}}$ will be the composite $M_0=\varepsilon\circ M:L\xrightarrow{M}\mathcal{C}\xrightarrow{\varepsilon}\widetilde{\mathcal{C}}$. Before we show this using 9.2.10(ii), we give another description of M_0. We note that the representable presheaves over $\mathcal{C}$ are already sheaves. So $M_0(s)$ is the representable *sheaf* $h_{M(1)}=h_{\mathbb{Z}[x]}$ over $\mathcal{C}$. For an object A of $\mathcal{C}=\mathcal{R}_f^{\mathrm{opp}}$, $h_{\mathbb{Z}[x]}(A)=\mathrm{Hom}_{\mathcal{C}}(A,\mathbb{Z}[x])=\mathrm{Hom}_{\mathcal{R}_f}(\mathbb{Z}[x],A)$ which last set can be identified with the *underlying set* $|A|$ of A: the elements $a\in|A|$ are in one-one correspondence with homomorphisms $\mathbb{Z}[x]\to A$. So, $M_0(s)$ is "the underlying set-functor" on $\mathcal{R}_f$. The operations in L have similar natural meanings in M_0.

Next we show that M_0 is indeed generic. First, ignore the topology on $\mathcal{C}$ and consider the left exact functors $\mathcal{C}\to\mathbf{Set}$, or briefly the $\mathcal{C}$-algebras. Let $\mathrm{Alg}(\mathcal{C})$ be the category of all $\mathcal{C}$-algebras (full subcategory of the functor-category $\mathbf{Set}^{\mathcal{C}}$). Given an algebra $F:\mathcal{C}\to\mathbf{Set}$, the composition $F\circ M:L\to\mathbf{Set}$:

$$\begin{tikzcd} L \arrow[r, "M"] \arrow[dr, "F\circ M"'] & \mathcal{C} \arrow[d, "F"] \\ & \mathbf{Set} \end{tikzcd}$$

is a ring; this is so because M is a $\mathcal{C}$-ring (i.e., $M:L\to\mathcal{C}$ satisfies the ring axioms) as we said above, moreover F is left exact and the ring axioms are formulated in terms of finite left limits only. Actually, more is true viz.

9.3.2 *The functor $F\mapsto F\circ M$ (compositon by M) establishes an equivalence between the categories* $\mathrm{Alg}(\mathcal{C})$ *and* $\mathcal{R}$, *the category of all commutative rings with* 1.

We leave the verification of this to the reader (c.f. also Section 4). We just note that the point is that (i) $F\circ M$ is the 'restriction' of F to the ring $(\mathbb{Z}[x],0,1,+.-,\cdot)$ and actually (ii) F is determined by this restriction i.e., by the effects of F on $\mathbb{Z}[x],0,1:1\rightrightarrows\mathbb{Z}[x]$, $+,-,\cdot:\mathbb{Z}[x]\times_{\mathcal{C}}\mathbb{Z}[x]\rightrightarrows\mathbb{Z}[x]$.

Now, the category of *models* $\mathcal{C} \to \mathbf{Set}$, with $\mathcal{C}$ the site with the Zariski topology, is the full subcategory of $\mathrm{Alg}(\mathcal{C})$ consisting of those algebras $F : \mathcal{C} \to \mathbf{Set}$ which, in addition, are continuous with respect to the Zariski topology, or equivalently carry each basic covering in $\mathcal{C}$ into a "true" covering in $\mathbf{Set}$.

Claim 9.3.3 *For any $F \in \mathrm{Ob}(\mathrm{Alg}(\mathcal{C}))$, $F : \mathcal{C} \to \mathbf{Set}$ is a model for the Zariski site $\mathcal{C}$ iff $F \circ M$ is a local ring.*

We argue as follows. Let F be an arbitrary algebra $\mathcal{C} \to \mathbf{Set}$. Denote $F(\mathbb{Z}[x])$ by A. The ring $F \circ M$ is nothing but $(A, 0_A, 1_A, +_A, -_A, \cdot_A)$ where, e.g., 1_A is the element $1_{\mathbf{Set}} \to A$ which is the F-image of the $\mathcal{R}_f$-morphism $\mathbb{Z}[x] \xrightarrow{x \mapsto 1} \mathbb{Z}$, and $\cdot_A$ is the operation $A \times A \to A$ which is the F image of the $\mathcal{R}_f$-morphism $\mathbb{Z}[x] \xrightarrow{x \mapsto x_1 \cdot x_2} \mathbb{Z}[x_1, x_2]$, etc. Below, we drop the subscripts A from the operations in the ring $F \circ M$ and we also write A for the ring $F \circ M$ itself. Next, we invite the reader to check that in the following commutative diagram in $\mathcal{R}_f$:

$$\begin{tikzcd}
\mathbb{Z} \arrow[r] & \mathbb{Z}[x, \frac{1}{x}] \\
\mathbb{Z}[x] \arrow[u, "x\mapsto 1"] \arrow[ur, "f_1", "x\mapsto x"'] \arrow[r, "x\mapsto x_1\cdot x_2"'] & \mathbb{Z}[x_1, x_2] \arrow[u, "x_1\mapsto x", "x_2\mapsto \frac{1}{x}"']
\end{tikzcd}$$

the rectangle is a push out (dual of pullback). Reversing the arrows we obtain a pullback diagram in $\mathcal{C}$. Since the algebra F preserves pullback of $\mathcal{C}$, we obtain that in

$$\begin{tikzcd}
1_{\mathbf{Set}} \arrow[d, "1_A"'] & F(\mathbb{Z}[x, \frac{1}{x}]) = B_1 \arrow[l] \arrow[dl, "F(f_1)"'] \arrow[d] \\
A & A \times A \arrow[l, "\cdot_A"]
\end{tikzcd}$$

the rectangle is a pullback, i.e., B as a subset of $A \times A$ is $\{\langle a, a' \rangle \in A \times A : a \cdot a' = 1\}$, with $F(f_1)$ being the restriction to B of the projection $\langle a, a' \rangle \mapsto a$. We have obtained that

9.3.4 *The image of the map $F(f_1)$ is precisely the set of invertible elements of A.*

Similarly, we get for $f_2 : \mathbb{Z}[x] \xrightarrow{x \mapsto x} \mathbb{Z}\Big[x, \frac{1}{1-x}\Big]$ we have

9.3.5 *The image of $F(f_2)$ is the set of a in A such that $1 - a$ is invertible.*

Now, by looking at the cocovering (ii)$'$ in the Zariski site, by 9.3.4 and 5 we see that F carries this into a 'real' covering in the sense of $\mathbf{Set}$ (an effective epimorphic family in $\mathbf{Set}$) *iff* for every $a \in A$, either a or $1-a$ is invariable, i.e., *iff* A is a local ring. Similarly, F will carry (i) into a 'real' covering *iff* F is a nontrivial ring $(0 \neq 1)$. Now recall that for a left exact $F : \mathcal{C} \to \mathbf{Set}$ to be continuous it suffices that F preserve the covering *generating* the topology, c.f. 1.1.5. Thus, we have a proof of 9.3.3.

Now, $\mathrm{Mod}(\mathcal{C}, \mathbf{Set})$ is the full subcategory of $\mathrm{Alg}(\mathcal{C})$ whose objects are models of $\mathcal{C}$, the category of local rings; $\mathrm{Mod}(\mathcal{T}, \mathbf{Set})$ is the full subcategory of the category of all rings whose objects are the local rings. Hence, putting 9.3.2 and 9.3.3 together we obtain the the functor $F \mapsto F \circ M$ establishes an equivalence

$$\mathrm{Mod}(\mathcal{C}, \mathbf{Set}) \xrightarrow{\sim} \mathrm{Mod}(\mathcal{T}, \mathbf{Set}).$$

But this is precisely the condition in 9.2.10(ii). Hence $M_0 = \varepsilon \circ M : L \to \mathcal{Z}$ for the Zariski topos $\mathcal{Z}$ is indeed the 'generic model' of $\mathcal{T}$.

Recapitulating the definition, the fact just verified means the following. Given any Grothendieck topos $\mathcal{E}$ with a local ring object A (i.e., an interpretation $L \xrightarrow{A} \mathcal{E}$ of the language of rings which satisfies the axioms for 'local rings'), there is a geometric morphism $U = (u_*, u^*, \phi) : \mathcal{E} \to \mathcal{Z}$, with the Zariski topos $\mathcal{Z}$, such that $u^* : \mathcal{Z} \to \mathcal{E}$ carries the generic local ring M_0 into A, moreover, U is essentially unique.

We hasten to add that the application of something like 9.2.10(ii) is far from being essential in the above verification. Rather, the situation is as in a common kind of application of the ordinary completeness theorem: one can use completeness to conclude that a particular sentence is derivable from a particular theory, but usually on can actually (with little or much work, depending on the situation) exhibit such a deduction, thereby eliminating the application of completeness. In fact, for the Zariski topos, it would not be hard to show directly that it is the classifying topos for the theory of local rings.

Actually, the content of the statement: "points are enough for classifying" specialized to the present situation is more faithfully expressed by saying that whenever the above argument involving local rings can be repeated for some arbitrary algebraic site $\mathcal{C}$, then $\widetilde{\mathcal{C}}$ is necessarily equivalent to the Zariski topos.

We note that quite similarly we can show that the Etale topos is the classifying topos of the theory of separably closed local rings (c.f. Wraith [?]). The Etale site is the category of all affine schemes with the Etale topology. Actually, all we need is the Etale equivalent of 9.3.3 above. The details are not given here.

§4 Appendix. Coste's construction of the classifying topos of a theory

In Coste and Coste [1975], there is a new construction of the classifying topos of a finitary coherent theory. In this section we extend this construction to an $L_{\infty\omega}$ theory, but for the particular case of a language with operation symbols only. The reason of this restriction is to compare the general theorem with the method of SGA 4 to construct classifying ringed topos, base on localizations on $\mathcal{R}_f^{\mathrm{opp}}$, the dual of the category of finitely presented rings.

We shall show that, under this restriction on languages, the natural extension of the procedure of SGA 4 is quite general and that localizations (Grothendieck topologies) differ only notationally form coherent axiomatizations.

We need first to recall some facts about Universal Algebra which, although widely known do not seem to be readily available in the literature.

Let L be a language with operation symbols only and let T_0 be any *equational* theory, i.e., whose axioms are coherent sequents of the form

$$\Rightarrow t = t'$$

where t, t' are terms of L.

A *finitely presented* T_0-algebra can be defined in one of the following equivalent ways:

1. (Gabriel-Ulmer [1971]) as a T_0-algebra (i.e. a model of T_0) A such that the representable functor $h^A : \mathrm{Mod}(T_0) = T_0\text{-algebras} \to \mathbf{Set}$ preserves filtered $\varinjlim$.

2. ("Conventional" way) as a T_0-algebra of the form $F[x_1, \ldots, x_n]/E$, where $F[x_1, \ldots, x_n]$ is the free T_0-algebra on a finite set of indeterminates, each having a given sort

and E is a congruence relation *defined* by a finite set Φ of equalities between "polynomials", i.e., $f \equiv g\ (E)$ iff $T_0 \vdash \Phi \Rightarrow f = g$.

A word about this equivalence: 2. $\Rightarrow$ 1. is trivial; to show 1. $\Rightarrow$ 2. one requires the fact (pointed out to us by A. Kock) that a retract of a finitely generated T_0-algebra in the sense of 2. is again finitely generated.

The category $\mathcal{C}_0$, the dual of the (full) subcategory of finitely presented T_0-algebras has a simple syntactical description. Indeed, define $\mathcal{C}(T_0)$, the category associated to T_0 as follows: as *objects* we take finite sets Φ of atomic formulas of L, i.e., finite sets of equalities between terms; as *morphisms* $\Phi(x_1, \dots, x_n) \to \Psi(y_1, \dots, y_m)$ equivalence classes of m-tuples of terms $\langle t_1, \dots, t_m\rangle$ with free variables among $x_1, \dots, x_n$ such that $T_0 \vdash \Phi \Rightarrow \Psi(t_1/y_1, \dots, t_m/y_m)$, under the equivalence relation $\langle t_1, \dots, t_m\rangle \sim \langle s_1, \dots, s_m\rangle$ iff $T_0 \vdash t_i = s_i$ for $i = 1, \dots, m$. Composition is defined by means of substitution in the obvious way.

Proposition 9.4.1 *$\mathcal{C}(T_0)$ is equivalent to $\mathcal{C}_0$, the dual of the category of finitely presented T_0-algebras.*

We sketch the proof. With the (finite) set $\Phi(x_1, \dots, x_n)$ of equalities, we associate "its coordinate algebra" $F[x_1, \dots, x_n]/C(\Phi)$, where $C(\Phi)$ is the congruence defined by Φ (see 2.) of the definition of finitely presented T_0-algebra). If $\langle t_1, \dots, t_m\rangle$ is a representative of a morphism $\Phi(x_1, \dots, x_n) \to \Psi(y_1, \dots, y_m)$ and $f \in F[y_1, \dots, y_m]$, we define a function $F[y_1, \dots, y_m] \to F[x_1, \dots, x_n]/C(\Phi)$ by sending $f(y_1, \dots, y_m)$ into the class of $f(t_1(x_1, \dots, x_n), \dots, t_m(x_1, \dots, x_n))$. This function factors through $F[y_1, \dots, y_m]/C(\Psi)$, precisely because of the definition of $C(\Psi)$. Indeed, let $f \equiv g\ C(\Psi)$. Hence $T_0 \vdash \Psi \Rightarrow f = g$. But $T_0 \vdash \Phi \Rightarrow \bigwedge\Psi(t_1/y_1, \dots, t_m/y_m)$. Therefore $T_0 \vdash \Phi \Rightarrow f(t_1/y_1, \dots, t_m/y_m) = g(t_1/y_1, \dots, t_m/y_m)$, i.e.,

$$f(t_1/y_1, \dots, t_m/y_m) \equiv g(t_1/y_1, \dots, t_m/y_m)\ (C(\Phi)).$$

We have shown that our association "coordinate algebra of" is functorial. This functor is (obviously) essentially surjective and faithful (exercise!). Let us check that it is full. In fact, let $F[x_1, \dots, x_n]/C(\Phi) \xrightarrow{f} F[y_1, \dots, y_m]/C(\Psi)$ be given. By composition, we obtain $F[x_1, \dots, x_n] \xrightarrow{\mathcal{T}_\Phi} F[x_1, \dots, x_n]/C(\Phi) \longrightarrow F[y_1, \dots, y_m]/C(\Psi)$. Then one can find $s_1, \dots, s_n \in F[y_1, \dots, y_m]$ such that $\mathcal{T}_\Psi(s_i) = f \circ \mathcal{T}_\Phi(x_i)$, since $\mathcal{T}_\Psi$ is onto. The class of $\langle s_1, \dots, s_n\rangle$ gives the desired morphism $\Psi \to \Phi$.

With either of these two descriptions of $\mathcal{C}_0$ one easily concludes

Proposition 9.4.2 *$\mathcal{C}_0$ has finite $\varprojlim$ and the functors $\mathcal{C}_0 \to$ **Set** which preserve these $\varprojlim$ are precisely the T_0-algebras.*

Now we can state

Theorem 9.4.3 *Let T be a coherent theory in $L_{\infty\omega}$, with language L having operation symbols only and let*

$$\mathrm{Alg}(T) = \{\ \ \Rightarrow t = t' | t, t' \text{ are terms of } L \text{ and } T \vdash\ \ \Rightarrow t = t'\}.$$

Furthermore, let $T_0 \subseteq \mathrm{Alg}(T)$ be any equational theory in L and let $\mathcal{C}_0$ be the dual of the finitely presented T_0-algebras. Then there is a localization on $\mathcal{C}_0$ such that $\mathrm{Sh}(\mathcal{C}_0) \simeq \mathcal{E}[\mathcal{T}]$, the classifying topos of T.

PROOF. We need the following result whose proof is an easy induction on formulas and which was proved as 8.3.2.

Lemma 9.4.4 *Every coherent formula of a $L_{\kappa\omega}$ language L (with κ regular) is equivalent to a disjunction of the form $\bigvee\{\exists x_1, \ldots, \exists x_{n_i} \bigwedge \Phi_i : i \in I\}$, where* $\mathrm{card}(I) < \kappa$ *and Φ_i is a finite set of atomic formulas.*

By 9.3.4, we may assume that T has a coherent axiomatization of the form

$$\Phi \Rightarrow \bigvee\{\exists y_1, \ldots \exists y_{n_i} \bigwedge \Psi_i : i \in I\}$$

where Φ and Ψ_i are (finite) sets of atomic formulas.

Each axiom and theorem (of this form) gives rise to the obvious "co-covering" family (in $\mathcal{C}_0^{\mathrm{opp}}$) $(F[x_1, \ldots, x_n]/C(\Phi) \xrightarrow{\tau_i} F[x_1, \ldots, x_n, y_1, \ldots, y_{n_i}]/C(\Phi \cup \Psi_i))_{i \in I}$, where $\langle x_1, \ldots, x_n \rangle$ is the sequence of free variables of the axiom (or theorem).

To show the stability of the covering families in $\mathcal{C}_0$ is equivalent to show that push-outs of "co-covering" families "co-cover".

The push-out of f and τ_i is given by the diagram

$$\begin{tikzcd}
F[z_1, \ldots, z_m]/C(\Theta) \arrow[r, "\tau_i'"] & F[z_1, \ldots, z_m, y_1, \ldots, y_{n_i}]/C(\Theta \cup \Psi_i') \\
F[x_1, \ldots, x_n]/C(\Phi) \arrow[u, "f"] \arrow[r, "\tau_i"'] & F[x_1, \ldots, x_n, y_1, \ldots, y_{n_i}]/C(\Phi \cup \Psi_i') \arrow[u, "f'"']
\end{tikzcd}$$

obtained as follows: the fullness of the functor "coordinate algebra of" gives us an n-tuple $\langle s_1, \ldots, s_n \rangle$ of elements of $F[z_1, \ldots, z_m]$ whose class is the morphism $\Phi \to \Theta$ sent by that functor into f. We let $\Psi_i' = \Psi_i(s_1/x_1, \ldots, s_n/x_n)$ and we let f' be the morphism which that functor associates with (the class of)

$$\langle s_1, \ldots, s_n, y_1, \ldots, y_{n_i} \rangle : \Theta \cup \Psi_i' \to \Phi \cup \Psi_i.$$

Notice that $T_0 \vdash \Theta \Rightarrow \Phi(s_1/x_1, \ldots, s_n/x_n)$, since the class of $\langle s_1, \ldots, s_n \rangle$ is a morphism. Furthermore, since

$$T \vdash \Phi \Rightarrow \bigvee\{\exists y_1, \ldots, \exists y_{n_i} \bigwedge \Psi_i : i \in I\}$$

it follows that

$$T \vdash \Phi(s_1/x_1, \ldots, s_n/x_n) \Rightarrow \bigvee\{\exists y_1, \ldots, \exists y_{n_i} \bigwedge \Psi_i' : i \in I\},$$

which means that $(\tau_i')_{i \in I}$ "co-covers".

To finish the proof, we notice that the continuous finite $\varprojlim$ preserving functors are precisely those T_0-algebras A such that every map

$$F[x_1, \ldots, x_n]/C(\Phi) \to A$$

factors through some

$$F[x_1, \ldots, x_n]/C(\Phi) \to F[x_1, \ldots, x_n, y_1, \ldots, y_{n_i}]/C(\Phi \cup \Psi_i),$$

which is precisely the condition

$$A \models \Phi \Rightarrow \bigvee\{\exists y_1, \ldots, y_{n_i} \bigwedge \Psi_i : i \in I\}.$$

As an *example*, let us consider the coherent theory of local rings, i.e., whose axioms are (besides those of the theory of rings)

(i) $\quad 0 = 1 \Rightarrow$

(ii) $\Rightarrow \bigvee\{\exists y(xy = 1), \exists y(x(1-y) = 1)\}$.

We take for T_0 the (equational) theory of rings. The localization in question is obtained by closing under push-outs the following "co-covering" families in $\mathcal{R}_f$, the category of finitely presented rings:

(i)′ $\mathbb{Z}/C(0 = 1)$ is co-covered by the empty family.

(ii)′

$$\begin{tikzcd}[row sep=tiny] & \mathbb{Z}[X,Y]/C(XY = 1) \\ \mathbb{Z}[X] \arrow[ur] \arrow[dr] & \\ & \mathbb{Z}[X,Y]/C((1-X)Y = 1) \end{tikzcd}$$

Spelling out these cocoverings in the more familiar language of ideals, we obtain the usual Zariski localization (and a "new proof" of Hakims's theorem, §3 of this chapter).

If we wish to eliminate non-zero nil-potent elements from this theory, we add the families

(iii)′ $\mathbb{Z}[X]/(X^n) \to \mathbb{Z}$, for each $n = 1, 2, \ldots$

(where the class of X goes to 0), obtained from the logical conditions

(iii) $X^n = 0 \Rightarrow X = 0$, for each $n = 1, 2, \ldots$.

§5 Appendix

In his paper [1975], Lawvere gives another definition of coherent object (in coherent topos) and asserts the equivalence of the two. His definition is closer in spirit to the original Serre's definition of an algebraic coherent sheaf (c.f. Serre [1955]).

An object A of a topos $\mathcal{E}$ is κ-presented (in the sense of Gabriel-Ulmer [1970]) if $h^A : \mathcal{E} \to \mathbf{Set}$ preserves κ-filtered $\varinjlim$. An object A is L_κ-coherent if given any two κ-presented objects P_1, P_2 and maps $P_1 \to A$, $P_2 \to A$, the pull-back $P_1 \times_A P_2$ is again κ-presented.

Proposition 9.5.1 (Lawvere [1975]). *Let $\mathcal{E}$ be a κ-coherent topos, $A \in |\mathcal{E}|$. Then A is κ-coherent iff A is L_κ-coherent.*

Proof. Let $\mathcal{E} \simeq \mathrm{Sh}(\mathcal{C})$, for some κ-algebraic site $\mathcal{C}$. From the "κ-version" of Theorem 1.23 [SGA 4, Exposé VI] we have

Lemma 9.5.2 *Let $\mathcal{E}$ be a topos and $A \in |\mathcal{E}|$.*

(i) *A is κ-quasi-compact iff for every κ-filtered inductive system $(X_i)_{i\in I}$, the natural application*

$$\varinjlim_i h^X(Y_i) \to h^X(\varinjlim_i Y_i)$$

is injective.

(ii) *If $\mathcal{E}$ is κ-coherent and A is κ-coherent, then A is κ-presented (i.e., the above application is bijective).*

Let A be a L_κ-coherent object. We claim that A can be covered by a κ-coherent object $C \to A$. In fact, from (i) of 9.5.2, A is κ-quasi-compact and hence it can be jointly covered by a family $(\varepsilon C_i)_{i\in I}$ of cardinality $< \kappa$. By 9.2.6, the C_i are κ-coherent.

The κ-version of Corollary 1.15 [loc. cit.] gives

Proposition 9.5.3 *A coproduct of a family of* card $< \kappa$ *of κ-coherent objects is κ-coherent.*

Letting now $C = \coprod_{i\in I} C_i$, we have that C is κ-coherent and $C \to A$.

Using (ii) of 9.5.2 we conclude that C is κ-presented and hence so is $C \times_A C$ (by the definition of A being L_κ-coherent). From 9.5.2(i) it follows that $C \times_A C$ is κ-quasi-compact.

In other words, we have shown that A can be exactly covered (in the sense of 1.4.11) via

$$Y : \mathrm{Coh}_\kappa(\mathcal{E}) \to \mathcal{E}.$$

By Lemma 1.4.11, A is κ-coherent.

For the converse, let A be κ-coherent. By 9.5.2(ii), it is κ-presented. Assume that P_1, P_2 are κ-presented and let $P_1 \to A$, $P_2 \to A$ be given. We claim that $P_1 \times_A P_2$ is κ-presented. We shall use the following "κ-version" of Corollary 1.24.2 of SGA 4, Exposé VI:

Lemma 9.5.4 *Let $\mathcal{E}$ be a coherent topos.*

(i) *An object A of $\mathcal{E}$ is κ-presented off there is a coequalizer diagram* $B \twoheadleftarrow C_0 \leftleftarrows C_1$ *with C_0, C_1 κ-coherent and a morphism $B \to A$ such that* $A \twoheadleftarrow B \overset{1_B}{\underset{w}{\leftleftarrows}} B$ *is again a co-equalizer and $w^2 = w \circ w = w$.*

(ii) *The full subcategory of κ-presented objects has $\varprojlim$ indexed by families of cardinality less that κ.*

To return to the proof of our proposition, assume first that P_1 is κ-coherent. By 9.5.4 applied to P_2, there is a diagram

$$\begin{tikzcd} A & P_2 \arrow[l] & B \arrow[l, two heads] & B \arrow[l, shift right, "1_B"'] \arrow[l, shift left, "w"] & C_0 \arrow[l, two heads] & C_1 \arrow[l, shift left] \arrow[l, shift right] \end{tikzcd}$$

such that C_0, C_1 are κ-coherent and both diagrams

$$\begin{tikzcd} P_2 & B \arrow[l, two heads] & B, \arrow[l, shift right, "1_B"'] \arrow[l, shift left, "w"] \end{tikzcd} \qquad \begin{tikzcd} B & C_0 \arrow[l, two heads] & C_1 \arrow[l, shift left] \arrow[l, shift right] \end{tikzcd}$$

are co-equalizers with $w^2 = w$.

Pulling back this diagram along $P_1 \to A$ and noticing the stability of $\varinjlim$ under pull-backs we obtain the new diagram

$$\begin{tikzcd} P_1 \arrow[d] & P_1 \times_A P_2 \arrow[l] \arrow[d] & P_1 \times_A B \arrow[l, two heads] \arrow[d] & P_1 \times_A B \arrow[l, shift right, "1"'] \arrow[l, shift left, "w'"] \arrow[d] & P_1 \times_A C_0 \arrow[l, two heads] & P_1 \times_A C_1 \arrow[l, shift left] \arrow[l, shift right] \arrow[d] \\ A & P_2 \arrow[l] & B \arrow[l, two heads] & B \arrow[l, shift right, "1_B"'] \arrow[l, shift left, "w"] & C_0 \arrow[l, two heads] & C_1 \arrow[l, shift left] \arrow[l, shift right] \end{tikzcd}$$

where ${w'}^2 = w'$ and all co-equalizers are preserved. Since $P_1 \times_A C_0$, $P_1 \times_A C_1$ are κ-coherent by Theorem 9.1. , $P_1 \times_A P_2$ is κ-presented, using 9.5.4(i) again.

The general case now follows from this diagram. Indeed, by the special case $P_1 \times_A C_0$ and $P_1 \times_A C_1$ are now κ-presented and part (ii) of 9.5.4 implies that $P_1 \times_A P_2$ is the κ-presented.

References

Antonius, W.
[1975] Théories cohérentes et prétopos, Thèse de Maîtrise, Université de Montréal.

[SGA4] Artin, M; A. Grothendieck; J.L. Verdier
[1972] Théorie des Topos et Cohomologie Etale des Schémas, Lecture Notes in Mathematics, vol. 1 no. 269, vol. 2: no. 270 (Springer Verlag).

Atiyah, M.F., and I.G. MacDonald
[1969] Introduction to Commutative Algebra (Addison-Wesley Publ. Co.)

Barr, M.
[1974] Toposes without points, J. Pure and Applied Algebra 5, pp. 265-280.

Barwise, J.
[1967] Infinitary logic and admissible sets. Doctoral Dissertation, Stanford University.

[1975] Admissible sets and Structures (Springer Verlag).

Bénabou, J.
[1973] Structures syntaxiques, notes by Roch Ouellet, April 1973.

[CK] Chang, C.C. and H.J. Keisler
[1973] Model Theory (North Holland).

Coste, M.
[1973] Logique du 1er ordre dans les topos élémentaires, Séminaire Bénabou, Paris.

Coste, M. and Coste, M.-F.,
[1975] Théories cohérentes et topos cohérents, Séminaire Bénabou, Paris.

Dionne, J.
[1973] Des théories élémentaires aux catégories conceptuelles, et Addendum, Thèse de M.Sc. Univ. de Montréal, July.

Feferman, S.
[1968] Lectures on proof theory, in: Lecture Notes in Mathematics, v. 70, pp. 1-107. (Springer Verlag).

Fitting, M.
[1969] Intuitionistic logic, model theory and forcing (North-Holland).

[1973] Model existence theorems for modal and intuitionistic logics, J. Symbolic logic 38, pp. 613-627.

Foruman, M.

[1977] The logic of topoi, in: Handbook of Mathematical Logic, ed. K.J. Barwise. (North-Holland).

Gabriel, P. and F. Ulmer
[1971] Lokal präsentierbare Kategorien, Lecture Notes in Mathematics, v. 221 (Springer Verlag).

Hakim, M.
[1972] Topos annelés et schémas relatifs, Ergebnisse Band 64 (Springer Verlag).

Higgs, D.
[1973] A category approach to Boolean valued set-theory, preprint.

Joyal, A. and G.E. Reyes
[?] Model completion in categorical logic, to appear.

Karp, C.R.
[1964] Languages with expressions of infinite length (North-Holland).

Keisler, H.J.
[1971] Model Theory for Infinitary Logic (North-Holland).

Kleene, S.C.
[1967] Mathematical Logic (John Wiley and Sons, Inc.)

Kock, A. and G.E. Reyes
[1977] Doctrines in categorical logic, in: Handbook of Mathematical Logic, ed. J. Barwise (North Holland).

Kripke, S.
[1965] Semantical analysis of intuitionistic logic I, in: Formal Systems and Recursive Functions, North Holland, pp. 92-130

Lawvere, F.W.
[1963] Functorial semantics of algebraic theories, Proc. Nat. Acad. Sci. U.S.A. 50, pp. 869-872.

[1965] Functorial semantics of elementary theories, in: Logic Colloquium, Leicester 1965 (North Holland).

[1969] Adjointness in foundations, Dialectica 23, pp. 281-296.

[1975] Continuously variable sets: algebraic geometry = geometric logic, in: Logic Colloquium 73 Bristol (North Holland), pp. 135-156.

[CWM] MacLane, S.
[1971] Categories for the Working Mathematician (Springer Verlag).

Makkai, M.
[1969] On the model theory of denumerably long formulas with finite strings of quantifiers, J. Symbolic Logic 34, pp. 437-459.

Makkai, M. and G.M. Reyes
[1976] Model-theoretic methods in the theory of topoi and related categories, Bulletin de l'Acad. Polon. Sci., ser. sci. math. astr. phys. 24, Part I: pp. 379-384, Part II: pp. 385-392.

Mansfield, R.
[1972] The completeness theorem for infinitary logic, J. Symbolic Logic 37, pp. 31-34.

Mitchell, W.

[1972] Boolean topoi and the theory of sets, J. Pure Appl. Algebra 2, pp. 261-274.

Osius, G.

[1973] The internal and external aspects of logic and set theory in elementary topoi, Cahiers Topologie Géom. Différentielle, 14, pp. 47-49.

Rasiowa, H. and R. Sikorski

[1963] The mathematics of metamathematics, (Panstowe Wydawnictwo Naukowe, Warsaw).

Reyes, G.E.

[1974] From sheaves to logic, in: M.A.A. Studies in Mathematics, vol. 9, (ed. A. Daigneault).

Robitaille-Giguère, M.

[1975] Modèles d'une catégorie logique dans des topos de préfaisceaux d'ensembles de Heyting, Thèse de Maîtrise, Université de Montréal.

Rosser, B.

[1969] Simplified independence proofs (Academic Press).

Schubert, H.

[1972] Categories (Springer Verlag).

Serre, J.-P.

[1955] Faisceaux algèbraiques cohérentes, Annals of Math., 61, pp. 197-278.

Tarski, A.

[1952] Some notions and methods on the borderline of algebra and metamathematics, in: Proc. Intern. Congr. of Mathematicians, Cambridge, Mass. 1950, 1 (Am. Math. Soc.), pp. 705-720.

Wraith, G.C.

[?] Separably closed local rings, to appear.