Toposes, Sets, and Cohen Forcing

Todd Schmid

April 25, 2020

Abstract

Shortly after its invention in the 1960's, Cohen's method of forcing found applications and generalizations to nearly every branch of mathematical logic. Due to the work of A. Blass, A. Joyal, W. Lawvere, G. Reyes, A. Scedrov (see [16]), and several others during the late 1970's and early 1980's, it is now understood that these various forms of forcing are subsumed by the construction of classifying toposes in category theory. Still, none but the few who are truly initiated in both logic and sheaf theory might suspect there to be such a deep connection between Cohen's efforts and Grothendieck's topos theory. The following document is essentially intended to be a readable set of notes introducing the concepts involved, motivated by a topos theoretic discussion of the continuum hypothesis. In particular, Boolean valued models of set theory are given a terse overview, as well as the theory of set-valued sheaves on a site. Finally, a sheaf model of set theory is given in which the relevant version of the continuum hypothesis is violated.

Contents

Introduction

The word *forcing* is used in reference to wide array of different methods of constructing new models from old; so wide, perhaps, that a single definition could not account for every context in which it appears. There is, on the other hand, a common theme: In the vaguest of terms, one could describe forcing as a method for piecing together models of a given elementary theory in such a way that the properties of the resulting model are controlled by the interrelations of the pieces. Typically, the "pieces" are called *conditions* and their "interrelations" form a partial order.

To the trained eye, this description might hint at a connection with sheaf theory, and indeed it has been said that "the various types of forcing can all be performed using sheaves" (see [15]). Drawing this sort of connection took many years and a great amount of effort on the parts of the logicians working in the area. Due to the work of A. Blass, A. Joyal, W. Lawvere, G. Reyes, A. Scedrov, M. Tierney, and several others during the late 1970's and early 1980's, it is now understood that these various forms of forcing are subsumed by the construction of classifying toposes in category theory (see [16]).

Unfortunately, this description of forcing is beyond the scope of this document, but there do happen to be a few specific examples of the equivalence that are relatively easy to explain. The motivating example in this set of notes is a *Boolean sheaf topos* called the *Cohen topos*. The Cohen topos satisfies a generalized version of the axiom of choice (the *internal* AC, IAC), but violates a version of the continuum hypothesis appropriate for toposes. The construction of this category will closely resemble Cohen's construction of a model of ZFC with κ-many *(Cohen) reals*, where $\kappa > 2^{\mathbb{N}}$, and will even make use of the same poset.

Objectively, there is little logical content in the argument below that does not already appear in Cohen's proof of the independence of the Continuum hypothesis. The process of construction cosey follows the Boolean valued model approach used by Timothy Chow in [4] for pedagogical purposes and introduced by Scott, Solovay, and Vopenka (its first appearance in print is in J. Bell's book [3]): A ground model is fixed, a partially ordered set of conditions in the ground model is chosen with some desired property in mind, and a new model is constructed from the the poset's *generic obects*.

In the classical situation, the ground model is a countable transitive model of ZFC provided by the Completeness and Lowenheim-Skölem theorems from first-order logic, the poset is the *Cohen poset*, and the generic objects are *interpretations* of *names*. Below, the ground model is the whole category $\mathbb{S}$**ets** of sets and functions and the generic objects will be *double-negation sheaves* for the Cohen poset. It was noticed by

Higgs, intially recorded in a well-known but unpublished document in 1975 (see [6] for a more general approach than Higgs'), that such a category of double-negation sheaves on a poset is naturally equivalent to the category of Boolean valued sets, and moreover that this equivalence respects relativization. This indicates that Cohen's original proof, the Boolean-valued model approach, and the sheaf-theoretic construction all contain essentially the same logical content.

1 Boolean Valued Models

The current section is a terse overview of the Boolean-valued model approach to forcing. As such, it is intended only to remind the reader of the concepts involved, not to provide first-time instruction. Sources for further/initial learning include [4] and [3] for Boolean-valued models, and [10] and [7] and [5] for the classical approach to forcing.

1.1 Boolean-valued sets

In set theory, the prevailing dogma is that "everything is a set". In other words, a set X is a special kind of conglomorate of elements, each of which a set. This seems circular at first, but defining the empty collection $\varnothing$ to be a set, recursion resolves the apparent circularity. Precisely the same philosophy carries over to *Boolean-valued sets*, with "set" replaced by "Boolean-valued set". Fix a complete Boolean algebra $\mathbb{B}$ as well as the symbols $\top$ and $\bot$ for the top and the bottom elements of $\mathbb{B}$ respectively.

Definition 1.1. A $\mathbb{B}$-*valued set* is a function $X \to \mathbb{B}$, where X is a set of $\mathbb{B}$-valued sets.

For instance, the empty set $\varnothing$ is a $\mathbb{B}$-valued set, as well as every *(global) element* $(\{\varnothing\} =)1 \to \mathbb{B}$ of $\mathbb{B}$. Fix a *common* model (M, E) of ZFC, ie. a transitive model whose elements are well-founded sets, and assume that $\mathbb{B} \in M$.

Definition 1.2. Write

$$M^{\mathbb{B}} = \{f \in M \mid f \text{ is a } \mathbb{B}\text{-valued set}\},$$

and let $\mathcal{L}_M$ be the set of finitary first-order formulae in the language $\{e, \subseteq\} \cup \{a \mid a \in M^{\mathbb{B}}\}$. Recursively define a function $[\![-]\!] : \mathcal{L}_M \to \mathbb{B}$ called the *valuation* for $M^{\mathbb{B}}$ as follows:

(a) $[\![-]\!]$ preserves with $\wedge$, $\vee$, and $\neg$,

(**b**) for any $\varphi(x)$ in $\mathcal{L}_M$,
$$[\![(\exists x)(\varphi(x))]\!] = \bigvee_{a\in M^{\mathbb{B}}} [\![\varphi(a)]\!],$$
(**c**) for any $\varphi(x)$ in $\mathcal{L}_M$,
$$[\![(\forall x)(\varphi(x))]\!] = \bigwedge_{a\in M^{\mathbb{B}}} [\![\varphi(a)]\!],$$
(**d**) for any $\mathbb{B}$-valued sets x, y,
$$[\![x \subseteq y]\!] = \bigwedge_{a\in \mathrm{dom}(x)} (x(a) \Rightarrow [\![aEy]\!]),$$
(**e**) and where by $x = y$ we mean $(x \subseteq y) \wedge (y \subseteq x)$ for variables x and y,
$$[\![xEy]\!] = \bigvee_{a\in \mathrm{dom}(y)} ([\![x = a]\!] \wedge y(a)).$$

Conditions (d) and (e) in the previous definition are recursively referential to one another in the following sense: Let the rank of an element $a \in M^{\mathbb{B}}$ be $\mathrm{rank}(a) = \sup\{\mathrm{rank}(b) + 1 \mid b \in \mathrm{dom}(a)\}$. Computing $[\![xEy]\!]$ or $[\![x \subseteq y]\!]$ involves a calculation of the other, but for elements of a strictly lower rank than that of x and y. For example, $[\![\varnothing \subseteq y]\!]$ is an empty join giving $\top$, and $[\![\varnothing Ey]\!] = y(\varnothing)$ for any $y : 1 \to \mathbb{B}$.

Note that for the above definition to make sense one could replace the assumption that $\mathbb{B}$ is complete with the weaker assumption that $\mathbb{B}$ is *M-complete*, meaning that every subset of $\mathbb{B}$ indexed by a set in M has a supremum and infimum.

Theorem 1.3. *Let $[\![-]\!]$ be the valuation for $M^{\mathbb{B}}$. For any axiom φ of ZFC, $[\![\varphi]\!] = \top$.*

The proof can be found in [3]. Any use of the phrase *Boolean-valued model* is in reference to the pair $(M^{\mathbb{B}}, [\![-]\!])$.

1.2 Truth Values?

Now, while Theorem 1.3 states there is a sense in which $M^{\mathbb{B}}$ is a model of ZFC, $M^{\mathbb{B}}$ is certainly not a common model. One fundamental difference between $M^{\mathbb{B}}$ and a common model of ZFC is its set of *truth values*. The set $2 = \{0, 1\}$ in a standard model of ZFC can alternatively be thought of as the set $\{\mathrm{false}, \mathrm{true}\}$. This allows 2 to play the role of a *subset classifier*, the unique set (up to bijection) that satisfies the following property: For any set X and $U \subseteq X$, there is a unique function $\mathrm{ch}(U) : X \to 2$ such that $\mathrm{ch}(U)\restriction_U$ is the constant function true and to any other function $h : Y \to X$ such that $\mathrm{ch}(U) \circ h = \mathrm{true}$ there corresponds a unique factorization of h through U (in other words, $\mathrm{img}(h) \subseteq U$). In fact, any formula $\varphi(x)$ of the form

$$(x \in A \wedge \psi(x))$$

is given uniquely by the function

$$\varphi : A \to 2$$

it induces (and conversely, functions into 2 are identifiable with bounded formulas such as $\varphi(x)$). Since there are two *truth values* that such a function can obtain, one calls a common model of ZFC *two-valued*.

The model $M^{\mathbb{B}}$, on the other hand, is not usually two-valued. Powersets in $M^{\mathbb{B}}$ are constructed by setting $\mathrm{dom}(\mathcal{P}(y)) = \mathbb{B}^X$ for any $y : X \to \mathbb{B}$ in $M^{\mathbb{B}}$ and

$$\mathcal{P}(y)(x) = [\![x \subseteq y]\!]$$

for any $x \in \mathbb{B}^X$. Following the usual construction of 2 in $M^{\mathbb{B}}$, we obtain the Boolean-valued set $\Omega = \mathcal{P}(\mathcal{P}(\varnothing)) = \mathbb{B}^{\mathbb{B}^{\varnothing}}$, which is isomorphic to $\mathbb{B}$ in $M^{\mathbb{B}}$. In other words, $M^{\mathbb{B}}$ has one truth value for each element of $\mathbb{B}$.

The completeness theorem of first-order logic is a statement about two-valued models. Thus, relying on the completeness theorem to prove the dependence or independence of a statement of ZFC using Boolean-valued models requires a method of obtaining a two-valued model from a general Boolean-valued model.

1.3 From Boolean-valued models to two-valued models

Recall that a $\mathbb{B}$-valued set is intended to be an assignment of probability to each member of X. Obtaining a two-valued model of ZFC from $M^{\mathbb{B}}$ consists of separating the "high" probabilities from the "low" probabilities. This is best done with filters.

Definition 1.4. A subset U of a poset $\mathbb{P}$ is called a *filter* in $\mathbb{P}$ if

(**a**) $U \neq \mathbb{P}$,

(**b**) $x \leqslant y$ and $x \in U$ implies $y \in U$, and

(**c**) if $x, y \in U$, then $(\exists z \in U)((z \leqslant x) \wedge (z \leqslant y))$.

Furthermore, U is called an *ultrafilter* if U is a maximal filter.

In particular, if $\mathbb{P}$ has a *bottom* $\perp$, then $\perp \notin U$. Additionally, a filter on a Boolean algebra $\mathbb{B}$ is an ultrafilter if and only if $(\forall p \in \mathbb{B})(p \in U \vee \neg p \in U)$. For a detailed exposition of the various properties of filters and ultrafilters, see [9].

Let $U \subseteq \mathbb{B}$ be an ultrafilter on $\mathbb{B}$. The most direct method of obtaining a two-valued model from a Boolean-valued model is by constructing the smallest two-valued model $M[U]$ of ZFC containing M as a subset and U as an element. It is highly unlikely that U is a member of M, but ensuring that $M[U]$ satisfies these desired properties actually requires that one arrange for U to *almost* be a member of M.

Definition 1.5. A subset D of a bottomless poset $D \subseteq \mathbb{P}$ is called *dense* in $\mathbb{P}$ if for any $p \in \mathbb{P}$, there is a $q \in D$ such that $q \leqslant p$. A subset of a Boolean algebra $\mathbb{B}$ is called *dense* if it is dense in $\mathbb{B}\backslash\{\bot\}$. A subset $G \subseteq \mathbb{B}$ is called *M-generic* if $G \cap D \neq \varnothing$ for any $D \subseteq \mathbb{B}$ dense in $\mathbb{B}$.

An M-generic set is very much like an *average element* of M, and will rarely actually be an element of M. One particularly amusing analogy goes something like this: The average person, every of whose traits places them at the precise mean of the general populous, does not exist, and if one were to introduce such a person to the world, the world would be as it were. Of course, one does not simply add U to M to get $M[U]$: The functions in and out of $M[U]$, intersections and unions, images, and so on, also need to be present.

The most direct method of constructing $M[U]$ goes as follows: Define the *interpretation* $\iota_U(y)$ of an element $y \in M^{\mathbb{B}}$ recursively by

$$\iota_U(y) = \{\iota_U(x) \mid x \in \mathrm{dom}(y) \wedge y(x) \in U\}.$$

Again, the idea is that the value $y(x)$ is the "probability" that x is a member of y. So, U is simply picking out the elements of $\mathrm{dom}(y)$ whose probability of "being in y" is "high". Finally, define

$$M[U] = \{\iota_U(y) \mid y \in M^{\mathbb{B}}\}.$$

There is a copy of every element of M in $M[U]$, as well as a copy of U in $M[U]$. Indeed, let the *name* of a set $x \in M$ be the $\mathbb{B}$-valued set defined by

$$\mathrm{dom}(\hat{x}) = \{\hat{z} \mid z \in x\}$$

and $\hat{x}(\hat{z}) = \top$ for each $z \in x$. The *generic name*, the name of U, is given alternatively by

$$\mathrm{dom}(\hat{U}) = \{\hat{u} \mid u \in \mathbb{B}\},$$

and $\hat{U}(\hat{u}) = u$ for each $u \in \mathbb{B}$. Names are first-order definable, and are therefore actually elements of M (it is in this sense that M *almost* contains U). Inductively,

$$\iota_U(\hat{x}) = \{\iota_U(\hat{z}) \mid z \in x \wedge \top \in U\} = \{z \mid z \in x\} = x,$$

and

$$\iota_U(\hat{U}) = \{\iota_U(\hat{u}) \mid u \in U\} = U.$$

The term *forcing* enters the picture when the *forcing relation* is defined.

Definition 1.6. We define for each φ in $\mathcal{L}_M$ and $p \in \mathbb{B}$

$$p \Vdash \varphi \text{ iff } p \leqslant [\![\varphi]\!].$$

The relation $p \Vdash \varphi$ is read "*p forces* φ".

Notice that, if $M^{\mathbb{B}}$ is a Boolean-valued model of ZFC, then $p \Vdash \varphi$ for any $p \in \mathbb{B}$ and any axiom φ of ZFC. Cohen's observation was that, if U is M-generic, then $M[U]$ is a two-valued model of ZFC satisfying every of the forced properties.

Theorem 1.7. *Let M be a transitive model of ZFC, and $\mathbb{B}$ be a Boolean algebra in M. If U is an M-generic ultrafilter on $\mathbb{B}$, then $M[U]$ is a transitive model of ZFC. Moreover, for any $\varphi \in \mathcal{L}_M$,*

$$p \Vdash \varphi \textit{ for some } p \in U \textit{ implies } M[\mathcal{U}] \models \varphi.$$

Now, suppose the consistency of a certain sentence φ with ZFC were called into question. Since $\top \in U$ for any filter U, to show that a model of $ZFC + \varphi$ exists, it suffices to find a $\mathbb{B}$-valued model of ZFC in which $[\![\varphi]\!] = \top$ and that there exists some M-generic ultrafilter on $\mathbb{B}$. In the case at hand, such an ultrafilter always exists.

Lemma 1.8. *Let M be a countable transitive model of ZFC, and $\mathbb{B}$ an M-complete Boolean algebra in M such that $\mathbb{B}$ is* atomless*, ie. satisfies*

$$(\forall p)(p \neq \bot \textit{ implies } (\exists q)(q < p)).$$

Then there exists an M-generic ultrafilter U on $\mathbb{B}$.

Proof. If $\{D_n \mid n \in \mathbb{N}\}$ is an enumeration of the dense subsets of $\mathbb{B}$ in M (recall that each of these is dense in $\mathbb{B}\backslash\{\bot\}$), one can construct an ultrafilter as follows: Let $d_0 \in D_0\backslash\{\bot\}$, and suppose $d_1 \in D_1, \cdots, d_n \in D_n$ have been chosen such that

$$d_n < d_{n-1} < \cdots < d_1 < d_0.$$

Since D_{n+1} is dense in $\mathbb{B}$, there is some $d_{n+1} \in D_{n+1}$ such that $d_{n+1} \leqslant d_n$. This gives a strictly descending sequence $\{d_n \mid n \in \mathbb{N}\} \subseteq \mathbb{B}$. Define

$$U_0 = \{p \in \mathbb{B} \mid (\exists n)(d_n \leqslant p)\}.$$

This defines a filter on $\mathbb{B}$. An application of Zorn's lemma shows that every filter is contained in an ultrafilter, so in particular there is an ultrafilter U on $\mathbb{B}$ containing U_0. Since $d_n \in U$ for any $n \in \mathbb{N}$, U is M-generic. $\square$

The following theorem is a direct consequence of this lemma.

Theorem 1.9. *Let M be a countable transitive model of ZFC and $\mathbb{B}$ be an M-complete atomless Boolean algebra in M. Then $ZFC + \varphi$ is consistent if $[\![\varphi]\!] = \top$ in $M^{\mathbb{B}}$.*

1.4 Independence of CH

The simplest example of Boolean-valued models in action is the proof of the consistency of $ZFC + \neg CH$, where $\neg CH$ is the formula

$$(\exists A)(|\mathbb{N}| < |A| < |2^{\mathbb{N}}|).$$

In view of Theorem 1.9, the goal will be to construct a Boolean-valued model in which $[\![\neg CH]\!] = \top$.

Fix a countable transitive model M of ZFC, and let $\kappa > 2^{\mathbb{N}}$ in this model. Let $\mathbb{P}$ be the set of functions $f : A \to 2$ in M with A a finite subset of $\kappa \times \mathbb{N}$, and order the elements of $\mathbb{P}$ by reverse inclusion: For any $f, g \in \mathbb{P}$, set $f \leqslant g$ (read f *extends* g) if and only if $g \subseteq f$. Since $\mathbb{P}$ is first-order definable with parameters in M, $\mathbb{P} \in M$. Observe that $\mathbb{P}$ is bottomless. The partially ordered set $\mathbb{P}$ is called the *Cohen poset*.

To obtain a Boolean algebra from $\mathbb{P}$, say that $D \subseteq \mathbb{P}$ is *downwards-closed* when

$$(\forall p \in D)(\forall q \leqslant p)(q \in D).$$

Let

$$\mathbb{B} = \{D \subseteq \mathbb{P} \mid D \in M \wedge D \text{ is dense} \wedge D \text{ is downwards-closed}\},$$

and order $\mathbb{B}$ by inclusion. Again, $\mathbb{B}$ is first-order definable in M, so $\mathbb{B} \in M$.

Lemma 1.10. *The partially ordered set $\mathbb{B}$ is a Boolean algebra under the operations*

$$D \wedge D' = D \cap D',\ \ D \vee D' = \{q \in \mathbb{P} \mid (\exists p \in D \cup D')(q \leqslant p)\},$$

and

$$\neg D = \{q \in \mathbb{P} \mid (\exists p \in \mathbb{P} \backslash D)(q \varepsilon p)\}.$$

In fact, a more general statement is true: If X is any topological space, say that an open subset $U \subseteq X$ is *regular* if it is equal to the interior of its closure, written $\text{int}(\text{cl}(U)) = U$. Then the partially ordered set of regular open sets $RO(X)$ is a Boolean algebra under the operations

$$U \wedge V = U \cap V,\ U \vee V = \text{int}(\text{cl}(U \cup V)),$$

and

$$\neg U = \text{int}(X \backslash U).$$

Taking the open subsets of $\mathbb{P}$ to be the downwards-closed subsets of $\mathbb{P}$, one sees that the interior and closure operations are given by

$$\text{int}(A) = \{p \in A \mid (\forall q \leqslant p)(q \in A)\} \text{ and } \text{cl}(A) = \{p \in \mathbb{P} \mid (\exists q \leqslant p)(q \in A)\},$$

and therefore

$$\text{int}(\text{cl}(A)) = \{p \in \mathbb{P} \mid (\forall q \leqslant p)(\exists r \leqslant q)(r \in A)\}.$$

Verifying

$$\text{int}(\text{cl}(\downarrow(p))) = \downarrow(p)$$

is a brief exercise. In fact, the operation $p \mapsto \downarrow(p)$ monotonically maps $\mathbb{P}$ onto an isomorphic copy of itself in $\mathbb{B}$.

Lemma 1.11. *The Boolean algebra just defined, $\mathbb{B}$, is the unique Boolean algebra (up to isomorphism) containing the Cohen poset as a dense subset.*

The proof that $M^{\mathbb{B}}$ believes $\neg CH$ with probability $\top$ is sketched as follows. Define the $\mathbb{B}$-valued set $D : \hat{\kappa} \times \hat{\mathbb{N}} \to \mathbb{B}$ by

$$D(\hat{\alpha}, \hat{n}) = \{p \in \mathbb{P} \mid p(\alpha, n) = 0\}$$

for any $(\alpha, n) \in \kappa \times \mathbb{N}$. Since D is first-order definable (with parameters in M), D is actually a $\mathbb{B}$-valued set in M.

Now, D gives rise to a function $g : \hat{\kappa} \to \mathbb{B}^{\hat{\mathbb{N}}}$ in M, defined

$$g(\hat{\alpha})(\hat{n}) = D(\hat{\alpha}, \hat{n}),$$

which one can check is a function "with probability" $\top$. The function g is actually injective with probability $\top$ as well, since the domain of each $p \in \mathbb{P}$ is finite. This exhibits the name $\hat{g}$ of an injective function $\hat{\kappa} \to \mathbb{B}^{\hat{\mathbb{N}}}$ in $M^{\mathbb{B}}$. Weilding the existence of such a function, one can show that

$$[\![\neg CH]\!] \geqslant [\![\mathbb{N} < 2^{\hat{\mathbb{N}}} < \kappa \leqslant \mathbb{B}^{\hat{N}}]\!] = \top$$

in $M^{\mathbb{B}}$. In other words, the continuum hypothesis fails with probability $\top$ in $M^{\mathbb{B}}$, and by Theorem 1.9, $ZFC + \neg CH$ is consistent.

1.5 Where to go next

The construction and argument above is very nearly the construction and argument found in the proceeding sections. There is a very important difference between the two, however. What is to follow is a shift in perspective that not only places Boolean valued models in the context of category theory, but has also allowed for a unified theory of forcing.

The motto is "forcing is sheaves". Essentially, a category-theoretic rephrasing of the notions of Boolean-valued model, name, and interpretation reveals that many of the various elements of forcing are specific examples of concepts already appearing in geometry, topology, and algebra.

2 A Crash Course

To fully understand the topos theory in the following section requires a comfortability with the language of categories. The current section is intended to be a crash course in the necessary ideas, but is by no means a self-contained and complete exposition of category theory. For a more in-depth study than the current section, consult any of [1], [14], or the classic [11].

2.1 Categories, functors, natural transformations

Definition 2.1. A *category* is a pair of classes $\mathbb{C} = (\text{obj}(\mathbb{C}), \text{arr}(\mathbb{C}))$ equppied with two functions $\text{dom}, \text{cod} : \text{arr}(\mathbb{C}) \to \text{obj}(\mathbb{C})$ and an operation $(f, g) \mapsto g \circ f \in \text{arr}(\mathbb{C})$ defined for each pair $(f, g) \in \text{arr}(\mathbb{C}) \times \text{arr}(\mathbb{C})$ for which $\text{dom}(g) = \text{cod}(f)$, altogether satisfying the following two conditions.[1]

(C1) For every $A \in \text{obj}(A)$, there exists $\text{id}_A \in \text{arr}(\mathbb{C})$ satisfying $f \circ \text{id}_A = f$ and $\text{id}_A \circ g = g$ whenever either is defined.

(C2) For any $f, g, h \in \text{arr}(\mathbb{C})$ such that both $h \circ g$ and $g \circ f$ are defined, $(h \circ g) \circ f = h \circ (g \circ f)$.

Members of $\text{obj}(\mathbb{C})$ are called the *objects* of $\mathbb{C}$, and $\text{arr}(\mathbb{C})$ are called the *arrows.* A pair (f, g) for which $g \circ f$ is defined is called *composable.* The notation $f : A \to B$ is reserved to denote an arrow with $\text{dom}(f) = A$ and $\text{cod}(f) = B$, A is the *source* or *domain* of f, and B is the *target*, or *codomain.*

[1] One could translate these conditions into a first-order *theory of categories* and define a *category* to be a model of this theory. In other words, category theory is an elementary theory, just like ZFC.

For example, the category $\mathbb{S}\mathbf{ets}$ of sets has its class of objects $\mathrm{obj}(\mathbb{S}\mathbf{ets}) = \{x \mid x \text{ is a set}\}$ and arrows $\mathrm{arr}(\mathbb{S}\mathbf{ets}) = \{f \mid f \text{ is a function between sets}\}$. Similar examples include $\mathbb{G}\mathbf{roups}$ and $\mathbb{R}\mathbf{ings}$ and the category $\mathbb{T}\mathbf{op}$ of topological spaces and continuous maps, with arrows given by homomorphisms and continuous maps. A category in which every arrow is an identity arrow is called a *discrete* category.

For a different flavour of example, consider any preordered set $\mathbb{P}$. One forms a category from $\mathbb{P}$ by taking $\mathrm{obj}(\mathbb{P})$ to be the points in $\mathbb{P}$, $\mathrm{arr}(\mathbb{P}) = \leqslant$, and $\mathrm{dom}(p \leqslant q) = p$ and $\mathrm{cod}(p \leqslant q) = q$ (in other words, there is a unique arrow $p \to q$ between any two p, q such that $p \leqslant q$). This gives a category structure to partially ordered sets (posets), Boolean algebras, and the like. Conversely, a category with a unique arrow between any two objects is a poset.

For brevity's sake, $g \circ f$ will almost always be shortened to gf. Moreover, if h and g are arrows, and there exists an arrow f such that $h = gf$, one says that h *factors through* g (via f).

Definition 2.2. Fix an arrow $f : A \to B$. A *section* of f is any arrow $s : B \to A$ such that $fs = \mathrm{id}_B$ (id_B factors through f), and a *retract* of f is any arrow $r : B \to A$ such that $rf = \mathrm{id}_A$ (id_A factors through r). One calls f an *isomorphism* if there is a $g : B \to A$ that is both a section and a retract of f. In such a case, g is called the *inverse* of f.

To every category $\mathbb{C}$ there curresponds a dual category $\mathbb{C}^{op}$, called its *opposite*, defined by the equations $\mathrm{obj}(\mathbb{C}^{op}) = \mathrm{obj}(\mathbb{C})$ and

$$\mathrm{arr}(\mathbb{C}^{op}) = \{f^{op} : B \to A \mid (f : A \to B) \in \mathrm{arr}(\mathbb{C})\}.$$

Essentially, $\mathbb{C}^{op}$ is the category $\mathbb{C}$ with all of its arrows reversed.

Definition 2.3. A *functor* $F : \mathbb{C} \to \mathbb{D}$ is a homomorphism of categories, ie. a pair of functions $F : \mathrm{obj}(\mathbb{C}) \to \mathrm{obj}(\mathbb{D})$ and $F : \mathrm{arr}(\mathbb{C}) \to \mathrm{arr}(\mathbb{D})$ such that $F(g \circ f) = F(g) \circ F(f)$ for any composable pair (f, g) and $F(\mathrm{id}_A) = \mathrm{id}_{FA}$ for any object A.

A functor of the form $F : \mathbb{C}^{op} \to \mathbb{D}$ is called a *contravariant* functor and identified with its corresponding pair of functions $F : \mathrm{obj}(\mathbb{C}) \to \mathrm{obj}(\mathbb{D})$ and $F : \mathrm{arr}(\mathbb{C}) \to \mathrm{arr}(\mathbb{D})$. In this context, $F(g \circ f) = F(f) \circ F(g)$ is written instead.

One easy example of a functor $\mathbb{C} \to \mathbb{D}$ is a *constant functor*: Given an object D of $\mathbb{D}$, define $\Delta D : \mathbb{C} \to \mathbb{D}$ to be the functor

$$\Delta D(C) = D, \;\; \Delta D(f : C \to C') = \mathrm{id}_D\,.$$

Another easy example is the *identity* functor $\mathrm{id}_{\mathbb{C}} : \mathbb{C} \to \mathbb{C}$, which does what one might hope: Nothing!

For a different set of examples, call a category with a unique object a *group* if every of its arrows is an isomorphism, and recall that a poset is simply a category with a unique arrow between any two objects. A functor between groups is then just a group homomorphism, and a functor between posets is simply a monotone function.

At this point, one might like to form the *category of categories* $\mathbb{C}\mathbf{at}$, whose arrows are functors between categories. To avoid the usual paradoxes, $\mathbb{C}\mathbf{at}$ is taken to be the category of *small* categories instead, where a category $\mathbb{C}$ is called *small* if $\mathrm{arr}(\mathbb{C})$ is a set.

Definition 2.4. A *natural transformation* $\eta : F \to G$ between functors $F, G : \mathbb{C} \to \mathbb{D}$ is a family of arrows $\{\eta_A : FA \to GA \mid A \in \mathrm{obj}(\mathbb{C})\}$ such that the following diagram[2] commutes for every $f : A \to B$.

$$\begin{tikzcd}
FA \arrow[r, "\eta_A"] \arrow[d, "F(f)"'] & GA \arrow[d, "G(f)"] \\
FB \arrow[r, "\eta_B"] & GB
\end{tikzcd}$$

A *natural isomorphism* is a natral transformation that has an inverse.

If $\eta : F \to G$ and $\sigma : G \to H$ are natural transformations, the composition $(\sigma \circ \eta)$ is defined by the equation $(\sigma \circ \eta)_A = \sigma_A \circ \eta_A$, and also defines a natural transformation[3]. For any functor $F : \mathbb{C} \to \mathbb{D}$, there is an *identity* natural transformation $\mathrm{id}_F : F \to F$ defined by $(\mathrm{id}_F)_A = \mathrm{id}_{FA}$. This gives the structure of a category to the set

$$\mathbb{D}^{\mathbb{C}} = \{F \mid F : \mathbb{C} \to \mathbb{D} \text{ is a functor}\}$$

whose arrows are natural transformations.

2.2 Presheaves and the Yoneda embedding

A functor of the form $\mathbb{C}^{op} \to \mathbb{S}\mathbf{ets}$ is called a *presheaf*, and the category of presheaves $\mathbb{S}\mathbf{ets}^{\mathbb{C}^{op}}$ is denoted $\widehat{\mathbb{C}}$. Instructive examples of presheaves used in analysis and topology

[2] A *diagram* (in a category $\mathbb{C}$) is a graph whose vertices are objects and edges are arrows. A diagram is said to *commute* if for any two objects A and B in the graph, composing arrows along any two paths from A to B in the graph give the same arrow $A \to B$.

[3] Notice that it follows that a natural transformation $\{\eta_A\}$ is a natural isomorphism if and only if η_A is an isomorphism for any A.

include the continuous and bounded function algebra functors $C, B : \mathbf{O}(X) \to \mathbb{R}$, where $\mathbf{O}(X)$ is the poset of open subsets of a topological space X ordered by inclusion.

Set
$$\mathrm{Hom}_{\mathbb{C}}(A, B) = \{f : A \to B \mid f \in \mathrm{arr}(\mathbb{C})\},$$
and call $\mathbb{C}$ *locally small* if $\mathrm{Hom}(A, B)$ is a set for any two $A, B \in \mathbb{C}$. For a locally small category $\mathbb{C}$ and any $C \in \mathbb{C}$, the assignment

$$(\mathbf{y}C)A = \mathrm{Hom}(A, C)$$

determines a map $\mathbf{y}C : \mathrm{obj}(\mathbb{C}) \to \mathrm{obj}(\mathbb{S}\mathbf{ets})$. Furthermore, define

$$f^* : \mathrm{Hom}(B, C) \to \mathrm{Hom}(A, C)$$

by the equation $f^*(h) = h \circ f$ and set $(\mathbf{y}C)(f) = f^*$. The resulting presheaf $\mathbf{y}C$ is called the *Yoneda embedding* of C.

The assignment $C \to \mathbf{y}C$ so far defines a function $\mathbf{y} : \mathrm{obj}(\mathbb{C}) \to \mathrm{obj}(\widehat{\mathbb{C}})$. To give $\mathbf{y}$ functor structure, set $(f_*)_A(h) = f \circ h$ for any $f : C \to C'$ and $h : A \to C$. The resulting family of functions

$$(f_*)_A : \mathrm{Hom}(A, C) \to \mathrm{Hom}(A, C')$$

determines a natural transformation $f_* : \mathbf{y}C \to \mathbf{y}C'$. Setting $\mathbf{y}(f) = f_*$ gives a functor $\mathbf{y} : \mathbb{C} \to \widehat{\mathbb{C}}$, called the *Yoneda embedding* of $\mathbb{C}$.

Lemma 2.5. **(Yoneda)** *Let $\mathbb{C}$ be a locally small category and $C \in \mathrm{obj}(\mathbb{C})$. For any presheaf $P : \mathbb{C}^{op} \to \mathbb{S}\mathbf{ets}$, the function*

$$\Phi_C : \mathrm{Hom}(\mathbf{y}C, P) \to PC$$

defined by the equation
$$\Phi_C(\eta) = \eta_C(\mathrm{id}_C)$$
is a bijection. Moreover, the family $\{\Phi_C\}$ defines a natural isomorphism $\Phi : \mathrm{Hom}(\mathbf{y}(-), P) \to P$.

Observe that, if $P = \mathbf{y}C'$ for some C', this gives a natural bijection

$$\mathrm{Hom}(\mathbf{y}C, \mathbf{y}C') \cong \mathrm{Hom}(C, C')$$

whose inverse is the Yoneda embedding. A useful consequence of this is that every natural transformation $\mathrm{Hom}(-, C) \to \mathrm{Hom}(-, C')$ is of the form f_* for some $f : C \to$

C'. Generally, one says that a functor $F : \mathbb{C} \to \mathbb{D}$ is *full* and *faithful* if its action on arrows

$$F : \mathrm{Hom}(A, B) \to \mathrm{Hom}(FA, FB)$$

is surjective and injective respectively.

From here on out, categories will generally be assumed to be locally small. This is not a huge drawback, notice, since any category of the form $\widehat{\mathbb{C}}$, when $\mathbb{C}$ is small, is locally small.

2.3 Equivalences and adjoints

Interestingly, the notion of isomorphism for categories is nearly useless to ordinary mathematicians. Consider the Gelfand-Naimark theorem or Stone's representation theorem, or any one of the other "equivalence results" from the 20th century: These equivalences are concerned with "invertible constructions", not bijective correspondances. This indicates that the notion of *isomorphism as equivalence* is too strong. The following notion, made possible by the laguage of category theory, specializes to each of the mentioned famous equivalences.

Definition 2.6. A functor $F : \mathbb{C} \to \mathbb{D}$ is an *equivalence of categories* if there is another functor $G : \mathbb{D} \to \mathbb{C}$ and natural isomorphisms $\eta : \mathrm{id}_{\mathbb{C}} \to GF$ and $\varepsilon : FG \to \mathrm{id}_{\mathbb{D}}$. If there is an equivalence of categories between $\mathbb{C}$ and $\mathbb{D}$, then $\mathbb{C}$ and $\mathbb{D}$ are called *eqivalent*, written $\mathbb{C} \simeq \mathbb{D}$.

This turns out to be the right notion of equivalence for categories: Essentially, if φ is an elementary property that a category may or may not have (here, $=$ is replaced by $\cong$ in the language) and $\mathbb{C} \simeq \mathbb{D}$, then $\mathbb{C}$ has the property φ if and only if $\mathbb{D}$ does.

There is an equivalence of categories that suggests some role in the analogy between $M^{\mathbb{B}}$ in the previous section and the categories of sets in the sequel. Given a category $\mathbb{C}$ and an object A in $\mathbb{C}$, we form the *slice* category $\mathbb{C}/A$ whose objects are given by

$$\mathrm{obj}(\mathbb{C}) = \{b \in \mathrm{arr}(\mathbb{C}) \mid \mathrm{cod}(b) = A\}$$

and every of whose arrows $h : b_1 \to b_2$ is given by an arrow $h : B_1 \to B_2$ such that

$$\begin{tikzcd}
B_1 \arrow[r, "h"] \arrow[dr, "b_1"'] & B_2 \arrow[d, "b_2"] \\
& A
\end{tikzcd}$$

commutes.

Theorem 2.7. *Let A be any set, thought of as a discrete category. There is an equivalence of categories* $\mathbb{S}\mathbf{ets}^A \simeq \mathbb{S}\mathbf{ets}/A$.

The idea behind the proof is simply that an object F of $\mathbb{S}\mathbf{ets}^A$ is a family of sets indexed by the elements of A, while an object in $\mathbb{S}\mathbf{ets}/A$ is a family of sets over A. The equivalence is given by the functor

$$T : \mathbb{S}\mathbf{ets}^A \longrightarrow \mathbb{S}\mathbf{ets}/A,$$

for which

$$T(F) : \bigsqcup_{a \in A} F(a) \to A$$

is defined by the equation

$$T(F)(x \in F(a)) = a.$$

The reverse construction is likely not hard to imagine.

In light of the theorem above, one might expect $\mathbb{S}\mathbf{ets}^{\mathbb{B}} \simeq \mathbb{S}\mathbf{ets}/\mathbb{B}$. Of course, $\mathbb{B}$ is not a discrete category (it is a poset), so $\mathbb{S}\mathbf{ets}^{\mathbb{B}}$ is far more restricted than $\mathbb{S}\mathbf{ets}/\mathbb{B}$. Dually, $\mathbb{S}\mathbf{ets}/\mathbb{B}$ is far too complicated to be a Boolean valued model: A Boolean valued set is not just any function from a set to $\mathbb{B}$! Instead, the category of Boolean valued sets is a strict *subcategory* of $\mathbb{S}\mathbf{ets}/\mathbb{B}$: Its objects are functions such as $x : X \to \mathbb{B}$ and $y : Y \to \mathbb{B}$ where X and Y contain only Boolean valued sets, and its arrows are subsets of the form $f \subseteq X \to Y$ such that

$$[\![(\forall r \in X)(\forall s, t)((r, s), (r, t) \in f \Rightarrow s = t)]\!] = \top,$$

ie. its arrows are relations that are "functions with probability $\top$". Transporting this subcategory of $\mathbb{S}\mathbf{ets}/\mathbb{B}$ along the equivalence in Theorem 2.7 produces a certain subcategory of $\mathbb{S}\mathbf{ets}^{\mathbb{B}}$ to be defined and explored later.

A mild weakening of the concept of equivalence is that of having a *left* or *right adjoint*, which encompasses "almost inverse" constructions such as the free-group construction, the hom-tensor adjunction, and the sheafification construction for spaces. Note that the definition below marks the first instance of a *universal mapping property* in this document, a formula of the form $(\forall f)(\exists! g)(D(f, g, \bar{a}))$ where the variables denote arrows and $D(f, g, \bar{a})$ denotes a diagram in which that arrows f, g and every a_i in $\bar{a}$ appear.

Definition 2.8. Let $F : \mathbb{C} \rightleftarrows \mathbb{D} : G$ be a pair of functors. One says that F is *left adjoint* to G (G is *right adjoint to* F), written $F \dashv G$, if there is a natural transformation $\eta : \mathrm{id}_{\mathbb{C}} \to GF$ satisfying the following universal mapping property:

For any $f : A \to GB$ in $\mathbb{C}$, $B \in \text{obj}(\mathbb{D})$, there is a unique $g : FA \to B$ such that the following diagram commutes.

$$\begin{tikzcd} A \arrow[r, "\eta_A"] \arrow[dr, "f"'] & GFA \arrow[d, "G(g)"] \\ & GB \end{tikzcd}$$

The natural transformation η is called the *unit* of the adjuntion.

Equivalently, for $\mathbb{C}$ and $\mathbb{D}$ locally small categories, $F \dashv G$ if and only if there is a bijection

$$\Phi_{A,B} : \text{Hom}_{\mathbb{C}}(A, GB) \cong \text{Hom}_{\mathbb{D}}(FA, B)$$

that is *natural* in A and in B. Here, *natural in* A means that the diagram

$$\begin{tikzcd} \text{Hom}(A, GB) \arrow[r, "\Phi_{A,B}"] \arrow[d, "f^*"] & \text{Hom}(FA, B) \arrow[d, "F(f)_*"] \\ \text{Hom}(A', GB) \arrow[r, "\Phi_{A',B}"] & \text{Hom}(FA', B) \end{tikzcd}$$

commutes for every $f : A' \to A$. What it means for $\Phi_{A,B}$ to be *natural in* B is similar. Any adjunction $F \dashv G$ is also given uniquely by its *counit* $\varepsilon : FG \to \text{id}_{\mathbb{D}}$, defined by the universal mapping property dual to that of the unit η: For any $g : FA \to B$ in $\mathbb{D}$, there is a unique $f : A \to GB$ such that the folowing diagram commutes.

$$\begin{tikzcd} FGB \arrow[r, "\varepsilon_A"] & B \\ FA \arrow[u, "F(f)"] \arrow[ur, "g"'] & \end{tikzcd}$$

While η basically supplies the map $\Phi_{A,B}$ defined above, ε supplies its inverse.

Lemma 2.9. *Adjoints are unique up to isomorphism. In other words, if $F \dashv G$ and $F \dashv G'$, then there is a natural isomorphism $G \cong G'$.*

There are many examples of adjoint functors in ordinary mathematics. For example, the *forgetful functor* U : $\mathbb{G}$**roups** $\to$ $\mathbb{S}$**ets**, which takes a group to its underlying set, has the free-group functor as a right adjoint: Rewording this in terms of the isomorphism Φ, all this says is that a homomorphism of a fixed free group is determined by its effect on generators.

One example from pure category theory is the *constant/section adjunction*. Recall that for any set A one can form the constant functor $\Delta A : \mathbb{C}^{op} \to \mathbb{S}\mathbf{ets}$. In fact, Δ is itself a functor, with

$$\Delta(f : A \to B) : \Delta A \to \Delta B$$

given by the constant family, $f_C = f$ for any C in $\mathbb{C}$. Explicitly, one obtains a functor of the form $\Delta : \mathbb{S}\mathbf{ets} \to \widehat{\mathbb{C}}$. Going in the other direction, define $\Gamma : \widehat{\mathbb{C}} \to \mathbb{S}\mathbf{ets}$ as follows: For any functor $F : \mathbb{C}^{op} \to \mathbb{S}\mathbf{ets}$,

$$\Gamma F = \mathrm{Hom}(\Delta 1, F),$$

and $\Gamma(\sigma : F \to G) = \sigma^*$ as usual. The calculation

$$\begin{aligned} \mathrm{Hom}(\Delta A, F) &\cong \mathrm{Hom}(\coprod_{a \in A} \Delta 1, F) \\ &\cong \prod_{a \in A} \mathrm{Hom}(\Delta 1, F) \\ &\cong \prod_{a \in A} \Gamma F \qquad \qquad \cong \mathrm{Hom}(A, \Gamma F) \end{aligned}$$

takes place in the category of sets, and reveals that $\Delta \dashv \Gamma$. In fact, one even has

$$\Gamma \Delta A = \mathrm{Hom}(\Delta 1, \Delta A) \cong \mathrm{Hom}_{\mathbb{S}\mathbf{ets}}(1, A) \cong A,$$

which is not uncommon in this sort of situation.

For a different sort of example, consider a Boolean algebra $\mathbb{B}$. Fix a $p \in \mathbb{B}$ and consider the map $p \wedge (-) : \mathbb{B} \to \mathbb{B}$ defined by $q \mapsto p \wedge q$. A routine calculation reveals that $p \wedge (-)$ is monotone, or in other words a functor on $\mathbb{B}$. Recall that the implication operator $\Rightarrow$, defined by the equation $p \Rightarrow q = \neg p \vee q$, is characterized by the property

$$(p \wedge q \leqslant r) \text{ if and only if } (q \leqslant p \Rightarrow r)$$

in $\mathbb{B}$. A second calculation reveals that $p \Rightarrow (-)$ is also a functor, so the characteristic formula gives $p \wedge (-) \dashv p \Rightarrow (-)$.

Definition 2.10. A lattice $\mathbb{H}$ with a top and bottom is called a *Heyting algebra* if $p \wedge (-) : \mathbb{H} \to \mathbb{H}$ has a right adjoint for any $p \in \mathbb{H}$.

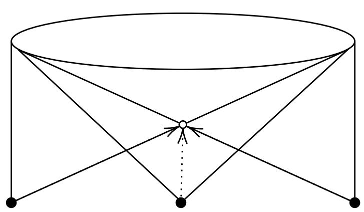
Where Boolean algebras were invented to study classical logic algebraically, Heyting algebras were formulated to study the logic of intuitionism adopted by Brouwer, Heyting, and even Poincaré. Intuitionistic logic differs from classical logic in that it is the logic of strictly *positive proofs*, rejecting any use of the law of excluded middle.

This manifests in the only difference between Heyting and Boolean algebras: If the identity $\neg\neg\varphi = \varphi$ holds for any φ in a Heyting algebra H, then H is a Boolean algebra.

Possibly the most useful property of right and left adjoints is that they can be used to transfer certain constructions, called *limits* and *colimits*, from one category to the next.

2.4 Limits

Definition 2.11. Let $P : \mathbb{C} \to \mathbb{D}$ be any functor. A *cone* for P is an object $D \in \mathbb{D}$ and a family of maps $\{d_A : D \to PA \mid A \in \mathrm{obj}(\mathbb{C})\}$ indexed by the objects of $\mathbb{C}$ such that for any arrow $f : A \to B$ in $\mathbb{C}$, $P(f) \circ d_A = d_B$. A cone $(X, \{\eta_A\})$ for P is *limiting*, or a *limiting cone*, if it satisfies the following universal mapping property: For any other cone $(D, \{d_A\})$, there is a unique *morphism of cones* $\sigma : D \to X$ in $\mathbb{D}$ such that $\eta_A \circ \sigma = d_A$.



The category $\mathbb{D}$ is said to be *complete* if every functor into $\mathbb{D}$ from a small category has a limiting cone. Similarly, $\mathbb{D}$ is said to have *finite limits* if every functor $P : \mathbb{C} \to \mathbb{D}$ with $\mathbb{C}$ a finite category (ie. $\mathrm{arr}(\mathbb{C})$ is finite) has a limiting cone.

Limiting cones are unique up to isomorphism, so one often writes $X = \varprojlim P$ and says that X is the *limit of* P if (X, η) is a limiting cone for P.

Functors from finite categories are often thought of as diagrams (and in fact, these two ideas are equivalent). For instance, the diagram

$$A \longrightarrow Y \longleftarrow B$$

corresponds to a functor from a category of the same shape,

$$\bullet \longrightarrow \bullet \longleftarrow \bullet$$

taking the first "$\bullet$" to A, the left "$\bullet \longrightarrow \bullet$" to $A \longrightarrow B$, and so on. The limit of a diagram of this shape is called a *pullback*, and is an important construction found throughout mathematics. Other prominent examples include the limit of a diagram of the form

$$\begin{tikzcd} \bullet \arrow[r, shift left] \arrow[r, shift right] & \bullet \end{tikzcd},$$

called an *equalizer*; the limit of an empty diagram, the *terminal object*; and the limit of a diagram of the form

$$\bullet \qquad\qquad\qquad \bullet$$

called a *product*. If A and B are objects, then the product of A and B is denoted $A \times B$, and the arrows $A \times B \to A$ and $A \times B \to B$ are called *projections*, often denoted π_A and π_B respectively.

Since the pullback construction is especially important to understand, some concrete examples are in order. Consider a poset $\mathbb{P}$, as well as $p, q, r \in \mathbb{P}$ such that $p \leqslant r$ and $q \leqslant r$. Then $p \to r \leftarrow q$ is a diagram in $\mathbb{P}$. The pullback of this diagram is an element $z \in \mathbb{P}$ such that for any cone $x \leqslant p, q, r$ one also has $x \leqslant z$, as in the following diagram.

$$\begin{tikzcd}
& p \arrow[r] & r \\
& z \arrow[u] \arrow[r] & q \arrow[u] \\
x \arrow[uur] \arrow[ur, dashed] \arrow[urr] & &
\end{tikzcd}$$

This makes $z = p \wedge q$. In other words, in a poset, pullbacks are products. In fact, in a category with a terminal object 1 and objects A and B, $A \times B$ is the pullback of the diagram $A \to 1 \leftarrow B$. For an example of a different flavour, let A, B, and C be sets and $f : A \to C$, $g : B \to C$. Define the set

$$D = \{(x, y) \mid f(x) = g(y)\} \subseteq A \times B,$$

and let $\pi_A : D \to A$ and $\pi_B D \to B$ be the projections onto A and B respectively. It follows right from the definition that $(D, \{\pi_A, \pi_B\})$ is a limiting cone for $A \xrightarrow{f} C \xleftarrow{g} B$.

One might wonder if, since products can be formed in a category with pullbacks and a terminal object, what other kinds of limits can be formed in such a category? This question has an illuminating answer, perhaps more illuminating than its proof.

<u>Lemma 2.12.</u> *Let $\mathbb{C}$ be any category. The following are equivalent:*

*(**a**) $\mathbb{C}$ has finite limits.*

*(**b**) $\mathbb{C}$ has pullbacks and a terminal object.*

(c) $\mathbb{C}$ *has equalizers and finite products.*

Dual to the concept of a limit is that of a *colimit*, which is simply a limit in the opposite category. Following the idea that colimits are simply limits "with the arrows reversed", the notation $\varinjlim F$ is used to denote the colimit of the functor F. Going from a category to its opposite, a pullback becomes a *pushforward*, a product becomes a *sum* (or *coproduct*), an equalizer becomes a *coequalizer*, and the terminal object becomes *initial*. Examples of sums include the disjoint union operation in $\mathbb{S}\mathbf{ets}$ and the $\vee$ operation in a lattice.

Untangling the definition, the sum operation can be seen as a direct generalization of the disjoint union operation in $\mathbb{S}\mathbf{ets}$. The coproduct of A and B is denoted $A \sqcup B$, and the maps $A \to A \sqcup B$ and $B \to A \sqcup B$ are called *inclusions*.

Definition 2.13. Let $F : \mathbb{C} \to \mathbb{D}$ and $P : \mathbb{I} \to \mathbb{C}$ be functors, and (A, d) be a cone for P. Then F is said to *preserve limits for* P if $(FA, F(d))$ is a limiting cone for FP when (A, d) is a limiting cone for P. Generally, F is said to be *continuous* if F preserves limits for any $P : \mathbb{I} \to \mathbb{C}$.

The analogous definition for colimits is evident, and a functor that preserves colimits is caled *cocontinuous*.

There are several continuity/cocontinuity results that will show up in the material to come. The most important one is that $\mathbf{y}$ is continuous, ie.

$$\mathrm{Hom}(C, \varprojlim P) \cong \varprojlim \mathrm{Hom}(C, P(-))$$

for any functor $P : \mathbb{I} \to \mathbb{C}$ and $C \in \mathbb{C}$, where the latter limit is taken in $\mathbb{S}\mathbf{ets}^{\mathbb{I}}$. Furthermore, the functor $\mathbf{y}C = \mathrm{Hom}(-, C)$ is *co*continuous. Since $\mathbf{y}C$ is contravariant, this takes the equational form

$$\mathrm{Hom}(\varinjlim P, C) \cong \varprojlim \mathrm{Hom}(P(-), C).$$

In particular, one has the familliar equations

$$\mathrm{Hom}(A, B \times C) \cong \mathrm{Hom}(A, B) \times \mathrm{Hom}(A, C)$$
$$\mathrm{Hom}(A \sqcup B, C) \cong \mathrm{Hom}(A, C) \times \mathrm{Hom}(B, C)$$

Where $\mathbb{C} = \mathbb{S}\mathbf{ets}$ above, one can actually see $\varprojlim$ and $\varinjlim$ as functors $\mathbb{S}\mathbf{ets}^{\mathbb{I}} \to \mathbb{S}\mathbf{ets}$, taking a functor to its limit and colimit respectively (their effect on arrows is computed using cones). In this setting, it is actually the case that $\varprojlim \dashv \Delta \dashv \varinjlim$. See [1] for a nice exposition of this fact.

The second continuity result that will appear later is known as *RAPL* (*right adjoints preserve limits*). Its applications extend far beyond those found in the proceeding sections.

Theorem 2.14. *Let $F \dashv G$ be a pair of adjoint functors. Then G preserves limits, and F preserves colimits.*

The third continuity result worth mentioning is actually more of a completeness result.

Theorem 2.15. *The category $\widehat{\mathbb{C}}$ is the* cocompletion *of the category $\mathbb{C}$, in the sense that every presheaf F can be written in the form*

$$F \cong \varinjlim(\mathbf{y} \circ P)$$

for some functor $P : \mathbb{I} \to \mathbb{C}$, and that every functor $\mathbb{C} \to \mathbb{D}$ extends to a unique functor $\widetilde{\mathbb{C}} \to \mathbb{D}$.

Suppose $F \cong \varinjlim(\mathbf{y} \circ P)$ and G is a presheaf on $\mathbb{C}$. Let $\iota_i : \mathbf{y}Pi \to F$ denote an inclusion, so that $(F, \{\iota_i\})$ is a limiting cone for $\mathbf{y} \circ P$. Then, since $(G, \{\eta\iota_i\})$ is a cone for G, the universal mapping property for limiting cones implies that η is determined uniquely by the maps $\eta \circ \iota_i$. In particular, if $\eta \neq \mu$ for some $\mu : F \to G$, then for some $i \in \mathrm{obj}(\mathbb{I})$, $\eta \circ \iota_i \neq \mu \circ \iota_i$. In such a case, one says that $\widehat{\mathbb{C}}$ is *generated* by the class of *representable* functors $\{\mathbf{y}C \mid C \in \mathrm{obj}(\mathbb{C})\}$.

2.5 The natural numbers

The statement of $\neg CH$ implies some a priori conception of the natural numbers. To make sense of the statement of $\neg CH$ for categories, the natural numbers need to be given a categorical characterization. To this end, it helps to consider the sort of process that $\mathbb{N}$ is used to describe: Indexing denumerable lists.

Definition 2.16. In a category with a terminal object 1, an *inductive system* (X, t, z) is any diagram of the form

$$1 \xrightarrow{\quad z \quad} X \xrightarrow{\quad t \quad} X\ .$$

The inductive systems form a category, an arrow in which is an arrow $X \to Y$ such that the following diagram commutes.

$$\begin{tikzcd}
1 \arrow[r, "z"] \arrow[dr, "z'"] & X \arrow[r, "t"] \arrow[d] & X \arrow[d] \\
& Y \arrow[r, "t'"] & Y
\end{tikzcd}$$

A *natural numbers system* (N, s, z) is an initial object in the category of inductive systems, and N is called a *natural numbers object.*

That

$$1 \xrightarrow{\varnothing} \mathbb{N} \xrightarrow{(-)+1} \mathbb{N}$$

is the natural numbers system for $\mathbb{S}\mathbf{ets}$ is an easy exercise. Recall that $\mathbb{N}$ is a set because the *axiom of infinity* deems it to be so. Indeed, one might consider the mere existence of a natural numbers object to define an axiom of infinity for categories.

This *internal* form of the axiom of infinity is not so rarely satified. Actually, many categories have natural numbers objects, as the following lemma indicates.

Lemma 2.17. *Let $F : \mathbb{C} \rightleftarrows \mathbb{D} : G$ be a pair of adjoint functors, $F \dashv G$, such that F preserves finite limits. If (N, s, z) is a natural numbers system for $\mathbb{C}$, then $(FN, F(s), F(z))$ is a natural numbers system for $\mathbb{D}$.*

Recall the adjunction $\Delta \dashv \Gamma$ of the constant and section functors $\Delta : \mathbb{S}\mathbf{ets} \leftrightarrows \mathbb{S}\mathbf{ets}^{\mathbb{C}} : \Gamma$ from earlier, and observe that Δ is a continuous functor. Thus, the lemma above implies that any category of the form $\widehat{\mathbb{C}}$ has the constant functor $\Delta\mathbb{N}$ as a natural numbers object.

3 Logic and Toposes

With a little category theory under one's belt, it is possible to begin answering questions such as what it means for a category to give a model of ZFC. One such answer would involve simply translating the axioms of ZFC into category theoretic language. However, as will soon become clear, category theoretic language allows for an entirely different depiction of mathematical foundations. The incredibly simple and flexible notion of a "category of sets", called a *topos*, will take center stage. Toposes interpret higher-order intuitionistic logic, but are also the home of *sheaves* (to be defined later). As will be discussed near the end of the next section, sheaves are the carriers of models of ZFC.

3.1 The Subobject Functor

Recall the model M, the Boolean algebra $\mathbb{B} \in M$, and the construction of $M^{\mathbb{B}}$ in the previous section. The structure $M^{\mathbb{B}}$ earns its title as a *Boolean*-valued model because of its Boolean algebra of *truth values*, or *values of formulas*. But, what exactly *are* truth values? Such an idea is captured beautifully by the laguage of category theory, inspired by the usual identification of formulas with the subsets they define.

The internal logic of M consists of the lattice of elements of $2 = \mathcal{P}(\mathcal{P}(0))$, the subset classifier. As an object in $\mathbb{S}\mathbf{ets}$, 2 is characterized by its universal mapping property. In order to make sense of "characteristic arrows", one needs a notion of "subobject".

Definition 3.1. Let $\mathbb{C}$ be a category. An arrow $m : B \to A$ in $\mathbb{C}$ is called *monic* if for any object C in $\mathbb{C}$ and any arrows $i, j : C \to B$, $mi = mj$ implies $i = j$ (equivalently, m_* is injective). Say that two monic arrows $m_1 : B_1 \to A$ and $m_2 : B_2 \to A$ are equivalent if there is an isomorphism $h : B_1 \to B_2$ such that $m_2 h = m_1$.

Dually, an arrow in $\mathbb{C}$ is called *epic* if it is monic in the *opposite category* $\mathbb{C}^{op}$. Unpacking that statement, an arrow $k : A \to B$ is epic if and only if $ik = jk$ implies $i = j$ for any arrows i and j out of B (equivalently, k^* is injective).

Write $\mathrm{Sub}(A)$ to denote the class of equivalence classes of monic arrows into A. An element of $\mathrm{Sub}(A)$ is called a *subobject* of A. In $\mathbb{S}\mathbf{ets}$, a function is monic if and only if it is injective, and epic if and only if it is surjective. An equivalence class of monic arrows into a fixed set A is identifiable with its image, a subset of A. Thus, if $B \to A$ is a monic arrow, we refer to its corresponding subobject with $B \subseteq A$.

Lemma 3.2. *Let $B, C \subseteq A$. Define a relation $\leqslant_A$ on $\mathrm{Sub}(A)$ by setting $B \leqslant_A C$ if and only if there is an arrow $B \to C$ such that the following diagram commutes*

$$\begin{tikzcd} & A \\ B \arrow[ur] \arrow[r] & C \arrow[u] \end{tikzcd}$$

Then $\leqslant_A$ is a poset relation on $\mathrm{Sub}(A)$.

Using $\mathbb{S}\mathbf{ets}$ as an example again, $\mathrm{Sub}(A)$ is identifiable with the Boolean algebra 2^A. It is in this sense that the set $\mathrm{Sub}(A)$ is given *internally* by the exponent 2^A of the subset classifier 2 in $\mathbb{S}\mathbf{ets}$. Translating the universal mapping property of 2, one obtains the notion of a *subobject classifier*.

Definition 3.3. In a category $\mathbb{C}$, a *subobject classifier* is an object Ω in $\mathbb{C}$ equipped with an arrow $\mathrm{true} : I \to \Omega$ such that for any $U \subseteq A$, there is a unique arrow $\mathrm{ch}(U) : A \to \Omega$ such that the following is a pullback diagram.

$$\begin{tikzcd}
A \arrow[r, "\mathrm{ch}(U)"] & \Omega \\
U \arrow[u] \arrow[r, "!"] & I \arrow[u, "\mathrm{true}"']
\end{tikzcd}$$

Actually, if Ω is a subobject classifier equipped with the arrow $\mathrm{true} : I \to \Omega$, then I is necessarily a terminal object for the category. Indeed, id_A is always monic, so $|\operatorname{Hom}(A, I)| \geqslant 1$ for any A. To see the reverse, observe that

$$\begin{tikzcd}
I \arrow[r, "\mathrm{true}"] & \Omega \\
I \arrow[u, "\mathrm{id}_I"] \arrow[r, "\mathrm{id}_I"] & I \arrow[u, "\mathrm{true}"]
\end{tikzcd}$$

is a pullback diagram. This makes true monic, so $|\operatorname{Hom}(A, I)| \leqslant 1$.

In set theory, the notation 2^A is used to denote the set of functions $\operatorname{Hom}(A, 2)$ from A to 2. Now, given the obvious equivalence between the powerset $\mathcal{P}(A)$ and $\mathrm{Sub}(A)$, it is traditional in set theory to identify $\mathrm{Sub}(A)$ with 2^A. In fact, define $\mathrm{Sub}(f : A \to B) : \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ to be the function

$$\mathrm{Sub}(f)(V) = f^{-1}(V) = \{a \in A \mid f(a) \in V\}$$

for any $V \subseteq B$. This gives $\mathrm{Sub}(-)$ functor structure, with respect to which the equivalence $\mathrm{Sub}(-) \cong \operatorname{Hom}(-, 2)$ is a natural isomorphism. It is this natural equivalence that is referred to when one says that 2 *represents* $\mathrm{Sub}(-)$ in $\mathbb{S}\mathbf{ets}$. More generally, in a category with finite limits and a subobject classifier, there is a canonical functor structure on $\mathrm{Sub}(-)$ with respect to which a natural isomorphism $\mathrm{Sub}(-) \cong \operatorname{Hom}(-, \Omega)$ exists.

The need for finite limits can be motivated by thinking of a category as a sort of *universe of discourse*. Essentially, one would like the object Ω to represent the *internal logic* of its ambient category, in the same way that the Boolean algebra 2 represents the internal logic of $\mathbb{S}\mathbf{ets}$ (ie. formulas can be identified with their corresponding functions into 2). The internal logic of the category $\mathbb{S}\mathbf{ets}$ is classical logic[4], while the internal

[4]Actually, higher-order cassical logic.

logic of (the category) $M^{\mathbb{B}}$ is $\mathbb{B}$-valued classical logic. The precise meaning of the phrase "internal logic" is somewhat complicated, but the basic idea is that a category with a subobject classifier is able to interpret some fragment of intuitionistic logic, and the richness of the fragment depends on the richness of the ambient category.

Fix a category $\mathbb{C}$ with finite limits. To give $\mathrm{Sub}(-)$ the structure of a functor $\mathbb{C} \to \mathbb{S}\mathbf{ets}$, define for each $f : A \to B$ in $\mathbb{C}$ and $U \subseteq B$ the obect $f^{-1}U$ to be the pullback of the diagram $A \xrightarrow{f} B \xleftarrow{\subseteq} U$, as in

$$\begin{tikzcd} A \arrow[r, "f"] & B \\ f^{-1}U \arrow[u] \arrow[r] & U \arrow[u] \end{tikzcd}$$

It follows from the next lemma that the arrow $f^{-1}U \to A$ above represents a subobject of A in $\mathbb{C}$.

<u>**Lemma 3.4.**</u> *Suppose the following diagram is a pullback diagram*

$$\begin{tikzcd} A \arrow[r] & B \\ P \arrow[u] \arrow[r] & C \arrow[u] \end{tikzcd}$$

Then $P \to A$ is monic if $C \to B$ is.

Therefore, $f^{-1} : \mathrm{Sub}(B) \to \mathrm{Sub}(A)$. The next lemma implies that f^{-1} is actually monotone (ie. a functor).

<u>**Lemma 3.5.**</u> *Consider the following commutative diagram*

$$\begin{tikzcd} A \arrow[r] & B \arrow[r] & C \\ P \arrow[u] \arrow[r] & Q \arrow[u] \arrow[r] & D \arrow[u] \end{tikzcd}$$

in which the right-hand square is a pullback diagram. Then the outer rectangle is a pullback if and only if the left-hand square is a pullback.

The same lemma can also be used to show that $\mathrm{Sub}(-)$ is a contravariant functor, in the sense that $(fg)^{-1} = g^{-1}f^{-1}$. The following lemma states that the isomorphism $\mathrm{Sub}(A) \cong \mathrm{Hom}(A, \Omega)$ is natural in A.

Lemma 3.6. *In a category $\mathbb{C}$ with finite limits and a subobject classifier, if $f : A \to B$ is any arrow in $\mathbb{C}$, then the following diagram commutes.*

$$\begin{tikzcd}
\mathrm{Sub}(A) \arrow[r, "\mathrm{ch}(-)"] & \mathrm{Hom}(A,\Omega) \\
\mathrm{Sub}(B) \arrow[u, "f^{-1}"] \arrow[r, "\mathrm{ch}(-)"] & \mathrm{Hom}(B,\Omega) \arrow[u, "f^*"']
\end{tikzcd}$$

Proof. The two smaller rectangles below are pullbacks by definition.

$$\begin{tikzcd}
A \arrow[r, "f"] & B \arrow[r, "\mathrm{ch}(V)"] & \Omega \\
f^{-1}V \arrow[u] \arrow[r] & V \arrow[u] \arrow[r, "!"] & 1 \arrow[u, "\mathrm{true}"']
\end{tikzcd}$$

By the previous lemma, the outer rectangle is a pullback. By uniqueness of characteristic functions, $\mathrm{ch}(f^{-1}V) = \mathrm{ch}(V)f = f^*(\mathrm{ch}(V))$. □

3.2 Exponentials and Powerobjects

One could also see $\mathrm{Sub}(-)$ as a *covariant* functor, ie. a functor $\mathbb{C} \to \mathbb{S}\mathbf{ets}$. This requires the slightest additional structure on $\mathbb{C}$.

Definition 3.7. In a category $\mathbb{C}$, the *exponential* B^A of B by A is an object equipped with an arrow $\varepsilon : A \times B^A \to A$ satisfying the following universal mapping property in $\mathbb{C}$: For any $h : A \times X \to A$, there is a unique $\mathrm{tr}(h) : X \to B^A$ such that the following diagram commutes.

$$\begin{tikzcd}
A \times B^A \arrow[r, "\varepsilon"] & A \\
A \times X \arrow[u, "\mathrm{id}_A \times \mathrm{tr}(h)"] \arrow[ur, "h"'] &
\end{tikzcd}$$

In a category $\mathbb{C}$ with a subobject classifier Ω, the *powerobject* $\mathcal{P}A$ of A is the exponential Ω^A of Ω by A.[5]

[5] There is an alternative definition of powerobject which does not require the presence of a subobject classifier: One could call a powerobject of A any object which represents the functor $\mathrm{Sub}((-) \times A)$, in the sense that $\mathrm{Hom}(B, \mathcal{P}A) \cong \mathrm{Sub}(B \times A)$ for any B. This definition removes the need to include subobject classifiers as part of the definition of a topos below, as any category with finite limits and power objects has $\mathcal{P}(1)$ as a subobject classifier.

The map ε above is often called *evaluation*, because that is precisely its role in $\mathbb{S}\mathbf{ets}$: The exponential B^A of a set B by a set A is precisely the set of functions $A \to B$, and $\varepsilon(a, f : A \to B) = f(a)$.

Actually, one can think of $A \times (-)$ and $(-)^A$ as functors. Let $F = A \times (-)$. To see F as a functor, let $f : B \to C$ be any arrow and define $F(f)$ to be the unique arrow such that the following diagram commutes.

$$\begin{tikzcd}
A \arrow[d, "\mathrm{id}_A"] & A \times B \arrow[l] \arrow[r] \arrow[d, "F(f)"] & B \arrow[d, "f"] \\
A & A \times C \arrow[l] \arrow[r] & C
\end{tikzcd}$$

Namely, $F(f) = \mathrm{id}_A \times f$. Similarly, let $G = (-)^A$. To see G as a functor, let $G(f)$ be the unique arrow such that the following diagram commutes

$$\begin{tikzcd}
A \times C^A \arrow[r] & C \\
A \times B^A \arrow[r] \arrow[u, "\mathrm{id}_A \times G(f)"] & B \arrow[u, "f"']
\end{tikzcd}$$

A rephrasing of Definition 3.7 is readily available: D is the exponential of B by A if and only if there is a bijection

$$\mathrm{Hom}(A \times C, B) \cong \mathrm{Hom}(C, D)$$

natural in B and C. This proves the following lemma.

Lemma 3.8. *Let $\mathbb{C}$ be a category with finite limits and exponentials.*[6] *Then, for each $A \in \mathrm{obj}(\mathbb{C})$, $A \times (-)$ and $(-)^A$ are functors, and $A \times (-) \dashv (-)^A$.*

Recall that in a Boolean algebra containing p, one finds the adjoint relation $p \wedge (-) \dashv p \Rightarrow (-)$. This essentially reveals that $p \Rightarrow q$ is the exponential of q by p for any point q.[7]

3.3 Toposes

Putting all of the above together, one obtains the definition of a topos.

[6] This kind of category is said to be *cartesian closed*.

[7] This should be no surprise, given modes ponens!

Definition 3.9. An *elementary topos* (or simply a *topos*) is a category with a subobject classifier, finite limits, and powerobjects.[8]

The following lemma is immediate from the definition of a topos and the universal mapping property for exponentials.

Lemma 3.10. *In a topos $\mathbb{E}$ with subobject classifier Ω, $\mathcal{P}(1) \cong \Omega$.*

The very simple list of properties that define toposes have many consequences, most of which highly nontrivial. Every of [13], [8], and [2] is a comprehensive guide to the topic, which I do not intend to give here. The following theorem lists the essential properties of toposes, whose proofs are found throughout [13]. From now on, let $\mathbb{E}$ denote an arbitrary topos.

Theorem 3.11. *The following statements hold in $\mathbb{E}$.*

(1) If an arrow is both monic and epic, then it is an isomorphism.

(2) $\mathbb{E}$ has finite colimits. In particular, $\mathbb{E}$ has coequalizers, coproducts, and an initial object 0.

(3) Pullbacks of epic arrows are epic.

(4) Every arrow of the form $A \to 0$ is an isomorphism.

3.4 The internal logic of a topos

As it will turn out, every $\mathrm{Sub}(A)$ in $\mathbb{E}$ is a Heyting algebra. This essentially says that the internal logic of $\mathbb{E}$ is intuitionistic. A paraphrasing of the proof found in [13] IV.§6 follows, as it is an instructive illustration of the properties listed in Theorem 3.11.

Lemma 3.12. *If $m : U \to A$ is a monic arrow in $\mathbb{E}$, then*

$$U \xrightarrow{\quad m \quad} A \overset{\mathrm{ch}(m)}{\underset{\mathrm{true}_A}{\rightrightarrows}} \Omega$$

is an equalizer diagram, where $\mathrm{true}_A = \mathrm{true}! : A \to 1 \to \Omega$. In other words, every monic arrow in $\mathbb{E}$ is an equalizer.

[8] In a previous footnote, an alternative definition was given for powerobjects which did not require the presence of a subobject classifier. Actually, the presence of powerobjects and a terminal object implies the presence of a subobject classifier, namely $\Omega = \mathcal{P}(1)$. This simplifies the definition of topos even further.

Given an arrow $f : A \to B$, call a monic arrow $m : M \to B$ the *image* of f if f factors through m and m factors through any other monic $n : N \to B$ that f factors through. If m is the image of f, one writes $M = f\text{"}A$ and $m = \mathrm{img}(f)$, as in the following diagram.

$$\begin{tikzcd}
A \arrow[r] \arrow[rr, bend left, "f"] \arrow[dr] & f\text{"}A \arrow[r, "\mathrm{img}(f)"] \arrow[d, dashed] & B \\
& N \arrow[ur, "n"] &
\end{tikzcd}$$

One should verify for themself that this indeed captures the notion of image for sets. By monicity of n, such a factorization is unique.

Corollary 3.13. (to Theorem 3.11) *Let $f : A \to B$ be an arrow in $\mathbb{E}$. There exists a monic $m : M \to B$ and an epic $h : A \to M$ such that m is the image of f and $f = mh$.*

Proof. Dual to the notion of *pullback* is that of *pushforward*, the colimit of a diagram of the form $\bullet \longleftarrow \bullet \longrightarrow \bullet$. By 3.11.(3), $\mathbb{E}$ has all pushforwards, and in particular there is a pushforward diagram

$$\begin{tikzcd}
B \arrow[r, "q_1"] & D \\
A \arrow[u, "f"] \arrow[r, "f"] & B \arrow[u, "q_2"']
\end{tikzcd}$$

in $\mathbb{E}$. Toposes have equalizers, so there is an equalizer diagram

$$\begin{tikzcd}
M \arrow[r, "m"] & B \arrow[r, shift left, "q_1"] \arrow[r, shift right, "q_2"'] & D
\end{tikzcd}$$

also in $\mathbb{E}$. Now, since (A, f) is a cone for the right hand side of the equalizer diagram, there is a unique arrow $h : A \to M$ such that $mh = f$. This gives a factorization of f. It is routine to check that equalizers are always monic, in particular m. Thus, to see that m is the image of f, it suffices to check its universal mapping property. Let $f = nk : A \to N \to B$ for some monic n. By Lemma 3.12, n is an equalizer, say of the pair $x_1, x_2 : B \rightrightarrows D'$. We have

$$x_1 f = x_1 nk = x_2 nk = x_2 f,$$

so by the universal property of pushouts, there is a unique $u : D \to D'$ such that $uq_i = x_i$, $i = 1, 2$. Moreover,

$$x_1 m = uq_1 m = uq_2 m = x_2 m,$$

so by the universal property of equalizers, m factors through n. This shows that m is the image of f.

To see that h is epic, first observe that f is epic if and only if m is an isomorphism. Indeed, if m is an isomorphism, then $q_1 = m^{-1}mq_1 = m^{-1}mq_2 = q_2$, which by the universal property of pushouts is eqivalent to the statement that f is epic (if f is epic, then $D = B$ and $q_1 = q_2 = \mathrm{id}_B$).

Next, apply the above factorization to h, say $h = nk : A \to N \to M$ with n monic. In light of Theorem 3.11.(1), it suffices to show that n is epic. Now, since mn is monic (any composition of monics is also monic) and $(mn)k = f$, by the universal mapping property for images there is a unique $t : M \to N$ such that $m = (mn)t$. The monicity of m then implies that $\mathrm{id}_M = nt$, so that t is a right inverse to n. Any arrow with a right inverse is epic, so n is epic. □

Recall that a *distributive lattice* is a poset L with a top $\top$ and bottom $\perp$ and two associative, commutative binary operations $\wedge, \vee : L \times L \to L$ such that $a \wedge b = a$ if and only if $a \leqslant b$ if and only if $a \vee b = b$, and $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$.

<u>**Lemma 3.14.**</u> *Given any object A of $\mathbb{E}$, $U, V \subseteq A$, define $U \vee V$ to be the image of the unique map $U \sqcup V \to A$ which restricts to the inclusions $U \to A, V \to A$. The operations $\wedge$, defined earlier, and $\vee$, as well as the identity id_A and the initial object $0 \xrightarrow{?} A$, give the poset $\mathrm{Sub}(A)$ a distributive lattice structure with $\top = \mathrm{id}_A$ and $\perp = ?$. Moreover, for any $f : A \to B$, $f^{-1} : \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ preserves $\wedge$.*

Proof. The arrows ? and id_A are trivially monic. Since every monic into A factors through id_A, $\mathrm{id}_A = \top$. Moreover, since ? factors through every monic into A, $? = \perp$.

Any monic $V \to U$ that factors through $U \to A$ turns the diagram

$$\begin{tikzcd} V \arrow[r] & A \\ V \arrow[u, "\mathrm{id}_V"] \arrow[r] & U \arrow[u] \end{tikzcd}$$

into a pullback diagram, so $V \leqslant U$ if and only if $V \wedge U = V$. To see the corresponding equation for $\vee$, observe that $U \leqslant V \vee U$ holds always, and that $V \vee U$ factors through $U \to A$ if and only if there is a factorization of $V \vee U \to A$ through $U \to A$.

Product and sum operations are always commutative, up to isomorphism, so $\wedge$ and $\vee$ are commutative.

It follows directly from the definition that $\sqcup$ is associative (up to isomorphism). To see that $\vee$ is associative, let $m_U, m_V, m_W : U, V, W \subseteq A$ respectively, and consider the following diagram.

$$\begin{tikzcd}
(U \vee V) \vee W \arrow[r] & A & U \vee (V \vee W) \arrow[l] \\
(U \vee V) \sqcup W \arrow[u] & U \sqcup V \sqcup W \arrow[l] \arrow[r] & U \sqcup (V \vee W) \arrow[u]
\end{tikzcd}$$

The inclusion $W \to U \sqcup V \sqcup W$ gives (by composition) an arrow $W \to U \vee (V \vee W)$. Since $U \vee V \to A$ is the image of the arrow $[m_U, m_V] : U \sqcup V \to A$ induced by m_U and m_V, it factors through $U \vee (V \vee W) \to A$. Together, $W \to U \vee (V \vee W)$ and $U \sqcup V \to U \vee (V \vee W)$ induce an arrow $(U \vee V) \sqcup W \to U \vee (V \vee W)$. This gives an alternate factorization of $(U \vee V) \sqcup W \to A$, so $U \vee (V \vee W) \to A$ factors through $(U \vee V) \vee W \to A$. Similarly, $U \vee (V \vee W) \to A$ factors through $(U \vee V) \vee W \to A$. By uniqueness of the arrows $U \vee (V \vee W) \leftrightarrows (U \vee V) \vee W$, they are mutually inverse.

That $\wedge$ is distributive follows from the second statement of the lemma. Indeed, if $m : U \to A$ is a monic arrow, then $m^{-1} : \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ is given by $m^{-1}(V) = U \wedge V$. Whence, if m^{-1} is a lattice homomorphism, then

$$U \wedge (V \vee W) = m^{-1}(V \vee W)m^{-1}(V) \vee m^{-1}(W) = (U \wedge V) \vee (U \wedge W).$$

So let $f : A \to B$. To see that f^{-1} preserves $\wedge$, consider the following commutative cube

$$\begin{tikzcd}
& A \arrow[rr, "f"] & & B \\
f^{-1}V \arrow[ur] \arrow[rr] & & V \arrow[ur] & \\
& f^{-1}U \arrow[uu] \arrow[rr] & & U \arrow[uu] \\
f^{-1}V \wedge f^{-1}U \arrow[uu] \arrow[ur] \arrow[rr, dashed] & & U \wedge V \arrow[uu] \arrow[ur] &
\end{tikzcd}$$

in which the dashed arrow is the arrow induced by the pullback that is its right-hand face. Every face of the cube but those adjacent to the dashed arrow are pullbacks by definition. Verifying that the bottom face is a pullback is routine. $\square$

One notable property of the subobject relation is its stability: Not only is the subobject relation transitive, but it respects $\wedge$ and $\vee$. In other words, if U and V are subobjects of B and B is a subobject of A, then the *intersection* $U \wedge V$ of U and V as subobjects of B is equal to the intersection $U \wedge V$ in A. In particular, the following is a pullback diagram for any $U, V \subseteq A$.

$$\begin{tikzcd} U \arrow[r] & U \vee V \\ U \wedge V \arrow[u] \arrow[r] & V \arrow[u] \end{tikzcd}$$

Actually, if $U \wedge V = 0$, this is also a pushforward diagram, so that $U \vee V \cong U \sqcup V$. Note also that this gives a direct relationship between $\vee$ and $\wedge$ without any explicit reference to A whatsoever.

An alternate proof of the equation $f^{-1}(U \wedge V) = f^{-1}U \wedge f^{-1}V$ can be obtained from Theorem 2.14 and the next lemma.

Lemma 3.15. *Let A and B be objects of $\mathbb{E}$ and $f : A \to B$. Thought of as a functor, $f^{-1} : \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ has a left adjoint $\exists_f : \mathrm{Sub}(A) \to \mathrm{Sub}(B)$.*

Proof. The left adjoint is defined as follows: Let $m : U \to A$ be monic. Then $\exists_f U : f"U \to B$ is defined to be the image of $fm : U \to B$. A routine calculation verifies that $\exists_f$ is monic. To see that $\exists_f \dashv f^{-1}$, one shows that $f"U \to B$ factors through a subobject V if and only if $U \to A$ factors through $f^{-1}V \to A$. The forward implication is obtained from the universal mapping property for pullbacks, and the backward direction is obtained from the universal mapping property for images. $\square$

Since $\mathrm{Sub}(-)$ is represented by Ω, the distributive lattice structure of $\mathrm{Sub}(A)$ can be represented *internally* in $\mathbb{E}$: Recall the natural isomorphism $\mathrm{Hom}(-, \Omega) \times \mathrm{Hom}(-, \Omega) \cong \mathrm{Hom}(-, \Omega \times \Omega)$. Since every natural transformation $\mathrm{Hom}(-, \Omega \times \Omega) \to \mathrm{Hom}(-, \Omega)$ is of the form f_* for some $f : \Omega \times \Omega \to \Omega$, the operations $\wedge$ and $\vee$ are given by unique arrows $\wedge, \vee : \Omega \times \Omega \to \Omega$. Moreover, it follows from naturality that the following diagrams commute.

$$\begin{tikzcd} \Omega \times \Omega \times \Omega \arrow[r, "\pi_1 \times \wedge"] \arrow[d, "\wedge \times \pi_3"'] & \Omega \times \Omega \arrow[d, "\wedge"] \\ \Omega \times \Omega \arrow[r, "\wedge"] & \Omega \end{tikzcd} \qquad \begin{tikzcd} \Omega \times \Omega \times \Omega \arrow[r, "\pi_1 \times \vee"] \arrow[d, "\vee \times \pi_3"'] & \Omega \times \Omega \arrow[d, "\vee"] \\ \Omega \times \Omega \arrow[r, "\vee"] & \Omega \end{tikzcd}$$

The two diagrams above indicate associativity.

$$\begin{tikzcd}
\Omega \times \Omega \times \Omega \arrow[r, "\pi_1 \times \vee"] \arrow[d, "(\pi_1 \times \pi_2, \pi_1 \times \pi_3)"'] & \Omega \times \Omega \arrow[d, "\wedge"] \\
\Omega \times \Omega \times \Omega \times \Omega \arrow[r, "\wedge(\vee \times \vee)"] & \Omega
\end{tikzcd}$$

This last diagram expresses distributivity. Similar diagrams express the remaining definitive properties of a lattice, and the object Ω is said to be an *internal lattice* of $\mathbb{E}$ because they a commute. In fact, in a topos, Ω is much more than a mere lattice.

Lemma 3.16. *In $\mathbb{E}$, the distributive lattice* $\mathrm{Sub}(1)$ *is a Heyting algebra.*

Proof. It suffices to show that for any $U \subseteq 1$, $U \wedge (-)$ has a right adjoint. This is constructed as follows. First, observe that for any object W, $W \subseteq 1$ if and only if $|\,\mathrm{Hom}(A, W)| \leqslant 1$ for every A. Given $V \subseteq 1$, the evaluation map $\varepsilon : U \times V^U \to V$ is the unique arrow between these objects. So, let $i, j : A \to U^V$ for some object A. A brief calculation reveals that $\pi_1 : U \times 1 \to U$ is an isomorphism. This leads to the commutative diagram below.

$$\begin{tikzcd}
U \times V^U \arrow[r, "\varepsilon"] & U \\
U \times A \arrow[u, shift left, "\mathrm{id}_U \times a"] \arrow[u, shift right, "\mathrm{id}_U \times b"'] \arrow[ur, "\pi_1"'] &
\end{tikzcd}$$

By the universal mapping property for exponentials, $a = b$. This makes $V^U \subseteq 1$.

To see that $U \wedge (-) \dashv (-)^U$, simply observe that $U \wedge V = U \times V$ in $\mathrm{Sub}(1)$, since

$$\begin{tikzcd}
U \arrow[r] & 1 \\
U \times V \arrow[u] \arrow[r] & V \arrow[u]
\end{tikzcd}$$

is a pullback diagram. $\square$

The right adjoint to $\wedge$ in a Heyting algebra is always denoted $\Rightarrow$ and called its *implication operator*. To see that Ω is an *internal Heyting algebra*, ie. that it has an implication operator $\Rightarrow : \Omega \times \Omega \to \Omega$, a brief return to slice categories is convenient.

Theorem 3.17. *For any object A in $\mathbb{E}$, $\mathbb{E}/A$ is a topos.*

Proof (Sketch). Power objects and pullbacks in $\mathbb{E}/A$ are evidently computed in $\mathbb{E}$, and the terminal object in $\mathbb{E}/A$ is id_A. This already shows that $\mathbb{E}/A$ has all finite limits by Lemma 2.12.

The subobject classifier in $\mathbb{E}/A$ is actually $\pi_1 : A \times \Omega \to A$, with the "truth" arrow being $\mathrm{id}_A \times \mathrm{true} : A \times 1 \to A \times \Omega$. Of course, $A \times 1 \cong A$, so actually $\pi_1 : A \times 1 \to A$ is terminal in $\mathbb{E}/A$. To see that $A \times \Omega \to A$ is the subobject classifier, let $m : U \to X$ be a monic arrow factoring through an arrow $a : X \to A$ in $\mathbb{E}/A$. Since m corresponds to the unique $\mathrm{ch}(m) : X \to \Omega$, one constructs the arrow $(a, \mathrm{ch}(m)) : X \to A \times \Omega$ and observes that the following diagram commutes.

$$\begin{tikzcd}
X \arrow[rr, "{(a,\mathrm{ch}(a))}"] \arrow[dr, "a"] & & A \times \Omega \arrow[dl] \\
& A & \\
U \arrow[uu, "m"] \arrow[ur, "u"] \arrow[rr, "{(u,!)}"] & & A \times 1 \arrow[ul] \arrow[uu, "\mathrm{id}_A \times \mathrm{true}"']
\end{tikzcd}$$

A routine check verifies that the outer square is a pullback. $\square$

For example, $\mathbb{S}\mathbf{ets}/A$ (and therefore also $\mathbb{S}\mathbf{ets}^A$) is a topos for any set A.

One remarkable aspect of slice categories is the several ways in which the slice operation is given functor structure $\mathbb{C} \to \mathbb{C}\mathbf{at}$: Let $f : A \to B$ be any arrow in $\mathbb{C}$. Then there is a contravariant functor $f^* : \mathbb{C}/B \to \mathbb{C}/A$ defined by the pullback

$$\begin{tikzcd}
A \arrow[r, "f"] & B \\
f^*X \arrow[u, "f^*b"] \arrow[r] & X \arrow[u, "b"']
\end{tikzcd}$$

There are two other ways to turn $\mathbb{C}/(-)$ into a functor, whose relation to f^* is recorded in the following lemma.

Lemma 3.18. *Let $f : A \to B$ in $\mathbb{C}$. The* pullback functor *$f^* : \mathbb{C}/B \to \mathbb{C}/A$ has both a left and a right adjoint,*

$$\sum_f \dashv f^* \dashv \prod_f .$$

Returning to the issue at hand, since $1 = \mathrm{id}_A$ in $\mathbb{E}/A$, a monic arrow into 1 is simply a subobject of A.

Lemma 3.19. *For any object A in $\mathbb{E}$, the map*

$$\mathrm{Sub}_{\mathbb{E}/A}(1) \to \mathrm{Sub}_{\mathbb{E}}(A),$$

which forgets the slice category structure, is a natural bijection in $\mathbb{E}$.

To make sense of naturality in A, simply observe that the pullback operator f^* readily extends to a functor $f^* : \mathbb{E}/B \to \mathbb{E}/A$.

Since $\mathrm{Sub}_{\mathbb{E}/A}(1)$ is a Heyting algebra, so is $\mathrm{Sub}_{\mathbb{E}}(A)$, for any A in $\mathbb{E}$. Thus, the implication operator for $\mathrm{Sub}(A)$ corresponds via the Yoneda lemma to an implication operator $\Rightarrow : \Omega \times \Omega \to \Omega$ in $\mathbb{E}$. This concludes a proof of the main theorem of this section.

Theorem 3.20. *The object Ω is an internal Heyting algebra with top $\top = \mathrm{true}$ and bottom $\bot = \mathrm{false}$, given in the following pullback diagram.*

$$\begin{tikzcd} 1 \arrow[r, "\mathrm{false}"] & \Omega \\ 0 \arrow[u] \arrow[r] & 1 \arrow[u, "\mathrm{true}"'] \end{tikzcd}$$

It is well-known to logicians that there is no material distinction between formulas and the sets they define. In the topos theory literature, this is embodied by the equating formulas with arrows into Ω. In fact, one can define a partial order, *entailment*, on the arrows into Ω by setting

$$\varphi \vdash \psi \text{ if and only if } \varphi \Rightarrow \psi = \mathrm{true},$$

where $\varphi \Rightarrow \psi = (\Rightarrow) \circ (\varphi \times \psi)$, for any $\varphi, \psi : A \to \Omega$. This partial order has the notable property that if $\varphi = \mathrm{ch}(U \subseteq A)$ and $\psi = \mathrm{ch}(V \subseteq A)$, then $\varphi \vdash \psi$ if and only if $U \leqslant_A V$. Furthermore, define $\varphi \wedge \psi = \wedge \circ (\varphi \times \psi)$ and $\varphi \vee \psi = \vee \circ (\varphi \times \psi)$. A calculation involving Lemma 3.5 is used to show that each of the diagrams below are pullback diagrams.

$$\begin{tikzcd} A \arrow[r, "\varphi\wedge\psi"] & \Omega \\ U \wedge V \arrow[u] \arrow[r] & 1 \arrow[u, "\mathrm{true}"'] \end{tikzcd} \quad \begin{tikzcd} A \arrow[r, "\varphi\vee\psi"] & \Omega \\ U \vee V \arrow[u] \arrow[r] & 1 \arrow[u, "\mathrm{true}"'] \end{tikzcd} \quad \begin{tikzcd} A \arrow[r, "\varphi\Rightarrow\psi"] & \Omega \\ U \Rightarrow V \arrow[u] \arrow[r] & 1 \arrow[u, "\mathrm{true}"'] \end{tikzcd}$$

One can further define $\neg\varphi = \varphi \Rightarrow \bot$, but an equivalent description of the operator $\neg$ is given by the observation that any arrow from 1, ie. any *global element*, is monic.

Hence, false : $1 \to \Omega$ is monic, so $\neg$ can be defined as in the following pullback diagram.

$$\begin{tikzcd} \Omega \arrow[r, "\neg"] & \Omega \\ 1 \arrow[u, "\text{false}"] \arrow[r] & 1 \arrow[u, "\text{true}"'] \end{tikzcd}$$

In other words, $\neg\,\text{false} = \text{true}$. Moreover, since both of the inner squares in the diagram

$$\begin{tikzcd} 1 \arrow[r, "\text{true}"] & \Omega \arrow[r, "\neg"] & \Omega \\ 0 \arrow[u] \arrow[r] & 1 \arrow[u, "\text{false}"] \arrow[r] & 1 \arrow[u, "true"'] \end{tikzcd}$$

are pullbacks, so is the outer square. Whence, $\text{false} = \text{ch}(? : 0 \to 1) = \neg\,\text{true}$ as well.

The other standard properties of Heyting algebras apply to the maps $\wedge, \vee, \Rightarrow : \Omega \times \Omega \to \Omega$ and $\neg : \Omega \to \Omega$ as well, including

$$x \leqslant \neg\neg x \tag{1}$$

$$\neg\neg\neg x = \neg x \tag{2}$$

$$\neg(x \wedge y) = \neg x \wedge \neg y \tag{3}$$

$$(x \vee y) \Rightarrow z = (x \Rightarrow z) \wedge (y \Rightarrow z), \tag{4}$$

and so on, for any $x, y, z : \bullet \to \Omega$.

3.5 Presheaves, Boolean toposes

One might wonder at this point whether the internal logic of a topos is actually Boolean. In fact, this is rarely the case.

Lemma 3.21. *Let $\mathbb{C}$ be any small category whatsoever. Then $\widehat{\mathbb{C}} = \mathbb{S}\mathbf{ets}^{\mathbb{C}^{op}}$ is a topos.*

Proof. In $\mathbb{S}\mathbf{ets}^{\mathbb{C}}$, limits (and colimits) are computed pointwise. That is, if $P : \mathbb{I} \to \mathbb{S}\mathbf{ets}^{\mathbb{C}}$ and $F : \mathbb{C} \to \mathbb{S}\mathbf{ets}$, then

$$(\varprojlim P)F = \varprojlim(P \circ F),$$

where the first limit is taken in $\mathbb{S}\mathbf{ets}^{\mathbb{C}}$ and the second limit is taken in $\mathbb{S}\mathbf{ets}$. The category of sets is complete (and cocomplete), so this actually reveals that $\mathbb{S}\mathbf{ets}^{\mathbb{C}}$ is complete as well. In particular, $\mathbb{S}\mathbf{ets}^{\mathbb{C}}$ has the constant functor $\Delta(\{\varnothing\})$ (ie. $\Delta 1$) as its terminal object. Replacing $\mathbb{C}$ with $\mathbb{C}^{op}$, one obtains the stated result.

To construct the subobject classifier of $\widehat{\mathbb{C}}$, the following definition is helpful.

Definition 3.22. For an object C in a category, a *sieve on* C is a set S of arrows such that

$$(f : A \to C) \in S \text{ and } (g : B \to A) \text{ implies } (f \circ g \in S).$$

Given a sieve S on C and an object B of $\mathbb{C}$, define

$$SB = \{f \in S \mid \mathrm{dom}(f) = B\}.$$

Moreover, for an arrow $f : A \to B$, define $S(f) : SB \to SA$ by

$$S(f)(h) = hf.$$

This gives a contravariant functor structure to S.

By definition, for any sieve S on C, $SB \subseteq (\mathbf{y}C)B$. This suggests that $S \subseteq \mathbf{y}C$ in $\widehat{\mathbb{C}}$, which is indeed the case. A short calculation verifies that the diagram

$$\begin{tikzcd} SA \arrow[r, "\subseteq"] & \mathrm{Hom}(A, C) \\ SB \arrow[u, "S(f)"] \arrow[r, "\subseteq"] & \mathrm{Hom}(B, C) \arrow[u, "f^*"'] \end{tikzcd}$$

commutes, and a natural transformation of functors into $\mathbb{S}$**ets** is monic if and only if every member of its family of functions is injective. The converse holds as well: Every subfunctor (subobject) S of $\mathbf{y}C$ defines a sieve for C, the set

$$\{f \mid (\exists B)(f \in SB)\},$$

also denoted S. Subfunctors of $\mathbf{y}C$ and sieves on C are henceforth referred to interchangeably.

It is worthwhile to untangle the definition of subfunctor, in particular to uncover the following observation: Given a pair of presheaves $U \subseteq F$, there is no harm in assuming that $UA \subseteq FA$ for any object A. Thus, since the inclusion $UA \hookrightarrow FA$ is natural, given any arrow $f : B \to A$ the function $F(f) : FA \to FB$ restricts to the function $U(f) : UA \to UB$. In other words, the image of UA under $F(f)$ is contained in UB.

Proof of Lemma 3.21 (continued). For each object C of $\mathbb{C}$, let ΩC denote the set of sieves on C. Given $f : C' \to C$, define $\Omega(f) : \Omega C \to \Omega C'$ as follows: Using the fact that $\widehat{\mathbb{C}}$ is complete, form the following pullback diagram.

$$\begin{tikzcd}
S \arrow[r, "\subseteq"] & \mathrm{Hom}(-, C) \\
f * S \arrow[u] \arrow[r, "\subseteq"] & \mathrm{Hom}(-, C') \arrow[u, "f_*"']
\end{tikzcd}$$

One can check that, in fact,

$$f * S = \{h : \bullet \to C \mid f \circ h \in S\}.$$

That the set $f * S$ is a sieve on C can be seen in two ways: On the one hand, the pullback of monic arrows is monic. More concretely, if $g : B \to A$, $h : A \to C$, and $h \in f * S$ (ie. $fh \in S$), then $(fh)g \in S$ because S is a sieve on C', and therefore $hg \in f * S$. The definition $\Omega(f)(S) = f * S$ gives a presheaf structure to Ω.

As the notation suggests, $\Omega : \mathbb{C}^{op} \to \mathbb{S}\mathbf{ets}$ is the subobject classifier of $\widehat{\mathbb{C}}$. Let $\mathrm{true} : 1 \to \Omega$ be the natural transformation defined at each object C by

$$\varnothing \longmapsto \mathbf{y}C = \{f \mid (\exists B \in \mathrm{obj}(\mathbb{C}))(f : B \to C)\}.$$

The presheaf $\mathbf{y}C$ is trivially a sieve, and is in fact the maximal sieve on C. For any presheaf $F : \mathbb{C}^{op} \to \mathbb{S}\mathbf{ets}$, $U \subseteq F$, and object C of $\mathbb{C}$, let $\mathrm{ch}(U)_C : FC \to \Omega C$ be the function defined

$$\mathrm{ch}(U)_C(s) = \{f \mid (\exists C')(f : C' \to C \text{ and } F(f)(s) \in UC')\}.$$

Indeed, $\mathrm{ch}(U)_C(s)$ is a sieve on C, for if $(f : C' \to C) \in \mathrm{ch}(U)_C(s)$ and $h : B \to C'$, then

$$F(fh)(s) = F(h) \circ F(f)(s) = F(h)(F(f)(s)) = U(h)(F(f)(s)) \in UB,$$

since $F(f)(s) \in UC'$.

The following diagram is a pullback in $\mathbb{S}\mathbf{ets}$.

$$\begin{tikzcd}
FC \arrow[r, "\mathrm{ch}(U)_C"] & \Omega C \\
UC \arrow[u, "\subseteq"] \arrow[r, "!"] & 1 \arrow[u, "\mathrm{true}"']
\end{tikzcd}$$

To check this, it suffices to verify that $\mathrm{ch}(U)_C(s) = \mathbf{y}C$ if and only if $s \in UC$. The reverse direction is by definition. For the forward direction, apply the definition of $\mathrm{ch}(U)_C(s)$ to the arrow id_C.

The power objects in $\widehat{\mathbb{C}}$ are slightly harder to find: Assuming they exist,

$$\mathcal{P}(F)C \cong \mathrm{Hom}(\mathbf{y}C, \mathcal{P}(F)) \cong \mathrm{Hom}(\mathbf{y}C \times F, \Omega) \cong \mathrm{Sub}(\mathbf{y}C \times F)$$

for any object C of $\mathbb{C}$. This motivates the definition

$$\mathcal{P}(F) = \mathrm{Sub}(\mathbf{y}(-) \times F).$$

Since $\mathbf{y}$ and $F \times (-)$ are covariant and $\mathrm{Sub}(-)$ is contravariant, this defines a contravariant functor $\mathcal{P}$ into $\mathbb{S}\mathbf{ets}$.

The evaluation map $\varepsilon : F \times \mathcal{P}(F) \to \Omega$ is defined by the equation

$$\varepsilon_C(s, U) = \mathrm{ch}(U)_C(\mathrm{id}_C, s)$$

for each $C \in \mathrm{obj}(\mathbb{C})$, $U \subseteq \mathbf{y}C \times F$, and $s \in FC$. Verifying that this defines an evaluation map is routine. □

Consider the topos $\widehat{\mathbb{P}}$ of presheaves on a poset $\mathbb{P}$. A sieve on a point $p \in \mathbb{P}$ is a downward closed subset of $\downarrow(p)$, and a subfunctor $F \subseteq 1$ in $\widehat{\mathbb{P}}$ is given uniquely by the sieve of points q for which $Fq = 1$. Hence, the internal logic of $\widehat{\mathbb{P}}$ is the Heyting algebra $RO(\mathbb{P})$ of sieves of downward closed subsets of $\mathbb{P}$. Notice that this is often not a Boolean algebra: For instance, if $\mathbb{P}$ is non-linear and bottomless[9].

Definition 3.23. A topos is called *Boolean* if its internal logic is classical, ie. $\mathrm{Sub}(1)$ is a Boolean algebra and Ω is an internal Boolean algebra.

Although Boolean toposes are somewhat rare in reality, there is a well-known method of producing Boolean toposes from non-Boolean toposes, called *Booleanization*. This requires a whole new idea, explored in the next section. For now, the following lemma (Proposition 1 from [13], VI.§3) provides a useful characterization of Boolean toposes.

Lemma 3.24. *In a topos $\mathbb{E}$, the following are equivalent.*

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

(ii) The negation operator $\neg : \Omega \to \Omega$ satisies $\neg\neg = \mathrm{id}_\Omega$.

(iii) For every A in $\mathbb{E}$, $\mathrm{Sub}(A)$ is a Boolean algebra.

(iv) For every $U \subseteq A$, $\neg U \vee U = A$.

(v) The map $[\mathrm{true}, \mathrm{false}] : 1 + 1 \to \Omega$ is an isomorphism.

[9]For a concrete example, let $\mathbb{P} = \{a, b, c\}$ with $a, b \leqslant c$. Then $\neg\neg\{b\}$ in $RO(\mathbb{P})$ is the set $\{b, c\}$, which is not $\{b\}$.

Proof. The equivalence between (i) and (ii) is well-documented, as well as the equivalence of (iii) and (iv). Points (i) and (iii) are equivalent via an interchange of structure between $\mathrm{Sub}(-)$ and $\mathrm{Hom}(-, \Omega)$. It therefore suffices to show that (iv) implies (v) implies (ii), since (iv) implies (ii).

Assume (iv). Thinking of true and false : $1 \to \Omega$ as subobjects of Ω, observe that $\mathrm{ch}(\mathrm{true} \vee \mathrm{false}) = \mathrm{ch}(\mathrm{true} \vee \neg\, \mathrm{false}) = \top$. Thus, $1 \vee 1 = \Omega$. By definition, $\mathrm{true} \wedge \mathrm{false} = \bot$, so actually $[\mathrm{true}, \mathrm{false}]$ is a monic arrow into Ω representing the subobject $\mathrm{true} \vee \mathrm{false} = \top$. This makes $[\mathrm{true}, \mathrm{false}]$ an isomorphism.

Assuming (v), $1 \sqcup 1$ is a subobject classifier equipped with the arrow $h \circ \mathrm{true} : 1 \to 1 \sqcup 1$, where $h : \Omega \cong 1 \sqcup 1$ is the inverse of $[\mathrm{true}, \mathrm{false}]$. The negation operator on $1 \sqcup 1$ is then

$$\neg' = h \circ \neg \circ [\mathrm{true}, \mathrm{false}] = h \circ [\neg\, \mathrm{true}, \neg\, \mathrm{false}] = h \circ [\mathrm{false}, \mathrm{true}].$$

Whence,

$$\neg' \circ \neg' = (h \circ \neg \circ [\mathrm{true}, \mathrm{false}]) \circ (h \circ [\mathrm{false}, \mathrm{true}]) = h \circ \neg \circ [\mathrm{false}, \mathrm{true}] = h \circ [\mathrm{true}, \mathrm{false}] = \mathrm{id}_{1 \sqcup 1}.$$

Since h is an isomorphism, this makes $\neg\neg = \mathrm{id}_\Omega$. $\square$

4 The construction of $\mathbf{Sh}(\mathbb{P}, \neg\neg)$

As was stated at the end of the previous section, few toposes are Boolean. However, every topos contains a Boolean topos as a full subcategory $\mathbf{Sh}_{\neg\neg}(\mathbb{E})$, which is maximal in the sense that the inclusion functor $i : \mathbf{Sh}_{\neg\neg}(\mathbb{E}) \to \mathbb{E}$ has a left adjoint

$$\mathbf{a} : \mathbb{E} \to \mathbf{Sh}_{\neg\neg}(\mathbb{E}).$$

The category $\mathbf{Sh}_{\neg\neg}(\mathbb{E})$ is the full subcategory of what are called *sheaves for* $\neg\neg$ in $\mathbb{E}$, and the functor $\mathbf{a}$ is called the *sheafification* functor. What exactly all of this means is essentially what will be recorded in this section.

4.1 Sheaves

Sheaves are best understood when observed from multiple perspectives. Possibly the quickest route is via modal logic.

As was observed in the previous section, toposes interpret intuitionistic propositional logic in the sense that arrows into Ω form a Heyting algebra. Modalities are traditionally given by unary operators on formulas. In the internal logic of a topos, then, modalities are simply arrows of the form $\Omega \to \Omega$. The sort of modality that is of interest in the *Booleanization* process is inspired by the Gödel-Gentzen embedding (or, perhaps, Glivenko's theorem) of classical logic into intuitionistic logic.

Definition 4.1. In a topos $\mathbb{E}$ whose subobject classifier is Ω, an arrow $j : \Omega \to \Omega$ is called a *Lawvere-Tierney modality* (*L-T modality*) if

$$jj = j, \ j \circ \mathrm{true} = \mathrm{true}, \ \text{and } j \circ \wedge = \wedge(j \times j).$$

<u>Lemma 4.2.</u> *The* double-negation operator $\neg\neg : \Omega \to \Omega$ *in* $\mathbb{E}$ *is a L-T topology.*

Proof. Recall the properties (1)-(4) of Heyting algebras stated at the end of §3.4. $\square$

Every arrow $j : \Omega \to \Omega$ induces a *closure operation* $U \mapsto \overline{U}$ on $\mathrm{Sub}(A)$ for every A, defined by $\mathrm{ch}(\overline{U}) = j \circ \mathrm{ch}(U)$. Translating the definition of L-T modality into properties of the closure operation, one sees that $U \mapsto \overline{U}$ satisfies

$$\overline{\overline{U}} = \overline{U}, \ U \subseteq \overline{U}, \ \text{and } \overline{U \wedge V} = \overline{U} \wedge \overline{V}$$

if and only if j is a L-T modality.

Given a L-T modality whose closure operation is $\overline{(-)}$, Call a monic $U \subseteq A$ *closed* if $\overline{U} = U$ and *dense* if $\overline{U} = A$. This terminology applies equally well to the subobjects if their associated monics are clear.

Definition 4.3. Let j be a L-T modality in $\mathbb{E}$. An object X of $\mathbb{E}$ is called a *sheaf* (*for* j) if for any dense $m : U \to A$, every arrow $x : U \to X$ extends to a unique arrow $\bar{x} : A \to X$, as in the following commutative diagram

$$\begin{tikzcd} U \arrow[r, "m"] \arrow[dr, "x"'] & A \arrow[d, dashed, "\bar{x}"] \\ & X \end{tikzcd}$$

The category $\mathbf{Sh}_j(\mathbb{E})$ is then the full subcategory of $\mathbb{E}$ consisting of the sheaves for j.

The objects of $\mathbf{Sh}_{\neg\neg}(\mathbb{E})$ are called $\neg\neg$-*sheaves* or *double-negation* sheaves.

Notice that an immediate consequence of the definition is that the terminal object 1 of $\mathbb{E}$ is a sheaf, since for every morphism $U \to X$ whatseoever, the terminal arrow $! : U \to 1$ extends to the unique terminal arrow $! : X \to 1$. This means that $\mathbf{Sh}_j(\mathbb{E})$ and $\mathbb{E}$ have the same temrinal object.

4.2 Grothendieck Topologies

Sheaves for a L-T modality in a topos of the form $\widehat{\mathbb{C}}$ also have a different form, given by something called a *Grothendieck topology* for $\mathbb{C}$. Starting with a L-T modality $j : \Omega \to \Omega$ in $\widehat{\mathbb{C}}$, its corresponding Grothendieck topology is given by the functor $J : \mathbb{C}^{op} \to \mathbb{S}\mathbf{ets}$ defined by

$$JC = \{S \in \Omega C \mid j_C(S) = \mathbf{y}C\},$$

and for an arrow $f : C \to C'$,

$$J(f)(S \in JC') = f * S,$$

which is well-defined by the naturality of j. Loosely, j determines a selection of sieves on each object of $\mathbb{C}$. These will be called *covering sieves of* C, and one says that a sieve S *covers* C if $S \in JC$.

Lemma 4.4. *The functor J satisfies the following three properties.*

(G1) For any object C of $\mathbb{C}$, $\mathbf{y}C \in JC$.

(G2) If $S \in JC$ and $f : D \to C$, then $f * S \in JD$.

(G3) If $S \in JC$ and R and sieve on C such that $f * R \in JD$ for any $f : D \to C$ in S, then $R \in JC$.

Proof. (G1) follows from $j \circ \text{true} = \text{true}$, since $\text{true}_C : 1 \to \Omega C$ is the maximal sieve for any C.

(G2) is a triviality due to the naturality of j and the fact that J is a subfunctor of Ω.

To see that (G3) holds, let S and R be as described. By definition of J, for any $f : D \to C$ one has $f * (j_C(R)) = j_D(f * R) = \mathbf{y}D$ by assumption. This puts $\text{id}_D \in f * (j_C(R))$, or equivalently $f \in j_C(R)$. This makes $S \subseteq j_C(R)$. Now, since $j \circ \wedge = \wedge \circ (j \times j)$, j commutes with intersections of sieves. In particular,

$$j_C(S) = j_C(S \cap j_C(R)) = j_C(S) \cap j_C j_C(R) = j_C(S) \cap j_C(R),$$

so that

$$j_C(R) \subseteq \mathbf{y}C \subseteq j_C(S) \subseteq j_C(R).$$

Thus, $R \in JC$. $\square$

Subfunctors of Ω satisfying (G1)-(G3) are generally called *Grothendieck topologies* for the category $\mathbb{C}$, and essentially all of classical sheaf theory can be deduced from these three properties. Moreover, every Grothendieck topology for $\mathbb{C}$ corresponds to a L-T modality in $\widehat{\mathbb{C}}$, and the two correspondances are inverse to one another. This makes the notion of a L-T modality a direct generalization of Grothendieck topologies, as L-T modalities are defined for arbitrary toposes.

Lemma 4.5. *Let $j : \Omega \to \Omega$ be a L-T topology in $\widehat{\mathbb{C}}$, J its corresponding Grothendieck topology, and F and object in $\widehat{\mathbb{C}}$. Then F is a sheaf for j if and only if for every object C of $\mathbb{C}$ and any covering sieve S for C, every natural transformation $x : S \to F$ extends to a unique natural transformation $\mathbf{y}C \to F$, as in*

$$\begin{tikzcd} S \arrow[r, "\subseteq"] \arrow[dr, "x"'] & \mathbf{y}C \arrow[d, dashed, "\bar{x}"] \\ & F \end{tikzcd}$$

Proof. Suppose F is a sheaf for j. By the universal mapping property for sheaves, it suffices to show that S is dense in $\mathbf{y}C$. Starting with the characteristic function for S in $\mathbf{y}C$, compute

$$\begin{aligned} \mathrm{ch}(S)_C(h : S \to S) &= \{f \mid (\exists C')(f : C' \to C \text{ and } (\mathbf{y}C)(f)(h) \in SC')\} \\ &= \{f \mid (\exists C')(f : C' \to C \text{ and } f^*(h) \in SC')\} \\ &= \{f \mid (\exists C')(f : C' \to C \text{ and } h \circ f \in SC')\} \\ &= \{f \mid h \circ f \in S\} \\ &= h * S. \end{aligned}$$

Hence,

$$\mathrm{ch}(\overline{S})_C(h) = j_C \circ \mathrm{ch}(S)_C(h) = j_C(h * S) = h * j_C(S) = \mathbf{y}C = \mathrm{true} \circ !(h),$$

so that $\overline{S} = \mathbf{y}C$.

In the converse direction, the goal is to associate each element $r \in XC$ with a natural transformation $\sigma^{(r)}$ from some covering sieve of C into F in such a way that if $r \in UC$, then the extended natural transformation $\overline{\sigma^{(r)}} : \mathbf{y}C \to F$ corresponds via the Yoneda lemma to the element $x_C(r)$.

To start, observe that

$$XC = \overline{U}C = \{r \mid j_C\,\mathrm{ch}(U)_C(r) = \mathrm{true}(r) = \mathbf{y}C\} = \{r \mid \mathrm{ch}(U)_C(r) \in JC\}$$

since U is dense in X. This makes $\mathrm{ch}(U)_C(r)$ a covering sieve for C, for any $r \in XC$. Next define the natural transformation $\sigma^{(r)} : \mathrm{ch}(U)_C(r) \to F$ given by

$$\sigma^{(r)}_{C'}(f : C' \to C) = x_{C'} \circ X(f)(r).$$

Verifying that $\sigma^{(r)}$ is a natural transformation is a routine calculation. By assumption, now, $\sigma^{(r)}$ extends to a unique natural transformation $\overline{\sigma^{(r)}} : \mathbf{y}C \to F$, which via the Yoneda lemma corresponds to a unique element $\bar{x}_C(r) \in FC$. This also shows that, if $\bar{x}$ is the desired natural transformation, then $\bar{x}$ is unique.

That defining $\bar{x} : X \to F$ this way produces the desired natural transformation can be seen as follows. To show naturality, recall the naturality of the isomorphism appearing in the Yoneda lemma, $\Phi_C : \mathrm{Hom}(\mathbf{y}C, F) \cong F$. For any $g : C' \to C$, one has

$$\begin{aligned} F(g) \circ \bar{x}_C(r) &= F(g) \circ \Phi_C\left(\overline{\sigma^{(r)}}\right) \\ &= \Phi_{C'} \circ \mathbf{y}(g)^* \left(\overline{\sigma^{(r)}}\right) \\ &= \Phi_{C'}\left(\overline{\sigma^{(r)}} \circ \mathbf{y}(g)\right), \end{aligned}$$

and also

$$\bar{x}_{C'} \circ X(g)(r) = \Phi_{C'}\left(\overline{\sigma^{(X(g)(r))}}\right).$$

Thus, to show naturality, it suffices to show that $\overline{\sigma^{(X(g)(r))}} = \overline{\sigma^{(r)}} \circ \mathbf{y}(g)$. This follows from the calculation

$$\sigma^{X(g)(r)}(h) = x \circ X(h) \circ X(g)(r) = x \circ X(gh)(r) = \sigma^{(r)}(gh) = \sigma^{(r)} \circ \mathbf{y}(g)(r).$$

The inverse to Φ_C is given by the isomorphism $\Psi_C : FC \to \mathrm{Hom}(\mathbf{y}C, F)$ defined by

$$\Psi_C(s)_{C'}(h : C' \to C) = F(h)(s).$$

Turning the definition of $\bar{x}$ inside-out,

$$\begin{aligned} \Psi_C(x_C(r))_{C'}(h) &= F(h) \circ x_C(r) \\ &= x_{C'} \circ U(h)(r) \\ &= x_{C'} \circ X(h)(r) \\ &= \sigma^{(r)}_{C'}(h) \\ &= \overline{\sigma^{(r)}_{C'}}(h). \end{aligned}$$

Hence,

$$x_C(r) = \Phi_C\left(\overline{\sigma^{(r)}_{C'}}(h)\right) = \bar{x}_C(r).$$

□

Presheaves that satisfy the latter condition for an arbitrary Grothendieck topology J are said to be *sheaves* (*for* J), and this is in fact a more common definition of the term.

The category of sheaves in $\widehat{\mathbb{C}}$ for a given Grothendieck topology J on $\mathbb{C}$ is denoted $\mathbf{Sh}(\mathbb{C}, J)$. In the case where J is given by the L-T modality j, the previous lemma says that $\mathbf{Sh}_j(\widehat{\mathbb{C}}) = \mathbf{Sh}(\mathbb{C}, J)$.

Consider again the $\neg\neg$ modality. A clear description of the covering sieves in its corresponding Grothendieck topology, called the *dense topology*, will play a role in the next section. In this document, the dense topology will only appear in the context of partially ordered sets, so it helps to understand at least these first.

Fix a partially ordered set $\mathbb{P}$. A sieve for an element p in $\mathbb{P}$ is a downward-closed subset of $\downarrow(p)$, so the subobject classifier Ω for $\widehat{\mathbb{P}}$ simply collects the downward closed subsets. As was already observed, the characteristic function of a monic $m : U \to F$ in $\widehat{\mathbb{P}}$ is a natural transformation $\mathrm{ch}(U) : F \to \Omega$ of the form

$$\mathrm{ch}(U)_p(s \in F(p)) = \{q \leqslant p \mid s{\restriction_q} \in U(p)\},$$

where $s{\restriction_q} = F(q \leqslant p)(s)$. Now, as one should verify for themself,

$$\Omega(q \leqslant p)(S \in \Omega(p)) = S{\restriction_q} = \{r \in S \mid r \leqslant q\},$$

and $\mathrm{false}_p(\varnothing) = \varnothing \in \Omega(p)$. This gives, for each monic $m : U \to F$, the identity

$$\neg_p(S \in \Omega(p)) = \mathrm{ch}(\mathrm{false})_p(S) = \{q \leqslant p \mid S{\restriction_q} \in \{\varnothing\}\} = \{q \leqslant p \mid \neg(\exists r \leqslant q)(r \in S)\}.$$

Thus, the double-negation modality is given by the identity

$$\neg\neg_p(S) = \{r \leqslant p \mid \{q \leqslant r \mid \neg(\exists r \leqslant q)(r \in S)\} = \varnothing\},$$

or in other words,

$$q \in \neg\neg_p(S) \text{ if and only if } q \leqslant p \text{ and } (\forall r \leqslant q)(\exists s \leqslant r)(s \in S).$$

By definition, a sieve S covers p if and only if $\neg\neg_p(S) = \downarrow(p)$, which is the case if and only if

$$(\forall r \leqslant p)(\exists s \leqslant r)(s \in S).$$

One calls a sieve S on p with this property *dense below* p. Letting $\neg\neg$ be the Grothendieck topology given by

$$\neg\neg(p) = \{S \subseteq \downarrow(p) \mid S \text{ is dense below } p\},$$

one has $\mathbf{Sh}_{\neg\neg}(\widehat{\mathbb{C}}) = \mathbf{Sh}(\mathbb{P}, \neg\neg)$.

4.3 Sheaves form a topos

The next goal is prove the following reslt.

Theorem 4.6. *For any L-T modality $j : \Omega \to \Omega$ in the topos $\mathbb{E} = \widehat{\mathbb{C}}$, $\mathbf{Sh}_j(\mathbb{E})$ is a topos. Equivalently, if J is the corresponding Grothendieck topology, $\mathbf{Sh}(\mathbb{C}, J)$ is a topos.*

This will be done in a pair of lemmas.

Lemma 4.7. *The category $\mathbf{Sh}(\mathbb{C}, J)$ is complete.*

Proof. Let $P : \mathbb{I} \to \mathbf{Sh}(\mathbb{C}, J)$ be a functor from a small category. It follows from the fact that the inclusion functor $i : \mathbf{Sh}(\mathbb{C}, J) \to \mathbb{E}$ is full and faithful that, if $L \cong \varinjlim P$, then

$$i(L) \cong \varinjlim(i \circ P).$$

Hence, it suffices to show that if for every $I \in \mathbb{I}$ the functor PI is a sheaf, then L is a sheaf as well. This follows from the following computation: Let U be a dense subobject of X in $\mathbf{Sh}_j(\mathbb{E})$. Then, if $L = \varinjlim(i \circ P)$,

$$\begin{aligned}
\mathrm{Hom}_{\mathbb{E}}(U, L) &= \mathrm{Hom}(U, \varinjlim(i \circ P)) \\
&\cong \varinjlim \mathrm{Hom}(U, i \circ P) \\
&\cong \mathrm{Hom}(X, \varinjlim(i \circ P)) \\
&= \mathrm{Hom}(X, L).
\end{aligned}$$

Hence, arrows of the form $U \to L$ extend to arrows $X \to L$. This makes L a sheaf. □

Notice that the above argument only used the completeness of $\widehat{\mathbb{C}}$. In other words, the same argument shows that $\mathbf{Sh}_j(\mathbb{E})$ has finite limits for an *arbitrary* topos $\mathbb{E}$.

Lemma 4.8. *For any sheaves P and F in $\mathbb{E}$, the object P^F in $\mathbb{E}$ is a sheaf. Moreover, P^F is the exponential of P by F in $\mathbf{Sh}_j(\mathbb{E})$.*

Proof. Since $U \to X$ is dense,

$$\begin{aligned}
\mathrm{ch}(\overline{U \times F} \subseteq X \times F) &= j \circ \mathrm{ch}(U \times F) \\
&= j \circ \mathrm{ch}(U) \circ \pi_1 \\
&= \mathrm{ch}(\mathrm{id}_X) \circ \pi_1 \\
&= \mathrm{true}_{X \times F},
\end{aligned}$$

so that $U \times F \to X \times F$ is dense. Hence,

$$\mathrm{Hom}_{\mathbb{E}}(U, P^F) \cong \mathrm{Hom}(U \times F, P) \cong \mathrm{Hom}(X \times F, P) \cong \mathrm{Hom}(X, P^F),$$

making P^F a sheaf as well. Since i is full and faithful, i preserves exponentials. $\square$

Again, no mention of J or of $\mathbb{C}$ actually entered into the above argument, so $\mathbf{Sh}_j(\mathbb{E})$ has exponentials as well.

In view of the previous two lemmas, $\mathbf{Sh}_j(\mathbb{E})$ is a topos for an arbitrary topos $\mathbb{E}$. In particular, $\mathbf{Sh}(\mathbb{C}, \neg\neg)$ is a complete topos whose exponents match those of the surrounding category $\widehat{\mathbb{C}}$. The subobject classifier, on the other hand, does not match.

Lemma 4.9. *Let $P \subseteq F$ in $\mathbb{E}$, and assume F is a sheaf. Then P is a sheaf if and only if P is closed.*

Proof. Suppose $m : P \to F$ is a closed monic, $U \subseteq X$ is a dense subobject, and $x : U \to P$ is any arrow. Since F is a sheaf, there is a unique $y : X \to F$ such that

$$\begin{tikzcd} X \arrow[r, "y"] & F \\ U \arrow[u, "\subseteq"] \arrow[r, "x"] & P \arrow[u, "m"'] \end{tikzcd}$$

commutes. Since the image function $\exists_y : \mathrm{Sub}(X) \to \mathrm{Sub}(F)$ is left adjoint to $y^{-1} : \mathrm{Sub}(F) \to \mathrm{Sub}(X)$, $U \subseteq y^{-1}(x\text{"}U)$. Of course, the image $x\text{"}U$ of U under x is a subobject of P, so actually $U \subseteq y^{-1}(P)$ by monotonicity of y^{-1}. The closure operation is also monotone, so

$$X = \overline{U} \subseteq \overline{y^{-1}(P)} = y^{-1}(\overline{P}) = y^{-1}(P).$$

This implies that y actually factors through m, as $X = y^{-1}(P)$ implies

$$\mathrm{true}_X = \mathrm{ch}(y^{-1}(P)) \circ \mathrm{id}_X = \mathrm{ch}(P) \circ y.$$

The factorization through m provides the desired $\bar{x} : X \to P$.

Conversely, assume P is a sheaf. Then, since P is dense in $\overline{P}$, the monic $\mathrm{id}_P : P \to P$ extends to a unique arrow $\overline{P} \to P$ through which $P \to \overline{P}$ factors. This arrow, $\overline{P} \to P$, is necessarily monic, so $P = \overline{P}$. $\square$

Thus, the subobject classifier of $\mathbf{Sh}_j(\mathbb{E})$ is the object in $\mathbb{E}$ that classifies the closed subobjects of sheaves. A closed subobject $P \subseteq F$ in $\mathbb{E}$ is precisely a monic for which the following diagram commutes.

$$\begin{tikzcd} F \arrow[r, "\mathrm{ch}(P)"] \arrow[dr, "\mathrm{ch}(P)"'] & \Omega \arrow[d, shift right, "\mathrm{id}_\Omega"'] \arrow[d, shift left, "j"] \\ & \Omega \end{tikzcd}$$

It is true, of course, that any arrow with this property factors uniquely through the equalizer of j and id_Ω. Let Ω_j be such an object, and observe that since $j \circ \mathrm{true} = \mathrm{id}_\Omega \circ \mathrm{true}$, true factors through Ω_j as well. The last statement corresponds to the maximal sieve's role as a covering sieve.

To see that Ω_j is indeed a subobject classifier for $\mathbf{Sh}_j(\mathbb{E})$ requires the following observation.

Lemma 4.10. *Let $m : U \to X$ be any dense monic. The map $\overline{(-)} \circ \exists_m$ gives a natural bijective correspondance between the closed subobjects of U and the closed subobjects of X, written*

$$\mathrm{ClSub}(U) \cong \mathrm{ClSub}(X).$$

Proof. Let the dense monic $m_U : U \to X$ represent $U \subseteq X$. Making the action of the map explicit, the forward direction of the correspondence is just the closure operator: Since $U \subseteq X$, every subobject of U is also a subobject of X. Taking the closure of a subobject of U in X gives,

$$V \subseteq U \longmapsto \overline{V} \subseteq X.$$

In the reverse direction, one takes a closed subobject $A \subseteq X$ and intersects it with U, giving the reverse map

$$U \wedge A \longleftarrow A.$$

Note that this intersection is precisely the map m^{-1} since m is monic. To see that this is well-defined, recall that the closure of U in U is simply U, and furthermore that the closure of A is A, so that

$$\begin{aligned} \mathrm{ch}_U\left(\overline{U \wedge A}\right) &= j \circ \mathrm{ch}_U(U \wedge A) \\ &= j \circ \mathrm{ch}_X(U \wedge A) \circ m_U \\ &= (j\,\mathrm{ch}_X(U) \wedge j\,\mathrm{ch}_X(A)) \circ m_U \\ &= (\mathrm{ch}_X(X) \wedge \mathrm{ch}_X(A)) \circ m_U \\ &= \mathrm{ch}_U(U \wedge A). \end{aligned}$$

Backward and then forward, one has

$$\overline{U \wedge A} = \overline{U} \wedge \overline{A} = X \wedge A = A,$$

where the closure is taken in X. Forward and then backward, the computation

$$\begin{aligned}
\mathrm{ch}_U(U \wedge \overline{V}) &= \mathrm{ch}_X(U \wedge \overline{V}) \circ m_U \\
&= (\mathrm{ch}_X(U) \circ m_U) \wedge (j \circ \mathrm{ch}_X(V) \circ m_U) \\
&= (\mathrm{ch}_U(U)) \wedge (j\, \mathrm{ch}_U(V)) \\
&= \mathrm{ch}(V)
\end{aligned}$$

reveals that

$$U \wedge \overline{V} = V,$$

again where the closure is taken in X. This shows that the two maps are inverse to one another, and are therefore bijections. $\square$

As was already pointed out, the universal mapping property for equalizers gives a natural isomorphism

$$\mathrm{ClSub}(U) \cong \mathrm{Hom}(U, \Omega_j).$$

Thus, an immediate corollary of the preceeding lemma is

$$\mathrm{Hom}(U, \Omega_j) \cong \mathrm{ClSub}(U) \cong \mathrm{ClSub}(X) \cong \mathrm{Hom}(X, \Omega_j)$$

for any closed subobject $U \subseteq X$. This proves that Ω_j is a sheaf. Since Ω_j classifies closed subobjects, Ω_j classifies *subsheaves*, subobjects in $\mathbf{Sh}_j(\mathbb{E})$. This completes the proof that $\mathbf{Sh}_j(\mathbb{E})$ is a topos with the subobject classifier Ω_j, for arbitrary $\mathbb{E}$. The next step, however, is to understand how the various features of $\mathbf{Sh}_j(\mathbb{E})$ manifest in $\mathbf{Sh}(\mathbb{P}, \neg\neg)$.

Limits in presheaf categories are computed pointwise, equalizers included. Thus, for each object C of $\mathbb{C}$,

$$\Omega_j C = \{S \in \Omega C \mid j_C(S) = \mathrm{id}_{\Omega C}(S)\} = \{S \in \Omega C \mid j_C(S) = S\}.$$

In other words, $\Omega_j(C)$ consists of the *closed* sieves on C.

In the particular case where $j = \neg\neg$, the set $\Omega_{\neg\neg} C$ is then the set of sieves on C for which $\neg\neg S = S$. Generally, for any Heyting algebra H, the subalgebra $\{b \mid \neg\neg b = b\}$ in H is a Boolean subalgebra of H, and this applies to $\Omega_{\neg\neg}$ equally well. In other words, $\mathbf{Sh}(\mathbb{P}, \neg\neg)$ is a Boolean topos! In fact, $\mathbf{Sh}(\mathbb{P}, \neg\neg)$ is notably similar to a Boolean-valued model.

Theorem 4.11. *Let $\mathbb{P}$ be a poset. The category of double-negation sheaves $\mathbf{Sh}(\mathbb{P}, \neg\neg)$ on $\mathbb{P}$ is a Boolean topos with a natural numbers object.*

The second statement in the theorem follows from Lemma 2.17 and the following theorem.

Theorem 4.12. *For any topos $\mathbb{E}$ and any L-T modality $j : \Omega \to \Omega$ for $\mathbb{E}$, the inclusion of $\mathbf{Sh}_j(\mathbb{E})$ into $\mathbb{E}$ has a left adjoint*

$$\mathbf{a} : \mathbb{E} \leftrightarrows \mathbf{Sh}_j(\mathbb{E}) : i.$$

Furthermore, $\mathbf{a}$ preserves finite limits.

The classical case of the theorem, as in for $\mathbb{E} = \widehat{\mathbb{C}}$, is proven using a two-step process called the *double-plus* construction. For a general topos, Lawvere outlines the process in [17] as follows: Given an object A of $\mathbb{E}$, first form the image of the composition

$$A \xrightarrow{\{\cdot\}} \Omega^A \longrightarrow \Omega_j^A,$$

(see the discussion immediately following Definition 5.1 for $\{\cdot\}$) where the latter map is given by the exponential of the the arrow $\Omega \to \Omega_j$ obtained from factoring j. Call the resulting subobject A^+. Then, thinking of A^+ as a subobject of Ω_j^A, define $\mathbf{a}(A) = \overline{A^+}$. The unit of the adjunction is then given by the composition $A \to A^+ \subseteq \mathbf{a}(A)$.

5 Independence of the Continuum Hypothesis

In the previous section, $\mathbf{Sh}(\mathbb{P}, \neg\neg)$ was shown to be a Boolean topos with a natural numbers object. In the following section, various properties of $\mathbb{P}$ will be shown to correspond to properties of $\mathbf{Sh}(\mathbb{P}, \neg\neg)$, in the usual forcing argument fashion. The driving motivation for this endeavour will be a proof of the independence of the continuum hypothesis, recorded in the current section, which will roughly follow the beautiful exposition in chapter VI of [13]. The continuum hypothesis is stated for toposes as follows.

Definition 5.1. Let $\mathbb{E}$ be a topos with a natural numbers object N. Then $\mathbb{E}$ is said to *violate the continuum hypothesis*, or *satisfy $\neg CH$*, if there is an object A in $\mathbb{E}$ such that $N \subseteq A \subseteq \Omega^N$ and there are no epimorphisms in $\mathbb{E}$ of the form $A \to \Omega^N$ or $N \to A$.

5.1 Inequalitities in toposes

For any object A in a topos, there is a canonical monic arrow $\{\cdot\} : A \to \Omega^A$ induced by the arrow $\mathrm{true}_{A\times A}$, as in

$$\begin{tikzcd} A \times \Omega^A \arrow[rr, "\in"] & & \Omega \\ A \times A \arrow[u, "\mathrm{id}_A \times \{\cdot\}"] \arrow[urr, "\mathrm{true}_{A\times A}"] & & \end{tikzcd}$$

Here, the arrow $\in$ is the evaluation map for the exponential Ω^A. To understand the notation, observe that in $\mathbb{S}$**ets**, $\in$ is the usual membership relation and $\{\cdot\}$ is the function $a \mapsto \{a\} \subseteq A$. This shows that $N \subseteq \Omega^N$. To see that a strict inequality holds (in fact to show that this sort of strict containment holds in general) it is convenient to introduce the following construction, which can be performed in any topos $\mathbb{E}$.

For any $A \subseteq B$ in $\mathbb{E}$, let $\mathrm{Epi}(A, B)$ denote the class of epimorphisms of the form $A \to B$. The key property of $\mathrm{Epi}(A, B)$, as an object of $\mathbb{S}$**ets**, is that $\mathrm{Epi}(A, B) \cong 0$ entails $A < B$. However, $\mathrm{Epi}(A, B)$ is not often an object $\mathbb{E}$ knows much about, in the sense that $\mathrm{Epi}(A, B)$ is not an object of the category. On the other hand, one is able to construct an analogous subobject $\mathrm{epi}(A, B) \subseteq B^A$ in $\mathbb{E}$ with the property that $\mathrm{epi}(A, B) \cong 0$ implies $\mathrm{Epi}(A, B) = \varnothing$. Toward this end, define

$$\mathcal{I}_X : \mathrm{Hom}(X, B^A) \to \mathrm{Hom}(X, \Omega^B)$$

by setting

$$\mathcal{I}_X(f) = \mathrm{tr}^B \circ \mathrm{ch} \circ \mathrm{img}(\mathrm{tr}_A(f), \pi_2),$$

where $\mathrm{tr}_A : \mathrm{Hom}(X, B^A) \cong \mathrm{Hom}(A \times X, B)$ and $\mathrm{tr}^B : \mathrm{Hom}(B \times X, \Omega) \cong \mathrm{Hom}(X, \Omega^B)$ and $\mathrm{img}(\pi_1, \mathrm{tr}_A(f))$ is the image of the map $(\mathrm{tr}_A(f), \pi_2)$. A diagrammatic expression of the map $\mathcal{I}_X$ can be given, as well:

$$\begin{tikzcd} \mathrm{Hom}(A \times X, B) \arrow[r, "(-,\pi_2)"] & \mathrm{Hom}(A \times X, B \times X) \arrow[r, "\mathrm{img}"] & \mathrm{Sub}(B \times X) \arrow[d, "\cong"] \\ \mathrm{Hom}(X, B^A) \arrow[u, "\cong"] \arrow[r, "\mathcal{I}"] & \mathrm{Hom}(X, \Omega^B) & \mathrm{Hom}(B \times X, \Omega) \arrow[l, "\cong"'] \end{tikzcd}$$

Given either definition of the natural transformation $\mathcal{I}$, the Yoneda lemma provides a unique arrow $\ell : B^A \to \Omega^B$ such that $\mathcal{I} = \ell_*$. In $\mathbb{S}$**ets**, $\mathcal{I}_X(f)$ is the function

$$x \longmapsto \{b \in B \mid (\exists a \in A)(f(x)(a) = b)\},$$

and the function ℓ is the unique map such that

$$\ell \circ f = \mathcal{I}_X(f).$$

In the case where $X = 1$, an arrow $f : 1 \to B^A$ is given uniquely by the element $y = f(\varnothing)$ of B^A. It follows that ℓ is the map

$$(y : A \to B) \longmapsto \{b \in B \mid (\exists a \in A)(y(a) = b)\} = \mathrm{img}(y).$$

Now define $\mathrm{epi}(A, B)$ to be the pullback of ℓ and $\lceil B \rceil$, where

$$\mathrm{tr}(\lceil B \rceil) = \mathrm{true}_B : B \cong B \times 1 \to \Omega,$$

as in the pullback diagram

$$\begin{tikzcd}
B^A \arrow[r, "\ell"] & \Omega^B \\
\mathrm{epi}(A,B) \arrow[u] \arrow[r, "!"] & 1 \arrow[u, "\lceil B \rceil"']
\end{tikzcd}$$

In $\mathbb{S}\mathbf{ets}$, $\lceil B \rceil(\varnothing) = B \in 2^B$, so the pullback of ℓ and $\lceil B \rceil$ is precisely the set $\mathrm{Epi}(A, B)$ of functions $y : A \to B$ such that $\mathrm{img}(y) = \ell(y) = \lceil B \rceil \circ !(y) = B$. Note also that gluing pullback diagrams together yeilds

$$\begin{tikzcd}
B^A \arrow[r, "\ell"] & \Omega^B \arrow[r, "\mathrm{ch}(\lceil B \rceil)"] & \Omega \\
\mathrm{epi}(A,B) \arrow[u] \arrow[r, "!"] & 1 \arrow[u, "\lceil B \rceil"'] \arrow[r] & 1 \arrow[u, "\mathrm{true}"']
\end{tikzcd}$$

The outer rectangle is then a pullback square, so

$$\mathrm{ch}(\mathrm{epi}(A, B)) = \mathrm{ch}(\lceil B \rceil) \circ \ell.$$

The following calculation provides the key property of $\mathrm{epi}(A, B)$, that it provides a sort of classifier for "parameterized families" of epic arrows: For any $h : X \to B^A$,

$$\begin{aligned}
\mathrm{tr}_A(h) : A \times X \to B \text{ is epic } & \text{ if and only if } (\mathrm{tr}(h), \pi_2) \text{ is epic}, \\
& \text{ if and only if } \mathrm{img}(\mathrm{tr}(h), \pi_2) = B \times X, \\
& \text{ if and only if } \mathrm{ch}(\mathrm{img}(\mathrm{tr}(h), \pi_2)) = \mathrm{true}_{B \times X}, \\
& \text{ if and only if } \mathrm{tr}^B(\mathrm{ch}(\mathrm{img}(\mathrm{tr}(h), \pi_2))) = \lceil B \rceil \circ !_X.
\end{aligned}$$

In other words, an arrow $h : X \to B^A$ factors through $\mathrm{epi}(A, B)$ if and only if $\mathrm{tr}_A(h)$ is epic. This has as a special case one of the intuitive properties of $\mathrm{Epi}(A, B)$.

Lemma 5.2. *In a topos $\mathbb{E}$ with objects A and B, if $1 \not\cong 0$ and $\mathrm{epi}(A, B) \cong 0$, then $\mathrm{Epi}(A, B) = \varnothing$.*

It is necessary to first observe that any arrow $1 \to 0$ is an isomorphism[10]. Indeed, $0 \to 1$ and $0 \to 0$ are unique, so $0 \to 1 \to 0$ is the identity arrow for 0. And in reverse, the arrow $1 \to 1$ is unique, so $1 \to 0 \to 1$ is the identity.

Any topos with the property $1 \not\cong 0$ is called *nondegenerate*. Of course, 1 and 0 in $\mathbf{Sh}(\mathbb{P}, \neg\neg)$ are not isomorphic, so $\mathbf{Sh}(\mathbb{P}, \neg\neg)$ is nondegenerate.

Proof. Suppose $\mathrm{epi}(A, B) \cong 0$. It suffices to show that $\mathrm{Epi}(A \times 1, B) = \varnothing$. Since an epic arrow $h : A \times 1 \to B$ induces an arrow $f = \mathrm{tr}^A(h) : 1 \to B^A$ that factors through $\mathrm{epi}(A, B)$, f factors through 0. However, $\mathrm{dom}(f) = 1$, so this would provide an arrow $1 \to 0$. By assumption, $1 \not\cong 0$, so no such epic exists. $\square$

In a nondegenerate topos, define the relation $<$ on objects of $\mathbb{E}$ so that $A < B$ if and only if $A \subseteq B$ and $\mathrm{epi}(A, B) \cong 0$. An even stronger violation of the continuum hypothesis might then be the condition that there is an object A of $\mathbb{E}$ such that $N < A < \Omega^N$. Observe that this is indeed a stonger condition by Lemma 5.2.

Lemma 5.3. *Recall the functor $\mathbf{a}\Delta : \mathbb{S}\mathbf{ets} \to \mathbf{Sh}(\mathbb{P}, \neg\neg)$. For any sets A and B, $\mathrm{Epi}(A, B) = \varnothing$ if $\mathrm{epi}(\mathbf{a}\Delta A, \mathbf{a}\Delta B) \cong 0$.*

Proof. By the previous lemma, $\mathrm{epi}(\mathbf{a}\Delta A, \mathbf{a}\Delta B) \cong 0$ implies $\mathrm{Epi}(\mathbf{a}\Delta A, \mathbf{a}\Delta B) = \varnothing$. Thus, the desired equation follows from the observation that $\mathbf{a}$ preserves colimits and therefore epimorphisms, as well as the observation that identifying functions $A \to B$ with natural transformations $\Delta A \to \Delta B$ yields $\mathrm{Epi}(A, B) = \mathrm{Epi}(\Delta A, \Delta B)$. In other words, if $\mathrm{Epi}(\mathbf{a}\Delta A, \mathbf{a}\Delta B) = \varnothing$, then $\mathrm{Epi}(A, B) = \varnothing$. $\square$

This is nice, but the converse is really the useful result. The converse only actually holds under certain circumstances, and will play into how one should choose the correct poset for violating CH. Before this poset is chosen, however, a few lemmas are needed.

Lemma 5.4. *Let $h : A \to B$ be an epic arrow in $\mathbb{E}$, and X be any object. The eponentiated arrow $h^X : A^X \to B^X$ restricts to an arrow $\mathrm{epi}(X, A) \to \mathrm{epi}(X, B)$.*

Proof. Let $m : \mathrm{epi}(X, A) \subseteq X^X$. Recall that h^X is defined to be the unique arrow for

[10] In fact, a stronger statement holds: Any arrow of the form $A \to 0$ is an isomorphism in a topos. Recall Theorem 3.11.(4).

which the diagram

$$\begin{tikzcd} X \times B^X \arrow[r, "\varepsilon_B"] & Z \\ X \times A^X \arrow[u, "\mathrm{id}_X \times h^X"] \arrow[r, "\varepsilon_Y"] & A \arrow[u, "h"'] \end{tikzcd}$$

commutes. Gluing this diagram to the triangle

$$\begin{tikzcd} X \times A^X \arrow[r, "\varepsilon_Y"] & A \\ X \times \mathrm{epi}(X, A) \arrow[u, "\mathrm{id}_X \times m"] \arrow[ur, "\mathrm{tr}_X(m)"'] & \end{tikzcd}$$

reveals the equation

$$\mathrm{tr}_X(h^X m) = h\varepsilon_A(\mathrm{id} \times m) = h \circ \mathrm{tr}_X(m).$$

Since m trivially factors through $\mathrm{epi}(X, A)$, $\mathrm{tr}_X(m)$ is epic, and therefore so is $h \circ \mathrm{tr}_X(m)$. This implies that $\mathrm{tr}_X(h^X m)$ is epic, or equivalently that $h^X \circ m$ factors through $\mathrm{epi}(X, B)$. □

Lemma 5.5. *In a Boolean topos, if $A < B \leqslant C$ and there is at least one arrow $1 \to B$, then $A < C$.*

Proof. Write $m : B \subseteq C$. Since the ambient topos is Boolean, $B \vee \neg B = C$ in $\mathrm{Sub}(C)$. Equating C with $B \vee \neg B$ generally, one obtains an arrow $r : C \to Z$ via the universal mapping property for coproducts, as in

$$\begin{tikzcd} & Z & \\ Z \arrow[ur, "\mathrm{id}"] \arrow[r, "m"] & C \arrow[u, "r"] & \neg Z \arrow[l] \arrow[ul, "x \circ !"'] \end{tikzcd}$$

Since $rm = \mathrm{id}$, r is epic. By the previous lemma, $r^A : B^A \to C^A$ restricts to an arrow of the form

$$\mathrm{epi}(A, C) \to \mathrm{epi}(A, B).$$

It follows from Theorem 3.11.(4) that since $\mathrm{epi}(A, B) \cong 0$, one has $\mathrm{epi}(A, C) \cong 0$ as well. □

It has already been shown that $\mathbf{Sh}(\mathbb{P}, \neg\neg)$ is a Boolean topos. It is not true, however, that a nonzero sheaf admits an arrow from 1: It could be the case that a nonzero sheaf is only at some indices nonempty. This is fortunately not so for *locally constant* sheaves, sheaves of the form $\mathbf{a}\Delta(X)$ for some set X, for a nonempty set X gives rise to a nonzero global element $1 \to \Delta(X)$ that is sent via $\mathbf{a}$ to a global element $1 \to \mathbf{a}\Delta(X)$.

5.2 The Cohen Topos

The rest of the desired properties of $\mathbf{Sh}(\mathbb{P}, \neg\neg)$ follow from specific combinatorial properties of $\mathbb{P}$. Thus, to pick the correct $\mathbb{P}$ finally involves a little inspiration from Cohen. Let κ be a cardinal strictly larger than the continuum $2^{\mathbb{N}}$. Recall the *Cohen poset* (*for* κ), the set of functions

$$\mathbb{P}_\kappa = \{f \mid (\exists A \subseteq \kappa \times \mathbb{N})(|A| < \mathbb{N} \text{ and } f : A \to 2)\},$$

ordered by reverse inclusion, ie. $f \leqslant g$ in $\mathbb{P}_\kappa$ if and only if $g \subseteq f$, or equivalently

$$\mathrm{dom}(g) \subseteq \mathrm{dom}(f) \text{ and } f\restriction_{\mathrm{dom}(g)} = g.$$

The *Cohen topos* (*for* κ) is then defined to be the topos $\mathbf{Sh}(\mathbb{P}_\kappa, \neg\neg)$.

Theorem 5.6. *The Cohen topos satisfies the following properties.*

(a) *The Cohen topos is a Boolean topos.*

(b) *The object* $\mathbf{a}\Delta(\mathbb{N})$ *is a natural numbers object in the Cohen topos.*

(c) *The Cohen topos is generated by the representable presheaves, ie. for any pair of arrows* $f, g : F \rightrightarrows G$, $f \neq g$ *if and only if there is a* $p \in \mathbb{P}$ *and an arrow* $x : \mathbf{y}p \to F$ *such that* $fx \neq gx$. *In fancier words, the* points *from* representables seperate arrows.

(d) *In the Cohen topos for* κ,

$$\mathbf{a}\Delta(\mathbb{N}) < \mathbf{a}\Delta(2^{\mathbb{N}}) < \Omega_{\neg\neg}^{\mathbf{a}\Delta(\mathbb{N})}.$$

Properties (a) and (b) have already been established. The objective will soon be to prove (d), but toward this end (c) is actually a necessary step.

In order to see that (c) holds, recall that $\widehat{\mathbb{P}}$ is the cocompletion of $\mathbb{C}$, or equivalently, every presheaf is a colimit of representable presheaves, meaning that every arrow between presheaves is determined by a morphism of cones of representable presheaves. It should be clear that this statement implies that representable presheaves always generate their presheaf categories. Hence, in order to see that (c) holds, it suffices to show the following.

Lemma 5.7. *Every representable presheaf is a sheaf in the Cohen topos.*

Proof. Let S be a sieve dense below $g \in \mathbb{P}_\kappa$, let $f \in \mathbb{P}_\kappa$, and let $S \to \mathbf{y}(f)$ be a natural transformation. It suffices to show that $\downarrow(g) \subseteq \downarrow(f)$, or equivalently $g \leqslant f$. Since $S(h) \to \mathbf{y}(f)h = \mathrm{Hom}(h, f)$ is a function in $\mathbb{S}\mathbf{ets}$ for every $h \in \mathbb{P}_\kappa$, $S(h) \neq \varnothing$ implies $h \leqslant f$ for every h. This says that $S \subseteq \mathbf{y}f = \downarrow(f)$. To see that $g \leqslant f$, let $(\alpha, n) \in \mathrm{dom}(f)$ and suppose for a contradiction that there were an extension $g' \leqslant g$ such that $g'(\alpha, n) \neq f(\alpha, n)$. Then, by density of S, there is a $h \in S$ extending g' and by virtue of being in S also extends f. This would give $g'(\alpha, n) = h(\alpha, n) = f(\alpha, n)$, which is impossible. $\square$

Observe that the category $\widehat{\mathbb{P}_\kappa}$ contains every possible piecing-together of a partial function whose finite approximations exist in $\mathbb{P}$. That's a lot! And they are not interrelated in any controlled way. To lasso the pieces that make up functions whose "domain" is the "entire set" $\kappa \times \mathbb{N}$ (really, these will be arrows from the locally constant sheaf at $\kappa \times \mathbb{N}$) is essentially the idea behind isolating the double-negation sheaves in $\widehat{\mathbb{P}_\kappa}$, since the double-negation topology demands density from its sieves.

The next question is how to obtain a monic arrow $\mathbf{a}\Delta(\kappa) \to \Omega_{\neg\neg}^{\mathbf{a}\Delta(\mathbb{N})}$ from $\mathbb{P}$. This essentially consists of a locally consistent choice of subsets of $\kappa \times \mathbb{N}$, for which there are really just *two* choices. Following [13], set

$$A(f) = \{(\alpha, n) \in \mathrm{dom}(f) \mid f(\alpha, n) = 0\}.$$

If $f \leqslant g$, $A(g) \subseteq A(f)$, so indeed $A \subseteq \Delta(\kappa \times \mathbb{N}) = \Delta(\kappa) \times \Delta(\mathbb{N})$ as a functor. This immediately induces the characteristic arrow $\mathrm{ch}(A) : \Delta(\kappa) \times \Delta(\mathbb{N}) \to \Omega$.

Lemma 5.8. *The subobject $A \subseteq \Delta(\kappa \times \mathbb{N})$ is double-negation closed.*

Proof. It has already been shown that $A \subseteq \neg\neg A$. To prove the reverse containment, let $(\alpha, n) \notin A(f)$ for some $f \in \mathbb{P}_\kappa$. Then either $(\alpha, n) \notin \mathrm{dom}(f)$ or $f(\alpha, n) = 1$. In either case, one can easily extend f to a function f' such that $f'(\alpha, n) = 1$. In other words, $(\alpha, n) \notin \neg\neg A(f)$. $\square$

Now, since $\Omega_{\neg\neg}$ classifies closed subobjects, $\mathrm{ch}(A) : \Delta(\kappa) \times \Delta(\mathbb{N}) \to \Omega$ factors through $\Omega_{\neg\neg}$. Define $\gamma = \mathrm{tr}(\mathrm{ch}(A)) : \Delta(\kappa) \to \Omega_{\neg\neg}^{\Delta(\mathbb{N})}$.

Lemma 5.9. *The arrow γ is monic.*

Proof. Given $\alpha \neq \beta$ in κ, the claim is that

$$\gamma_f(\alpha) \neq \gamma_f(\beta) \in \left(\Omega_{\neg\neg}^{\Delta(\mathbb{N})}\right)(f) = \mathrm{Hom}(\mathbf{y}f \times \Delta(\mathbb{N}), \Omega_{\neg\neg}).$$

The explicit calculation of $\gamma_f(-)$ goes as follows: For any $g \leqslant f$ and $n \in \mathbb{N}$,

$$\gamma_f(\alpha)(g,n) = \varepsilon_f(\mathrm{id}_{\mathbb{N}} \times \gamma_f)(n,\alpha) = \mathrm{ch}(A)_f(\alpha,n) = \{q \leqslant f \mid q(\alpha,n) = 0\}.$$

Hence, since the domain of f is infinite, there is an $n_0 \in \mathbb{N}$ such that $(\alpha, n_0), (\beta, n_0) \notin \mathrm{dom}(f)$. This allows for an extension f' of f such that $f'(\alpha, n_0) \neq f'(\beta, n_0)$, implying $\gamma_f(\alpha) \neq \gamma_f(\beta)$. $\square$

Since $\mathbf{a}$ preserves finite limits, $\mathbf{a}$ preserves monics. Thus, the previous lemma implies that $\mathbf{a}\Delta(\kappa) \subseteq \mathbf{a}\left(\Omega_{\neg\neg}^{\Delta(\mathbb{N})}\right)$. In fact, the latter object is precisely the "continuum" in the Cohen topos: Observe that the Yoneda lemma in conjunction with the calculation

$$\begin{aligned}
\mathrm{Hom}(X, \Omega_{\neg\neg}^{\mathbf{a}\Delta(\mathbb{N})}) &\cong \mathrm{Hom}(\mathbf{a}(X), \Omega_{\neg\neg}^{\mathbf{a}\Delta(\mathbb{N})}) && (\mathbf{a} \dashv i) \\
&\cong \mathrm{Hom}(\mathbf{a}\Delta(\mathbb{N}) \times \mathbf{a}(X), \Omega_{\neg\neg}) && \text{(product/exponential adjunction)} \\
&\cong \mathrm{Hom}(\mathbf{a}(\Delta(\mathbb{N}) \times X), \Omega_{\neg\neg}) && (\mathbf{a} \text{ preserves } \times) \\
&\cong \mathrm{Hom}(\Delta(\mathbb{N}) \times X, \Omega_{\neg\neg}) && (\Omega_{\neg\neg} \text{ a sheaf}, \mathbf{a} \dashv i) \\
&\cong \mathrm{Hom}(X, \Omega_{\neg\neg}^{\Delta(\mathbb{N})}) && \text{(product/exponential adjunction)}
\end{aligned}$$

for an arbitrary presheaf X implies that $\Omega_{\neg\neg}^{\Delta} \cong \mathbf{a}\left(\Omega_{\neg\neg}^{\Delta(\mathbb{N})}\right)$. This says that $\Omega_{\neg\neg}^{\Delta(\mathbb{N})}$ is a sheaf, and therefore

$$\mathbf{a}\left(\Omega_{\neg\neg}^{\Delta(\mathbb{N})}\right) \cong \Omega_{\neg\neg}^{\mathbf{a}\Delta(\mathbb{N})}.$$

Hence,

$$\mathbf{a}\Delta(\kappa) \subseteq \Omega_{\neg\neg}^{\mathbf{a}\Delta(\mathbb{N})}.$$

The special combinatorial property of the Cohen poset that plays a role in the grand finale of this section is the *countable antichain condition*. For posets, the *countable antichain condition* says that any pairwise incompatible family of points is at most countable. The statement for general categories is as follows.

Definition 5.10. A *c.a.c. object* (or *object with the countable antichain condition*) in a category is a nonzero object A for which $\mathrm{Sub}(A)$ has the countable antichain condition. A *c.a.c. category* is a category generated by a family of c.a.c. objects.

As has already been observed, the Cohen poset is generated by the representable objects. A classical and straighforward excercise reveals that the Cohen poset satisfies the countable antichain condition, and therefore so does $\downarrow(f)$ for each $f \in \mathbb{P}_\kappa$. This implies that the poset of closed subobjects $\mathrm{ClSub}(\downarrow(f))$ has the countable antichain condition as well, making $\mathbf{y}f$ a c.a.c. object, and by extension $\mathbf{Sh}(\mathbb{P}_\kappa, \neg\neg)$ a c.a.c. category. The following lemma is therefore the last step in the proof of Theorem 5.6.(d).

Lemma 5.11. *In any c.a.c. topos of the form* $\mathbf{Sh}(\mathbb{C}, J)$*, given infinite sets* X *and* Y*, one has* $\mathrm{Epi}(X, Y) = \varnothing$ *if and only if* $\mathrm{epi}(\mathbf{a}\Delta(X), \mathbf{a}\Delta(Y)) \cong 0$.

Proof. The reverse direction of this is a special case of Lemma 5.3. To see the forward direction, consider its contrapositive: Write $E = \mathrm{epi}(\mathbf{a}\Delta(X), \mathbf{a}\Delta(Y))$, and assume that $E \ncong 0$. It suffices to generate a family of pairwise disjoint subobjects $\{U_{x,y} \mid x \in X\}$ for each $y \in Y$ of A such that $Y = \bigcup W_x$, where $W_x = \{y \in Y \mid U_{x,y} \ncong 0\}$ for each $x \in X$. Note that since $U_{x,y} \wedge U_{x,y'} \neq 0$ implies $y = y'$, $|W_x| \leqslant |\mathbb{N}|$ for any x. Such a family would demonstrate that

$$|Y| \leqslant \sum_{x \in X} |W_x| \leqslant \sum_{x \in X} |\mathbb{N}| = |\mathbb{N}| \cdot |X| = |X|.$$

The second inequality is due to A being a c.a.c. object, and the last equality is due to X being infinite.

To construct this family of subobjects of A, first equate subobjects of A with subobjects of $1 \times A$, as these objects are isomorphic. Since $E \ncong 0$, E has at least two distinct subobjects (0 and E, say). These give rise to two distinct formulas $E \rightrightarrows \Omega_j$. Since $\mathbf{Sh}(, J)$ is a c.a.c. category, then, there is some (nonzero) c.a.c. object A and an arrow $k : A \to E$ to seperate the two formulas. Compose $k : A \to E$ with the monic representing $E \subseteq \mathbf{a}\Delta(Y)^{\mathbf{a}\Delta(X)}$ to produce the arrow $k' : A \to \mathbf{a}\Delta(Y)^{\mathbf{a}\Delta(X)}$, and define $g = (\mathrm{tr}(k'), \pi_2) : \mathbf{a}\Delta(X) \times A \to \mathbf{a}\Delta(Y) \times A$ to be the epic arrow corresponding to k'.

For each element $x : 1 \to X$, or $y : 1 \to Y$ in $\mathbb{S}\mathbf{ets}$, set $\hat{x} = \mathbf{a}\Delta(x)$ and $\hat{y} = \mathbf{a}\Delta(y)$. The objects $U_{x,y}$ are constructed as in the following adjacent pullback diagrams.

$$\begin{tikzcd}
1 \times A \arrow[r, "\hat{x} \times \mathrm{id}_A"] & \mathbf{a}\Delta(X) \times A \arrow[r, "g"] & \mathbf{a}\Delta(Y) \times A \\
U_{x,y} \arrow[u] \arrow[r] & V_y \arrow[u] \arrow[r] & 1 \times A \arrow[u, "\hat{y} \times \mathrm{id}_A"']
\end{tikzcd}$$

To see that $U_{x,y} \wedge U_{x,y'} = 0$ for distinct $y, y' \in Y$, it suffices to see that $V_y \wedge V_{y'} = 0$ for such y, y'. Of course, since $\mathbf{a}$ preserves pullbacks, 0, and 1, the following is a pullback square.

$$\begin{tikzcd}
1 \arrow[r, "\hat{y}"] & \mathbf{a}\Delta(Y) \\
0 \arrow[u] \arrow[r] & 1 \arrow[u, "\hat{y}'"']
\end{tikzcd}$$

The functor $(-) \times A$ also preserves pullback squares, so

$$\begin{tikzcd}
1 \times A \arrow[r, "\hat{y} \times \mathrm{id}"] & \mathbf{a}\Delta(Y) \times A \\
0 \arrow[u] \arrow[r] & 1 \arrow[u, "\hat{y}' \times \mathrm{id}"']
\end{tikzcd}$$

is a pullback square as well. Now consider the following commutaive cube.

$$\begin{tikzcd}
& 1 \times A \arrow[rr, "\hat{y}' \times \mathrm{id}"] & & \mathbf{a}\Delta(Y) \times A \\
0 \arrow[ur] \arrow[rr] & & 1 \times A \arrow[ur, "\hat{y} \times \mathrm{id}"'] & \\
& V_{y'} \arrow[uu] \arrow[rr] & & \mathbf{a}\Delta(X) \times A \arrow[uu, "g"'] \\
0 \arrow[uu] \arrow[ur] \arrow[rr] & & V_y \arrow[uu] \arrow[ur] &
\end{tikzcd}$$

Using a similar argument to the one found in the proof for Lemma 3.14, since it is known that every square but the bottom square is a pullback, so must the bottom square be a pullback. It follows that $V_y \wedge V_{y'} = 0$ for $y \neq y'$. It follows that $\{U_{x,y} \mid y \in Y\}$ is a pairwise disjoint famiy of subobjects of $1 \times A$.

Showing that $Y = \bigcup_{x \in X} W_x$ consists of three applications of *LAPCL* (*left adjoints preserve colimits*). The following are the adjoints involved:

- $(-) \times A \dashv (-)^A$,
- $\mathbf{a} \dashv i$, and
- $h_* \dashv \prod_h$ for any arrow h.

The first item in this list shows that $(-) \times A$ preserves coproducts. It follows from the second item in the list that

$$\mathbf{a}\Delta(X) \cong \mathbf{a}\Delta\left(\bigsqcup_{x \in X} 1\right) \cong \bigsqcup_{x \in X} \mathbf{a}\Delta(1) = \bigsqcup_{x \in X} 1,$$

so one also has

$$\mathbf{a}\Delta(X) \times A \cong \bigsqcup_{x \in X} (1 \times A).$$

For any $y \in Y$, this gives (via the third item in the list) the pullback diagram

$$\begin{tikzcd} V_y \arrow[r] & \mathbf{a}\Delta(X) \times A \\ \coprod_{x\in X} U_{x,y} \arrow[r] \arrow[u, "\cong"] & \coprod_{x\in X}(1 \times A) \arrow[u, "\cong"] \end{tikzcd}$$

Hence, $V_y \cong \coprod_{x\in X} U_{x,y}$. Recall that in a topos, pullbacks of epics are epic. The arrows $V_y \to 1 \times A$ are therefore all epic, since g is, so $V_y \ncong 0$ for any y. This makes $U_{x,y} \ncong 0$ for some $x \in X$, putting $y \in \bigcup_{x\in X} W_x$. The element $y \in Y$ was arbitrary, so $Y = \bigcup_{x\in X} W_x$. $\square$

Since κ was chosen specifically to be larger than $2^{\mathbb{N}}$, and $2^{\mathbb{N}}$ is strictly larger than $\mathbb{N}$,

$$\operatorname{epi}(\mathbf{a}\Delta(\mathbb{N}), \mathbf{a}\Delta(2^{\mathbb{N}})) \cong \operatorname{epi}(\mathbf{a}\Delta(2^{\mathbb{N}}), \mathbf{a}\Delta(\kappa)) \cong 0.$$

This makes

$$\mathbf{a}\Delta(\mathbb{N}) < \mathbf{a}\Delta(2^{\mathbb{N}}) < \mathbf{a}\Delta(\kappa) \leqslant \Omega_{\neg\neg}^{\mathbf{a}\Delta(\mathbb{N})}.$$

The Cohen topos is Boolean, and the usual global elements of $2^{\mathbb{N}}$ carry over to global elements of $\mathbf{a}\Delta(2^{\mathbb{N}})$, so 5.6.(d) follows directly from Lemma 5.5.

References

[1] S. Awodey. *Category Theory*. Oxford Logic Guides. Ebsco Publishing, 2006. ISBN: 9780191513824. URL: `https://books.google.com/books?id=IK%5C_sIDI2TCwC`.

[2] M. Barr and C. Wells. *Toposes, Triples and Theories*. Grundlehren der mathematischen Wissenschaften. Springer New York, 2013. ISBN: 9781489900234. URL: `https://books.google.com/books?id=FlySngEACAAJ`.

[3] J.L. Bell. *Set Theory: Boolean-Valued Models and Independence Proofs*. Oxford Logic Guides. Clarendon Press, 2005. ISBN: 9780191620829. URL: `https://books.google.co.uk/books?id=Z9fJEYd8G94C`.

[4] Timothy Y. Chow. “A beginner’s guide to forcing”. In: *arXiv e-prints*, arXiv:0712.1320 (Dec. 2007), arXiv:0712.1320. arXiv: `0712.1320 [math.LO]`.

[5] Kenny Easwaran. “A Cheerful Introduction to Forcing and the Continuum Hypothesis”. In: *arXiv e-prints*, arXiv:0712.2279 (Dec. 2007), arXiv:0712.2279. arXiv: `0712.2279 [math.LO]`.

[6] Michael P. Fourman. “Sheaf models for set theory”. In: *Journal of Pure and Applied Algebra* 19 (1980), pp. 91–101. ISSN: 0022-4049. DOI: https://doi.org/10.1016/0022-4049(80)90096-1. URL: http://www.sciencedirect.com/science/article/pii/0022404980900961.

[7] T. Jech. *Set Theory: The Third Millennium Edition, revised and expanded*. Springer Monographs in Mathematics. Springer Berlin Heidelberg, 2007. ISBN: 9783540447610. URL: https://books.google.co.uk/books?id=CZb-CAAAQBAJ.

[8] P.T. Johnstone. *Topos Theory*. Dover Books on Mathematics. Dover Publications, 2014. ISBN: 9780486493367. URL: https://books.google.co.uk/books?id=zO8WAgAAQBAJ.

[9] Alex Kruckman. *Notes on Ultrafilters*. Nov. 2012.

[10] K. Kunen. *Set Theory An Introduction To Independence Proofs*. Studies in Logic and the Foundations of Mathematics. Elsevier Science, 2014. ISBN: 9780080570587. URL: https://books.google.com/books?id=wWniBQAAQBAJ.

[11] S.M. Lane. *Categories for the Working Mathematician*. Graduate Texts in Mathematics. Springer New York, 2013. ISBN: 9781475747218. URL: https://books.google.co.uk/books?id=gfI-BAAAQBAJ.

[12] F. William Lawvere. “An Elementary Theory of the Category of Sets”. In: *Proceedings of the National Academy of Science* 52.6 (Dec. 1964), pp. 1506–1511. DOI: 10.1073/pnas.52.6.1506.

[13] S. MacLane and I. Moerdijk. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*. Universitext. Springer New York, 2012. ISBN: 9781461209270. URL: https://books.google.com/books?id=LZWLBAAAQBAJ.

[14] E. Riehl. *Category Theory in Context*. Aurora: Dover Modern Math Originals. Dover Publications, 2016. ISBN: 9780486809038. URL: https://books.google.com/books?id=Sr09DQAAQBAJ.

[15] David Roberts. *Class forcing and topos theory*. June 2018. DOI: 10.4225/55/5b2252e3092af. URL: https://adelaide.figshare.com/articles/Class_forcing_and_topos_theory/6530843/2.

[16] A. Scedrov. *Forcing and Classifying Topoi*. American Mathematical Society: Memoirs of the American Mathematical Society no. 295. American Mathematical Soc., 1984. ISBN: 9780821860397. URL: https://books.google.com/books?id=flekxiR3QroC.

[17] Myles Tierney. “Toposes, Algebraic Geometry and Logic: Dalhousie University, Halifax, January 16-19, 1971”. In: Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2006, pp. 13–42. ISBN: 9783540376095. URL: https://books.google.co.uk/books?id=TOH7CAAAQBAJ.