Second European Summer School

Sheaves and Logic

Ieke Moerdijk
Mathematisch Instituut
Rijksuniversiteit Utrecht

Language, Logic and Information

Topological Methods in Mathematical Logic

Ieke Moerdijk
Mathematical Institute
University of Utrecht

The purpose of these lectures is to present some applications of topological methods in mathematical logic. Here 'topology' is interpreted in the wide sense of the Grothendieck school, so as to include the theory of Grothendieck topologies, of sheaves and of topoi.

In the first lecture, I present an outline of the basic facts of sheaf theory. This material is quite standard, and goes back to Grothendieck et. al. (SGA4). I have tried not to present more than what is motivated by the applications which follow.

In the second lecture, I present the interpretation of higher order logic in sheaf categories, via forcing. This type of forcing generalizes some of the well-known forcing techniques such as those of Cohen (1966) and Kripke (1965). In the generality presented here, it is primarily due to A. Joyal, and is often called "Kripke-Joyal semantics".

In the next two lectures, I consider two special cases of this interpretation. First, I present Freyd's proof of the independence of the axiom of choice (Freyd (1980)), and next I present what seems to me to be the smoothest proof of the consistency with repect to higher order intuitionistic logic of the continuity of all functions from the reals to the reals. Lectures II-IV are essentially taken from my forthcoming book with S. MacLane.

In the fifth and last lecture, I will use sheaf theory to give a "semantical" proof (due again to A. Joyal) of the fact that in intuitionistic logic, every definable function $\mathbf{R} \to \mathbf{R}$ is provably continuous. Finally, I will conclude with a sheaf theoretic proof that there are no definable non-principal ultrafilters in ZFC (or in classical higher order logic with choice).

There are now several introductions to topos theory available which are written for a logical audience, e.g. Lambek-Scott (1986), Bell (1989). Apart from SGA4, the principal reference work in the field remains Johnstone's comprehensive exposition (1977). However, this book might be a little hard-going for readers with a background primarily in logic (and some of the many logical applications contained in Johnstone's book are perhaps not immediately recognized as such).

In a short course like the present one, it is always hard to find a balance between the general theory and the concrete examples. I have tried to focus on the applications, rather than to develop a lot of abstract theory. A number of applications are presented which seem basic, but are not included in the text books just mentioned. Naturally, this course represents only a tip of the iceberg, and I have not even mentioned 'standard' applications such as the theory of geometric logic and 'generic models' (Makkai-Reyes (1977)), synthetic differential geometry (Kock (1981), Moerdijk-Reyes (1987)), or the intriguing connection between Kleene realizability and topos theory (Hyland (1982), (1988)). Nevertheless, I hope that this short survey does demonstrate that logic and geometry are intimately related.

Utrecht, July 1990, I.M.

Lecture I: Grothendieck topologies and sheaves.

§1. Grothendieck topologies.

In what follows, $\mathbf{C}$ will denote a small category; $\mathbf{C}^{op}$ is the corresponding opposite category.

1.1. Presheaves. A presheaf on $\mathbf{C}$ is a functor

$$P : \mathbf{C}^{op} \to \text{Sets}.$$

Thus P is given by a system of sets $P(C)$ (where C is any object from $\mathbf{C}$), together with a function $P(f)$ for any morphism $f : D \to C$ in $\mathbf{C}$:

$$P(f) : P(C) \to P(D), x \mapsto P(f)(x) = x \cdot f = x \mid f,$$

in such a way that the following identities hold:

$$\begin{aligned} (x \cdot f) \cdot g &= x \cdot (fg) \\ x \cdot id &= x \end{aligned}$$

(for any $x \in P(C), f : D \to C, g : E \to D$).

A *morphism* (or *natural transformation*) $\tau : P \to Q$ between presheaves P and Q on $\mathbf{C}$ is a system of functions

$$\tau_C : P(C) \to Q(C) \quad (C \in \mathbf{C})$$

such that for any $f : D \to C$ and any $x \in P(C)$,

$$\tau_D(x \cdot f) = \tau_C(x) \cdot f.$$

The presheaves on $\mathbf{C}$ and the morphisms between them form a category, denoted $\text{Sets}^{\mathbf{C}^{op}}$. The set of morphisms from P to Q is denoted $\underline{\text{Hom}}(P, Q)$.

1.2 Representable presheaves. Each object A of $\mathbf{C}$ gives rise to a presheaf

$$\underline{y}(A) = \mathbf{C}(-, A)$$

defined by

$$\underline{y}(A)(C) = \mathbf{C}(C, A) = \text{ the set of arrows } C \to A \text{ in } \mathbf{C}.$$

For an arrow $f : D \to C$ in $\mathbf{C}$, the restriction $\underline{y}(A)(C) \to \underline{y}(A)(D)$ is given by composition: for any $g \in \underline{y}(A)(C)$ we have $(\underline{y}(A)(f))(g) = g \circ f$. An arrow $h : A \to B$ yields a morphism of presheaves

$$\underline{y}(h) : \underline{y}(A) \to \underline{y}(B),$$

defined again by composition. In this way, $\underline{y}$ becomes a functor $\mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{op}}$, called the *Yoneda embedding.*

Yoneda lemma. *For any presheaf P, there is a natural bijection* $\underline{\text{Hom}}\,(\underline{y}(A), P) \cong P(A)$, *for any* $A \in \mathbf{C}$.

Proof. The bijection associates to a morphism $\tau : \underline{y}(A) \to P$ the element $\tau_A(id) \in P(A)$, where $id \in \underline{y}(A)(A)$ is the identity on A.

1.3 Subpresheaves. A subpresheaf of a presheaf P is a presheaf S such that $S(C) \subset P(C)$ for each object C in $\mathbf{C}$, and such that the restrictions of S are those of P (i.e. $P(f)(x) = S(f)(x)$ for $f : D \to C$ and $x \in S(C)$). Let $Subp(P)$ denote the set of subpresheaves on P. If $\tau : Q \to P$ is a morphism of presheaves, τ induces an operation

$$\tau^{-1} : Subp(P) \to Subp(Q)$$

by

$$\tau^{-1}(S)(C) = \tau_C^{-1}(S(C)) = \{x \in Q(C) \mid \tau_C(x) \in S(C)\}.$$

1.4 Sieves. A sieve on an object C of $\mathbf{C}$ is by definition a subpresheaf S of the representable presheaf $\underline{y}(C)$. Notice that we may also view S as a collection of arrows into C, with the property that if $f : D \to C$ is in this collection, then so is $f \circ g$ for any arrow g with codomain f. (I will use both points of view without distinguishing them notationally.)

1.5 Grothendieck topologies. A Grothendieck topology on $\mathbf{C}$ assigns to each object C a collection of sieves $J(C)$ on C, such that
(i) $\underline{y}(C) \in J(C)$; ($\underline{y}(C)$ is subpresheaf of itself)
(ii) if $g : B \to C$ and $S \in J(C)$, then $\underline{y}(g)^{-1}(S) \in J(D)$;
(iii) for any sieve R on C, if there exists an $S \in J(C)$ such that for any arrow $(g : B \to C) \in S$ we have $\underline{y}(g)^{-1}(R) \in J(B)$, then also $R \in J(C)$.

The sieves in $J(C)$ are called *covers* of C; (ii) is called the stability axiom, and (iii) the transitivity axiom. A *site* is a pair $(\mathbf{C}, J)$ where J is a Grothendieck topology on the category $\mathbf{C}$.

Perhaps the principal example to keep in mind is the following. Let X be a topological space, and let $\mathcal{O}(X)$ be the collection of open subsets of X. Make $\mathcal{O}(X)$ into a category, with exactly one arrow $U \to V$ iff $U \subset V$. For a sieve S on $U \in \mathcal{O}(X)$, say $S \in J(U)$ iff $U = \cup\{W \subset U \mid (W \to U) \in S\}$. This defined a Grothendieck topology J on $\mathcal{O}(X)$.

A list of further examples occures in §II.2 below.

§2. Sheaves.

In this section, $(\mathbf{C}, J)$ is a fixed site.

2.1 Definition of sheaf. Let F be a presheaf on $\mathbf{C}$. F is called a sheaf (for J) if for any object C of $\mathbf{C}$ and any cover $S \in J(C)$, any morphism $\tau : S \to F$ can be uniquely extended to a morphism $\tau' : \underline{y}(C) \to F$, as in

$$\begin{array}{ccc} S & \hookrightarrow & \underline{y}(C) \\ {\scriptstyle\tau}\downarrow & \swarrow {\scriptstyle\tau'} & \\ F & & \end{array}$$

So a sheaf is a special kind of presheaf. A morphism between sheaves is a morphism as defined for presheaves; $\underline{\text{Hom}}\,(F, G)$ again denotes the set of morphisms from a sheaf F to another one G. Thus we have a category $Sh(\mathbf{C}, J)$ of sheaves on $\mathbf{C}$ (a full subcategory $\text{Sets}^{\mathbf{C}^{op}}$).

By definition, a *Grothendieck topos* is a category of the form $Sh(\mathbf{C}, J)$ for some site $(\mathbf{C}, J)$.

2.2 The one-point sheaf. The presheaf $1 \in \text{Sets}^{\mathbf{C}^{op}}$ defined by $1(C) = \{*\}$ (some one-point set) is a sheaf. It is characterized by the property that for any sheaf F, there is exactly one morphism $F \to 1$ (i.e. 1 is a terminal object of $Sh(\mathbf{C}, J)$; CWM p 20).

2.3 Products of sheaves. Let P and Q be presheaves on $\mathbf{C}$ and form their "pointwise" product $P \times Q$, i.e.

$$(P \times Q)(C) = P(C) \times Q(C)$$

with the evident restrictions. This presheaf $P \times Q$ is the *product* in the category $\text{Sets}^{\mathbf{C}^{op}}$ (CWM p 69): there are projection morphisms

$$P \xleftarrow{\pi_1} P \times Q \xrightarrow{\pi_2} Q,$$

and for any other presheaf R and any morphisms $\alpha : R \to P, \beta : R \to Q$ there is exactly one morphism $(\alpha, \beta) : R \to P \times Q$ such that $\pi_1(\alpha, \beta) = \alpha$ and $\pi_2(\alpha, \beta) = \beta$. It follows immediately that if P and Q are sheaves then so is $P \times Q$. So $Sh(\mathbf{C}, J)$ is a category with *products.*

2.4 Function sheaves. Again, we first consider presheaves P and Q. There is a presheaf Q^P defined by

$$Q^P(C) = \underline{\text{Hom}}(\underline{y}(C) \times P, Q);$$

for $f : D \to C$ the restriction $Q^P(f) : Q^P(C) \to Q^P(D)$ sends a morphism $\alpha : \underline{y}(C) \times P \to Q$ to the composition

$$\underline{y}(D) \times P \xrightarrow{\underline{y}(f) \times id} \underline{y}(C) \times P \xrightarrow{\alpha} Q.$$

This presheaf Q^P is characterized by the fact that for any other presheaf R, there is a natural bijection between sets of morphisms

$$\mathrm{Hom}(R, Q^P) \cong \mathrm{Hom}(P \times R, Q).$$

For $R = Q^P$, the identity on the left corresponds to a morphism

$$ev : P \times Q^P \to Q \tag{1}$$

called *evaluation*; explicitly, for $C \in \mathbf{C}, x \in P(C)$ and $(\alpha : \underline{y}(C) \times P \to Q) \in Q^P(C)$,

$$ev_C(x, \alpha) = \alpha_C(id, x) \in Q(C).$$

(It follows by naturality that the bijection (1) sends $\tau : R \to Q^P$ to the composite $ev \circ (id \times \tau) : P \times R \to Q$.)

Although a little harder than for products, one can check that if P and Q are sheaves then so is Q^P. For sheaves P and Q, we call Q^P the *function sheaf*, or *exponential*. It follows that $Sh(\mathbf{C}, J)$ is a *cartesian closed category* (CWM p 95).

2.5 Subsheaves. Let P be a presheaf on $\mathbf{C}$, and let $U \subset P$ be a subpresheaf (as in 1.3). We call U *closed* (with respect to J) if for any object C of $\mathbf{C}$ and any cover $S \in J(C)$, for any commutative square as below there exists a (necessarily unique) diagonal morphism such that both triangles commute:

$$\begin{array}{ccc} S & \to & U \\ \downarrow & \nearrow & \downarrow \\ \underline{y}(C) & \to & P \end{array}$$

One can spell this out via the Yoneda lemma, and get the following equivalent description: $U \subseteq P$ is closed iff for any $C \in \mathbf{C}$, and any $S \in J(C)$ and $x \in P(C)$, if for every arrow $g \in S$ it holds that $x \cdot g \in U(\mathrm{dom}(g))$, then $x \in U$. (Informally, any element of P which is "locally" in U, is in U.) One readily verifies:

Lemma. *Let P a sheaf, and let $U \subseteq P$ be a subpresheaf. Then U is closed iff U is itself a sheaf. (U is then called a subsheaf* of P.)

For a sheaf F, we denote by $Sub(F)$ the collection of subsheaves of F. If $\tau : F \to G$ is a morphism of (pre)sheaves, then clearly for a closed subpresheaf $U \subseteq G$, the subpresheaf $\tau^{-1}(U) \subseteq F$ (as defined in 1.3) is again closed. In other words, the operation τ^{-1} on subpresheaves restricts to an operation on subsheaves:

$$\tau^{-1} : Sub(G) \to Sub(F).$$

(For a sheaf F, $Sub(F)$ has a great deal of structure which we will not go into here; e.g. it is always a complete Heyting algebra.)

2.6 Powersheaves. For a sheaf F, we can define a powersheaf $\mathcal{P}(F)$ which plays the role of the power set. Let for an object $C \in \mathbf{C}$,

$$\mathcal{P}(F)(C) = CSubp(\underline{y}(C) \times F) = \text{the set of } \textit{closed} \text{ subpresheaves of } \underline{y}(C) \times F. \quad (1)$$

For an arrow $f : D \to C$ in $\mathbf{C}$, the restriction $\mathcal{P}(F)(f) : \mathcal{P}(F)(C) \to \mathcal{P}(F)(D)$ sends a closed subpresheaf $\subseteq \underline{y}(C) \times F$ to the inverse image subpresheaf $(\underline{y}(f) \times id)^{-1}(U)$, where $\underline{y}(f) \times id : \underline{y}(D) \times F \to \underline{y}(C) \times F$; this inverse image is again closed.

The presheaf $\mathcal{P}(F)$ is characterized by a universal property: For any presheaf R, there is a natural bijection

$$\underline{\text{Hom}}(R, \mathcal{P}(F)) \cong CSubp(R \times F). \quad (2)$$

Using this bijection, one readily shows that $\mathcal{P}(F)$ is a sheaf if F is — it is called the *powersheaf* of F. Among sheaves R and F, we can rewrite the bijection (2) as

$$\underline{\text{Hom}}(R, \mathcal{P}(F)) \cong Sub(R \times F), \quad (3)$$

by the lemma in 2.5. As with exponentials, this bijection is completely determined by what it does to the identity for the special case where $R = \mathcal{P}(F)$: one gets a subsheaf

$$E \subseteq \mathcal{P}(F) \times F \quad (4)$$

which is the 'membership relation': For $C \in \mathbf{C}, x \in F(C)$ and $U \in \mathcal{P}(F)(C)$, that is $U \subseteq \underline{y}(C) \times F$ closed, we have

$$(U, x) \in E(C) \text{ iff } (id_C, x) \in U(C). \quad (5)$$

2.7 Other limits. Pullback and equalizers (CWM p 70,71) of sheaves are constructed pointwise (as products, see 2.3). For example, if $\tau : G \to F$ and $\sigma : H \to F$ are morphisms of sheaves then we can define a presheaf $G \times_F H$ by

$$(G \times_F H)(C) = \{(x, y) \mid x \in G(C), y \in H(C), \tau_C(x) = \sigma_C(y)\},$$

with evident restrictions. It is easily seen that $G \times_F H$ is a sheaf (assuming F, G, H are sheaves). Indeed, this follows from the fact that $G \times_F H$ is characterized by a universal property of the projections $G \xleftarrow{\pi_1} G \times_F H \xrightarrow{\pi_2} H$: for any two morphisms $G \xleftarrow{\varphi} K \xrightarrow{\psi} H$ from some (pre)sheaf K, if $\tau\varphi = \sigma\psi$ then there is a unique $\chi : K \to G \times_F H$ with $\pi_1\chi = \varphi, \pi_2\chi = \psi$ (cf. CWM p 71).

2.8 The zero-sheaf. Suppose F is a sheaf and C is an object of $\mathbf{C}$ such that $\emptyset \in J(C)$. Then the unique map of presheaves $\emptyset \to F$ must have a unique extension $\underline{y}(C) \to F$. By the Yoneda lemma, this means that $F(C)$ is a set with exactly one element. It follows that in general, the empty presheaf $\emptyset$ is not a sheaf. Nonetheless, there is a smallest sheaf 0, defined by

$$0(C) = \begin{cases} \{*\}, \text{ if } \emptyset \in J(C) \\ \emptyset, \text{ otherwise.} \end{cases}$$

Restrictions $0(C) \to 0(D)$ for $f : D \to C$ are defined because by the stability axiom, $\emptyset \in J(C)$ implies $\emptyset \in J(D)$. The sheaf 0 is called the zero-sheaf; it is an initial object of $Sh(\mathbf{C}, J)$ (cf. CWM p 20).

2.9 Sums of sheaves. Let $\{F_i : i \in I\}$ be a collection of sheaves. Their *sum* (or *coproduct*) $\sum F_i$ is to be a sheaf, together with morphisms $\sigma_j : F_j \to \sum F_i$ (for each $j \in I$), having the following universal property: given any sheaf G and any family of morphisms $\theta_j : F_j \to G (j \in I)$, there is a unique morphism $\theta : \sum F_i \to G$ with $\theta\sigma_j = \theta_j$ for all j.

Define $\sum F_i$ as follows. For $C \in \mathbf{C}$, elements of $(\sum F_i)(C)$ are equivalence classes $[\alpha]$ of families of morphisms $\alpha = \{\alpha_i : S_i \to F_i\}_{i \in I}$, where for each $i \in I, S_i \hookrightarrow \underline{y}(C)$ is a sieve on C such that $\cup S_i \in J(C)$, and whenever $i \neq j$ the sieves S_i and S_j are "disjoint", in the sense that if an arrow $h : D \to C$ belongs to $S_i \cap S_j$ then $\emptyset \in J(D)$. Two such $\{\alpha_i : S_i \to F_i\}$ and $\{\beta_i : T_i \to F_i\}$ are equivalent of there exist sieves $R_i \subseteq S_i \cap T_i$ such that $\cup R_i$ covers C and $\alpha_i \mid R_i = \beta_i \mid R_i$ for each $i \in I$.

If $f : D \to C$ is a morphism and $[\alpha] \in (\sum F_i)(C)$, then the restriction $[\alpha] \cdot f$ is represented by the family of maps $f^*(\alpha_i) : \underline{y}(f)^{-1}(S_i) \to F_i$ induced by the α_i, as in

$$\begin{array}{ccccc} \underline{y}(D) & \hookleftarrow & \underline{y}(f)^{-1}(S_i) & & \\ {\scriptstyle \underline{y}(f)} \downarrow & & \downarrow & \searrow {\scriptstyle f^*(\alpha_i)} & \\ \underline{y}(C) & \hookleftarrow & S_i & \xrightarrow{\alpha_i} & F_i. \end{array}$$

The 'universal' morphism $\sigma_j : F_j \to \sum F_i$ sends an element $x \in F_j(C)$ to the class of the family $\{\alpha_i : S_i \to F_i\}$ where $S_i = \emptyset$ if $i \neq j$, and $S_j = \underline{y}(C)$ while $\alpha_j : y(C) \to F_i$ corresponds to x via the Yoneda lemma.

One readily verifies that $\sum F_i$ is a sheaf with the univeral property as stated.

Notice also the following distributive law, for a sheaf G and a family of sheaves F_i:

$$G \times \sum F_i \cong \sum (G \times F_i).$$

2.10 Constant sheaves. For a set I, the sum of the family of sheaves $\{1 \mid i \in I\}$ of I-many copies of 1 is denoted by $\Delta(I)$, and called the *constant sheaf* corresponding to the set I:

$$\Delta I = \sum_{i \in I} 1.$$

For any sheaf F, write $\Gamma F = \underline{\text{Hom}}\ (1, F)$. The universal property of the sum then yields a natural bijection

$$\underline{\text{Hom}}(\Delta(I), F) \cong \underline{\text{Hom}}\ (I, \Gamma F)$$

(on the right, $\underline{\text{Hom}}$ is just the set of function). In other words, Δ is *left adjoint* to Γ (CWM p).

Lecture II: Logic.

§1. Forcing.

Throughout this section, $(\mathbf{C}, J)$ is a fixed site.

1.1 Language. Let us fix a language $\mathcal{L}$ for many-sorted predicate logic. Recall that $\mathcal{L}$ is given by the following data: a set of *sorts* (or *types*) $A, B, ..$ etc; with each sort A infinitely many variables $x^A, y^A, ..$ (I shall usually drop the A on the variables, and/or write $\forall x \in A\varphi$ for $\forall x^A \varphi$.) Furthermore the language has a set of constants c, functions f and relations R; each constant $c = c^A$ has a sort A assigned to it; with each function we are given its number of argument n say), the sort of each of these arguments (say $A_1, ..., A_n$ respectively) and the sort of its values (say B) — I abbreviate this as

$$f : A_1 \times ... \times A_n \to B.$$

Finally, each relation R is given with its number of arguments and the sorts of these arguments, say $A_1, ..., A_n$ — we write

$$R \subseteq A_1 \times ... \times A_n.$$

Terms and formulas are built up in the usual way.

1.2 Interpretation. A model $\mathcal{M}$ of $\mathcal{L}$ (over the site $(\mathbf{C}, J)$) is given by the following data (besides the site $(\mathbf{C}, J)$):

- for each sort A a sheaf $A^{\mathcal{M}}$
- for each constant c of sort A an arrow $c^{\mathcal{M}} : 1 \to A^{\mathcal{M}}$
- for each function $f : A_1 \times ... \times A_n \to B$ a morphism of sheaves $f^{\mathcal{M}} : A_1^{\mathcal{M}} \times ... \times A_n^{\mathcal{M}} \to B^{\mathcal{M}}$
- for each relation $R \subseteq A_1 \times ... \times A_n$ a subsheaf $R^{\mathcal{M}} \subseteq A_1^{\mathcal{M}} \times ... \times A_n^{\mathcal{M}}$.

By induction, one now defines for a term $t(x_1, ..., x_n)$ of sort B with free variables x_i of sort A_i, an arrow

$$t^{\mathcal{M}} : A_1^{\mathcal{M}} \times ... \times A_n^{\mathcal{M}} \to B^{\mathcal{M}}$$

in the standard way:

- if $t = x_i$ is a variable, $t^{\mathcal{M}}$ is the projection $A_1^{\mathcal{M}} \times ... \times A_n^{\mathcal{M}} \to A_i^{\mathcal{M}}$
- if $t = c^A$ is a constant, $t^{\mathcal{M}}$ is the given $c^{\mathcal{M}} : 1 \to A^{\mathcal{M}}$
- if $t = f(t_1(x_1, ..., x_n), ..., t_m(x_1, ..., x_n))$ and $t_i^{\mathcal{M}} : A_1^{\mathcal{M}} \times ... \times A_n^{\mathcal{M}} \to B_i^{\mathcal{M}}$ has been defined, then $t^{\mathcal{M}}$ is the composite

$$A_1^{\mathcal{M}} \times \dots \times A_n^{\mathcal{M}} \xrightarrow{<t_1^{\mathcal{M}},\dots,t_m^{\mathcal{M}}>} B_1^{\mathcal{M}} \times \dots \times B_m^{\mathcal{M}} \xrightarrow{t^{\mathcal{M}}} B^{\mathcal{M}}.$$

Convention: from now on, I will often delete the superscript $\mathcal{M}$.

Next, one defines by induction the forcing relation

$$C \Vdash \varphi(a_1, \dots, a_n)$$

where C is an object from the site $\mathbf{C}$, $\varphi(x_1, \dots, x_n)$ is a formula with free variables (among) x_i of sort A_i, and A_i is an element of $A_i(C) = A_i^{\mathcal{M}}(C)$. The clauses for atomic formulas are

$$C \Vdash R(t_1(a_1, \dots, a_n), \dots, t_m(a_1, \dots, a_n)) \text{ iff } ((t_1^{\mathcal{M}})_C(a_1, \dots, a_n), \dots, (t_m^{\mathcal{M}})_C(a_1, \dots, a_n)) \in R^{\mathcal{M}}(C).$$

$$C \Vdash t_1(a_1, \dots, a_n) = t_2(a_1, \dots, a_n) \text{ iff } (t_1^{\mathcal{M}})_C(a_1, \dots, a_n) = (t_2^{\mathcal{M}})_C(a_1, \dots, a_n).$$

Furthermore,

$C \Vdash \varphi \wedge \psi(a_1, \dots, a_n)$ iff $C \Vdash \varphi(a_1, \dots, a_n)$ and $C \Vdash \psi(a_1, \dots, a_n)$

$C \Vdash \varphi \vee \psi(a_1, \dots, a_n)$ iff there is a cover $S \in J(C)$ such that for any arrow $g : D \to C$ in S, $D \Vdash \varphi(a_1 \cdot g, \dots, a_n \cdot g)$ or $D \Vdash \psi(a_1 \cdot g, \dots, a_n \cdot g)$.

$C \Vdash \bot$ iff $\phi \in J(C)$

$C \Vdash \neg\varphi$ iff $C \Vdash (\varphi \to \bot)$

$C \Vdash \forall y \in B\varphi(a_1, \dots, a_n, y)$ iff for any arrow $g : D \to C$ and any $b \in B^{\mathcal{M}}(D)$, $D \Vdash \varphi(b, a_1 \cdot g, \dots, a_n \cdot g)$.

$C \Vdash \exists y \in B\varphi(a_1, \dots, a_n, y)$ iff there is a cover $S \in J(C)$ such that for any arrow $g : D \to C$ in S there is some $b_g \in B(D)$ with $D \Vdash \varphi(b_g, a_1 \cdot g, \dots, a_n \cdot g)$.

1.3 Basic properties. By induction on φ, one easily verifies:

Lemma. *The forcing relation has the following two properties:*
(i) *(monotonicity) If $C \Vdash \varphi(a_1, \dots, a_n)$ then $D \Vdash \varphi(a_1 \cdot g, \dots, a_n \cdot g)$ for any arrow $g : D \to C$.*
(ii) *(local character) If there exists a cover $S \in J(C)$ such that $D \Vdash \varphi(a_1 \cdot g, \dots, a_n \cdot g)$ for any $(g : D \to C) \in S$, then also $C \Vdash \varphi(a_1, \dots, a_n)$.*

Consider for $\varphi(x_1, \dots, x_n)$ as above and $C \in \mathbf{C}$ the subset

$$\{(x_1, \dots, x_n) \mid \varphi\}^{\mathcal{M}}(C) =_{\text{def}} \{(a_1, \dots, a_n) \mid a_i \in A_i(C) \text{ and } C \Vdash \varphi(a_1, \dots, a_n)\}.$$

Then (i) of the lemma says that $\{(x_1, ..., x_n) \mid \varphi\}^{\mathcal{M}}$ is a subpresheaf of $A_1 \times ... \times A_n$, while (ii) says that in fact it is a subsheaf (cf. I.2.5). Thus any formula defines a subsheaf by "comprehension". Notice that if φ doesn't have any free variables, then $\{\cdot \mid \varphi\}^{\mathcal{M}}$ is a subsheaf of the empty product 1.

A formula φ is said to be *valid in the model* $\mathcal{M}$ *over* $(\mathbf{C}, J)$, notation $M \models \varphi$, if for any object C of $\mathbf{C}$ and any $a_i \in A_i(C)$, $C \Vdash \varphi(a_1, ..., a_n)$. When $\mathcal{M}$ is understood (or when $\mathcal{M}$ is the standard interpretation, cf. 1.5 below), I also write $Sh(\mathbf{C}, J) \models \varphi$, rather than $\mathcal{M} \models \varphi$.

Soundness theorem. *All formulas provable in many-sorted intuitionistic predicate logic are valid in any model* $\mathcal{M}$ *over a site* $(\mathbf{C}, J)$.

There is also a completeness theorem, see II. 2.1 below. (For a list of axioms for intuitionistic predicate logic see e.g. Boileau-Joyal (1981), Kleene (1952), Dummett (1977). There are some subtleties involved when one doesn't assume sorts to be inhabited, but I don't go into this here.)

1.4 Higher order logic (or intuitionistic type theory). This is many-sorted intuitionistic predicate logic, with some additional structure on the collection of sorts. For instance, one requires the existence of (i) *finite products*; i.e. a sort 1 (the empty product), and for any sorts A and B a product sort $A \times B$, together with *projection functions* $p_1 : A \times B \to A$ and $p_2 : A \times B \to B$. (ii) For any two sorts A and B there is a *function-sort* B^A, together with an *evaluation function* $ev : A \times B^A \to B$. (iii) For any sort A there is a *power-sort* $\mathcal{P}A$, which comes with a *membership relation* $E \subseteq \mathcal{P}A \times A$. (As usual, I write $t(s)$ for $ev(s,t)$ and $t \in s$ for $E(s,t)$.) There are axioms (in addition to those of first-order intuitionistic logic) which ensure that these constructions have their intended meaning:

$$\text{(1 is a singleton)} \quad \exists! x \in 1 (x = x)$$

$$\text{(pairing)} \quad \forall x \in A \forall y \in B \exists! z \in A \times B (p_1 z = x \wedge p_2 z = y)$$

$$\text{(function} - \text{comprehension)} \quad \forall x \in A \exists! y \in B \varphi(x,y) \to \exists! f \in B^A \forall x \in A \varphi(x, f(x))$$

$$\text{(set} - \text{comprehension)} \quad \exists! x \in \mathcal{P}A \forall y \in A (y \in x \leftrightarrow \varphi(y)).$$

(In these axioms, φ may have additional free variables. Recall that while $y \in x$ abbreviates $E(x,y), \forall x \in A$ is just a formal device to indicate that x is a variabele of sort A. One often writes $*$ for the unique x in the first axiom, $< x, y >$ for the unique z in the second axiom, $\{y \in A \mid \varphi(y)\}$ for the unique x in the last axiom.)

A (*standard*) interpretation of type theory over a site $(\mathbf{C}, J)$ is given by an operation $\mathcal{M}$ as before, which "preserves" the structure on the types. Thus $(A \times B)^{\mathcal{M}}$ is the sheaf product $A^{\mathcal{M}} \times B^{\mathcal{M}}$, and $(p_i)^{\mathcal{M}}$ are the projections; and $1^{\mathcal{M}} = 1$. Furthermore $(B^A)^{\mathcal{M}}$ is the exponential $(B^{\mathcal{M}})^{A^{\mathcal{M}}}$ of I. 2.4 while $ev^{\mathcal{M}}$ is the evaluation morphism. Finally $(\mathcal{P}A)^{\mathcal{M}} = \mathcal{P}(A^{\mathcal{M}})$, the powersheaf of I. 2.6, while $E^{\mathcal{M}}$ is the membership-subsheaf as defined there.

For the standard interpretation, there is a *soundness theorem* for intuitionistic higher order logic (its proof is routine).

1.5 Arithmetic. Peano arithmetic (PA) denotes some standard version of classical first order arithmetic; Heyting arithmetic (HA) is its intuitionistic counterpart. Higher order Heyting arithmetic (HHA) is higher order intuitionistic logic as in 1.4, with in addition a distinguished sort N (for the natural numbers), a constant 0 of type N (zero), and a successor function $(\cdot + 1) : N \to N$, together with the usual axioms for (intuitionistic) arithmetic, including *full* higher order induction $\forall U \in \mathcal{P}N(0 \in U \wedge \forall n \in N(n \in U \to n+1 \in U) \to \forall n \in N(n \in U))$. The standard interpretation of HHA is the standard interpretation as above, such that in addition $N^{\mathcal{M}} = \Delta(\mathbf{N})$, the constant sheaf corresponding to the set of integers (cf. I 2.10). This interpretation is sound (but not complete, see 2.6 below).

1.6 (Intuitionistic) set theory (cf. Fourman (1980)). IZF denotes Zermelo-Fraenkel set theory with intuitionistic logic (for a list of axioms see Fourman, loc. cit.). IZF differs from HHA in many respects: IZF has no types (there is only one "improper" type, the universe V) and has unbounded quantification. Nonetheless, any site $(\mathbf{C}, J)$ gives rise to a sheaf model of IZF in a standard way, closely related to the standard interpretation of HHA. This model is obtained by imitating the construction of the cummulative hierarchy in $Sh(\mathbf{C}, J)$: For each ordinal α we define a sheaf V_α on $(\mathbf{C}, J)$, as

$$V_\phi = 0 \quad (\text{zero} - \text{sheaf})$$

$$V_{\alpha+1} = \mathcal{P}(V_\alpha) \quad (\text{power} - \text{sheaf})$$

$$V_\lambda = \varinjlim_{\alpha<\lambda} V_\alpha \quad (\lambda \text{ a limit ordinal})$$

Here $\varinjlim$ is the colimit in the category of sheaves, which I haven't described explicitly. There is no sheaf $V = \bigcup_\alpha V_\alpha$ (unless we allow "class-valued" sheaves). Roughly, one can now proceed as follows: add infinite conjunction to HHA (with the obvious rule for forcing), and define a translation from set theory into this extension of HHA, with the V_α's as types. (A first approximation is that an IZF-formula $\forall x\varphi(x)$ translates into $\bigwedge_\alpha \forall x^\alpha \in V_\alpha \varphi(x^\alpha)$; this simple idea has to be adapted somewhat for nested quantifiers.) Using this translation, one can now define forcing and validity for set theory in terms of forcing and validity for (extended) HHA. This yields a sound interpretation of IZF. I will not discuss further details. As a rule of thumb: any independence proof for HHA is also an independence proof for IZF, as long as all the interpretations of the types involved are contained in some V_α.

1.7 Classical logic. Many sheaf categories $Sh(\mathbf{C}, J)$ have the additional property that the standard interpretation is sound for classical logic, so that one gets models for classical higher order logic and classical ZF. Classical higher order logic is HHA with an additional axiom for the "excluded middle":

$$\forall p \in \mathcal{P}(1)(* \in p \vee \neg(* \in p)) \tag{1}$$

The most frequently occurring case in which the axiom of the excluded middle is valid is that of the *dense topology* or *double negation topology* $J = \neg\neg$. This topology is defined, for $C \in \mathbf{C}$, and $S \subseteq \underline{y}(C)$, by: $S \in \neg\neg(C)$ iff for any arrow $f : D \to C$ there exists an arrow $g : E \to D$ such that $fg \in S(E)$.

Proposition. *Any standard interpretation in* $Sh(\mathbf{C}, \neg\neg)$ *is classical.*

(Indeed, one easily verifies that (1) is valid in $Sh(\mathbf{C}, \neg\neg)$.)

§2. Examples.

In this section I review very briefly how sheaf models generalize some well-known types of semanties.

2.1 Kripke models (Kripke (1966)). Let $J = triv$ be the trivial topology on a category $\mathbf{C}$, where the *only* cover $S \rightarrowtail \underline{y}(C)$ of C is $\underline{y}(C)$ itself. Then every presheaf is a sheaf:

$$Sh(\mathbf{C}, triv) = \mathrm{Sets}^{\mathbf{C}^{op}}.$$

For this special topology, the forcing clauses for disjunction and existential quantification can be simplified. For instance (omitting free variables from notation)

$$C \Vdash \varphi \vee \psi \text{ iff } C \Vdash \varphi \text{ or } C \Vdash \psi.$$

In this way, one obtains a description of forcing equivalent to that of Kripke (loc. cit.). Notice that Kripke in fact considered more special models: for the category $\mathbf{C}$ he took a partially ordered set $\mathbf{P}$ (viewed as a category with exactly one arrow $p \to q$ iff $p \leq q$ in $\mathbf{P}$), and he considered only presheaves $A : \mathbf{P}^{op} \to \mathrm{Sets}$ all whose restriction maps $A(p) \to A(q)$ (for $q \leq p$) are inclusions. By Kripke's completeness theorem for first order intuitionistic logic, we conclude a fortiori that our sheaf semantics is complete.

2.2 Beth-models (Beth (1956)). Consider the set $\mathbf{N}^{<\mathbf{N}}$ of finite sequences u of natural numbers. Write $u \leq v$ if u is an extension of v. A *subtree* $T \subseteq \mathbf{N}^{<\mathbf{N}}$ is a collection of such sequences containing the empty sequence $\langle\ \rangle$ and closed under intial segments ($v \geq u \in T$ implies $v \in T$). The poset $(T, \leq)$ can be viewed as a category, cf. 2.1. A *path* in T is a function $\alpha : \mathbf{N} \to \mathbf{N}$ such that $\langle\alpha(0), ..., \alpha(n)\rangle \in T$ for every $n \geq 0$. If $u \in T$ and $\langle\alpha(0), ..., \alpha(n)\rangle = u$ for some n then we say that α *goes through* u. A *bar* for $u \in T$ is a set $B \subseteq T$ of extensions of u such that every path in T which goes through u also goes through some $v \in B$. Define a Grothendieck topology $J = \mathrm{Beth}$ on the category T by setting for each $u \in T$ and each $S \subseteq \underline{y}(u)$,

$$S \in \mathrm{Beth}\ (u) \Leftrightarrow \quad \text{there is a bar } B \text{ for } u \text{ such that } (v \to u) \in S \text{ for each } v \in B.$$

The resulting sheaf semanties is exactly that of Beth (who in fact only considered *constant* sheaves on T, cf. I. 2.10).

2.3 Boolean-valued models. Let $\mathbf{B} = (B, \leq, \wedge, \vee, 0, 1)$ be a complete Boolean algebra (cBa). $\mathbf{B}$ is a poset, hence a category. Define a Grothendieck topology J on $\mathbf{B}$, called the *supremum-topology* by

$$S \in J(b) \text{ iff } b = \bigvee\{c \in B \mid (c \to b) \in S\},$$

for any $b \in B$ and any sieve S on b. Sheaf semantics over the site $(\mathbf{B}, J)$ is essentially equivalent to the theory of Boolean-valued models.

If $\mathbf{P}$ is a poset, equipped with the $\neg\neg$-topology (cf. II 1.7), then the subsheaves of $1 \in Sh(\mathbf{P}, \neg\neg)$ form a cBa $\mathbf{B}$, and

$$Sh(\mathbf{P}, \neg\neg) = Sh(\mathbf{B}, \mathrm{sup})$$

where sup is the supremum-topology on $\mathbf{B}$. This explains the correspondence between Cohen-type forcing and Boolean-valued models.

2.4 Heyting-valued models. A complete Heyting algebra (cHa) is a distributive complete lattice $\mathbf{A} = (A, \leq, \wedge, \vee, 0, 1)$ for which in addition, the infinite distributive law

$$a \wedge \bigvee_{i \in J} b_i = \bigvee_{i \in I} a \wedge b_i \qquad (1)$$

holds. Implication $\Rightarrow$ is then defined as $a \Rightarrow b = \bigvee\{x \in A \mid x \wedge a \leq b\}$. Here A is a poset, hence a category, and we can define a supremum-topology on $\mathbf{A}$ exactly as for cBa's. (The identity (1) corresponds precisely to the stability axiom for a Grothendieck topology.) The resulting semantics is the semantics of "*Heyting-valued sets*". The phrase "*topological models*" refers to the special case where $\mathbf{A}$ is the lattice $\mathcal{O}(X)$ of open subsets of a topological space X. The sup-topology is then precisely the Grothendieck topology on $\mathcal{O}(X)$ defined in I. 1.5. One generally writes

$$Sh(X) = Sh(\mathcal{O}(X), \mathrm{sup}).$$

Topological models go back to Scott (1968), (1970); for more on Heyting-valued models, see Fourman-Scott (1979), Fourman-Hyland (1979), Grayson (1979), (1981), (1982).

2.5 Permutation-models. Let T be a topological space, and let G be a group acting continuously on T; that is, there is a map $G \times T \to T, (g, t) \mapsto g \cdot t$, which is continuous for the discrete topology on G, and satisfies the usual identities

$$1 \cdot t = t \quad g \cdot (h \cdot t) = (gh) \cdot t.$$

Consider the category T_G whose objects are the open subsets $U \subseteq T$, and whose arrows $U \to V$ are elements $g \in G$ with the property that $g \cdot t \in V$ for any $t \in U$. Define a Grothendieck topology J on T_G by setting for an open subset $U \subseteq T$ and a sieve $S \hookrightarrow \underline{y}(U)$,

$$S \in J(U) \text{ iff } U = \{g \cdot t \mid (g : V \to U) \in S(V), t \in V\}$$

$Sh(T_G, J)$ is the category of so-called (G)-equivariant sheaves on T.

For a *topological group* G acting on a space T, it is again possible to construct a site for the G-equivariant sheaves, but the construction (explicitly described in Moerdijk (1988)) is somewhat more involved. One can generalize further and consider, instead of the open subsets of a topological space T, the elements of a complete Heyting algebra A. In other words, instead of spaces one takes the more general notion of a locale, see §V. 3.

2.6 General remarks. Freyd (1979) showed that in some sense, models of the form $Sh(\mathsf{A}_G) =$ (G-equivariant sheaves on A), where G is a topological group acting on a complete Heyting algebra, are enough. Indeed, if (C, J) is an arbitrary site, then there exists an embedding functor $Sh(P\mathsf{C}, J) \hookrightarrow Sh(\mathsf{A}_G)$, for suitable such A and G. This embedding preserves all the relevant structure: products, exponentials, power-objects, etc. If one replaces "group" by "groupoid", then this embedding functor can be replaced by an equivalence; see Joyal-Tierney (1984), Joyal-Moerdijk (1990) for more precise statements. In the latter paper, one also finds the result that for any site (C, J) there is an embedding $Sh(\mathsf{C}, J) \hookrightarrow Sh(\mathsf{A})$, for a suitable *cHa* A, which preserves e.g. products and exponentials, but not powersheaves.

I mentioned that by Kripke's result, sheaf semantics is complete for first order logic (see 2.1). However, the standard interpretation for sheaves on a site is not complete for HHA. It is an interesting open problem to describe the set of HHA-sentences which are valid in all sheaf-models. (For instance, it is a folklore fact that first order *classical* Peano arithmetic is valid. Also, in sheaf-models every small complete category is a pre-order, although this is not provable in HHA (cf. *CWM*. p , Hyland (1988).) In the context of analysis, one has e.g. that "every Cauchy-sequence has a norm of convergence" is valid, but not provable in HHA (Moerdijk 1982, unpublished).

In these lectures, I have not considered the more general notion of an *elementary topos* (Lawvere-Tierney). Every Grothendieck topos is an elementary topos, but not conversely. For elementary topoi there is a trivial completeness proof, since the "term-model" of HHA is an elementary topos, the so-called *free topos* (see Lambek-Scott (1986)).

Lecture III: Freyd's proof of the independence of the axiom of choice

In this lecture, I shall prove the following theorem:

Theorem 1. *There exists a Grothendieck topos $\mathcal{F} = Sh(\mathsf{F}, J)$ and a sequence of sheaves $E_n (n \in \mathbf{N})$ on F such that*
(i) *The logic of $\mathcal{F}$ is classical* (cf. II. 1.7).
(ii) *For each n, the morphism $E_n \to 1$ is an epimorphism.*
(iii) *The product $\prod_{n \in \mathbf{N}} E_n$ is isomorphic to the zero-sheaf.*
(iv) *For all n, E_n is (isomorphic to) a subsheaf of $\mathcal{P}(N)$.*

(Here $N = \Delta(\mathbf{N})$ is the standard interpretation of the natural numbers in $\mathcal{F}$.) By translating these categorical properties of $\mathcal{F}$ via the forcing definition, one obtains:

Corollary 2. *The following sentence is valid in $\mathcal{F}$ (for the standard interpretation of (classical) higher order logic):*

$$\exists E \in \mathcal{P}(N \times \mathcal{P}N)[\forall n \in N \exists x \in \mathcal{P}N (n, x) \in E \wedge \neg(\exists f \in \mathcal{P}(N)^N \, \forall n \in N (n, f(n)) \in E)]$$

Thus, the axiom of choice is not provable in classical type theory. Since all the objects involved ($N, \mathcal{P}(N), \mathcal{P}(N)^N$, etc.) are contained in the cummulative hierarchy, theorem 1 also implies the independence from ZF; see II. 1.6. (In fact for the special case of the topos $\mathcal{F}$, any object of $\mathcal{F}$ is contained in some V_α.)

I have divided this lecture into three parts. The first section deals with some generalities concerning sheaves (and is really a continuation of §I. 2). In section 2 I will give a proof of theorem 1, and in a third section I will show how this theorem can be translated into the independence of the axiom of choice as stated in corollary 2.

The original source for this lecture is Freyd (1980). In the references, I have also listed a useful exposition of Freyd's method which has recently been published by Blass and Scedrov.

§1. More sheaf theory.

1.1 Monomorphisms. In any category, an arrow f is called a monomorphism if for any two arrows g and h which both have the domain of f as their codomain, $fg = fh$ implies $g = h$; cf. *CWM* p 9. It is not difficult to prove that in the category $Sh(\mathsf{C}, J)$ of sheaves on a site, a morphism $\tau : F \to G$ between sheaves is mono iff $\tau_C : F(C) \to G(C)$ is injective for each C. (For the purpose of these lectures, one may also take this as the definition of "monomorphism of sheaves".)

Let $\{\tau_i : F_i \to G_i\}_i$ be a family of monomorphisms of sheaves. There is an induced morphism $\tau = \sum \tau_i$: $\sum F_i \to \sum G_i$ (cf. I 2.9), and it follows by the explicit description of sums given above that τ is mono if each τ_i is.

1.2 Epimorphisms. An arrow f in some category is called an epimorphism if for any two arrows g and h with the codomain of f as their domain, $gf = hf$ implies $g = h$ (cf. CWM, p. ??). For instance, in the category of sets an epimorphism is just a surjection.

It follows immediately from the universal property of sums (I. 2.9) that if $\tau_i : F_i \to G_i$ is an epimorphism of sheaves for each i, then so is the induced morphism $\tau : \sum F_i \to \sum G_i$.

Epimorphisms of sheaves can explicitly be described as follows:

Lemma. *A morphism $\tau : F \to G$ of sheaves on a site $(\mathbf{C}, J)$ is an epimorphism in $Sh(\mathbf{C}, J)$ iff for any object C of $\mathbf{C}$ and any $x \in G(C)$, the sieve*

$$S_x = \{g : B \to C \mid x \cdot g = \tau_B(y) \text{ for some } y \in F(B)\}$$

is a cover of C. (Recall that we identify a sieve $S \subseteq \underline{y}(C)$ with a set of arrows into C.)

If $\tau : F \to G$ is any morphism of sheaves, one can define a subsheaf $\tau(F)$ of G, called the *image* of τ, by

$$\tau(F)(C) = \{x \in G(C) \mid S_x \in J(C)\}.$$

Then τ factors as $F \twoheadrightarrow \tau(F) \hookrightarrow G$, where the first arrow is an epimorphism and the second is the inclusion of a subsheaf. If τ is a monomorphism, then $F \to \tau(F)$ is an isomorphism; thus F is isomorphic to a subsheaf of G in this case. If τ is an epimorphism then $\tau(F) = G$, by the lemma.

From the preceding lemma, one easily deduces:

Lemma. *In a category $Sh(\mathbf{C}, J)$ of sheaves on a site, epimorphisms are stable under pull-back; that is, if $\tau : F \to G$ is an epimorphism then so is the projection $\pi_1 : H \times_G F \to H$, for any morphism $\sigma : H \to G$.*

$$\begin{array}{ccc} H \times_G F & \xrightarrow{\pi_2} & F \\ \pi_1 \downarrow & & \downarrow \tau \\ H & \xrightarrow{\sigma} & G. \end{array}$$

1.3 Infinite products. Let $\{F_i : i \in I\}$ be a family of sheaves on a site $(\mathbf{C}, J)$. Their product $\prod F_i$ is to be a sheaf, equipped with projections $p_i : \prod F_i \to F_i$ (for each i), such that the following universal property holds: For any other sheaf G and any family of morphisms $\tau_i : G \to F_i$, there is a unique morphism $\tau : G \to \prod F_i$ with $p_i \circ \tau = \tau_i$.

If I is finite, we have already seen in lecture I that this product can simply be constructed as the pointwise product. The same simple construction works for infinite products.

1.4 Associated sheaves. A basic theorem of sheaf theory states that the inclusion functor

$$Sh(\mathbf{C}, J) \hookrightarrow \mathrm{Sets}^{\mathbf{C}^{op}}$$

has a left adjoint, denoted $\underline{a}$. This means that for any presheaf $P : \mathbf{C}^{op} \to \mathbf{Sets}$ there is a "best approximation" of P by a sheaf. This approximation is given by a morphism $\eta : P \to \underline{a}(P)$, where $\underline{a}(P)$ is a sheaf; it is "best" in the sense that for any other morphism $\sigma : P \to F$ into a sheaf F, there is a unique $\tau : \underline{a}P \to F$ with $\tau\eta = \sigma$. The general construction of $\underline{a}(P)$ is somewhat involved, but need not detain us here, since the following special case suffices.

A presheaf P on $\mathbf{C}$ is called *separated* if it satisfies the uniqueness part of the definition of a sheaf; that is, for any covering sieve $S \hookrightarrow \underline{y}(C)$ any map $\sigma : S \to P$ has *at most* one extension $\underline{y}(C) \to P$. For a separated presheaf P, its associated sheaf $\underline{a}(P)$ can simply be described by the formula

$$\underline{a}(P)(C) = \varinjlim{}_{S \in J(C)} \underline{\mathrm{Hom}}\,(S, P). \tag{1}$$

Explicitly, an element of $\underline{a}(P)(C)$ is an equivalence class of morphisms $\alpha : S \to P$ where S is a cover of C; two such $\alpha : S \to P$ and $\alpha' : S' \to P$ are equivalent if there exist a "finer" cover $T \in J(C)$, $T \subseteq S \cap S'$, such that the restriction of α and α' to T coincide. For a morphism $f : D \to C$ in $\mathbf{C}$, the restriction $\underline{a}(P)(C) \to \underline{a}(P)(D)$ is defined by the stability axiom for Grothendieck topologies: if $[\alpha] \in \underline{a}(P)(C)$ is represented by $\alpha : S \to P$, then $[\alpha] \cdot f$ is represented by the dotted arrow as indicated in the following diagram:

$$\begin{array}{ccccc} \underline{y}(D) & \hookleftarrow & \underline{y}(f)^{-1}(S) & & \\ \underline{y}(f)\downarrow & & \downarrow & \searrow & \\ \underline{y}(C) & \hookleftarrow & S & \xrightarrow{\alpha} & P \end{array}$$

It is not hard to prove (with the transitivity axiom from I. 1.5) that $\underline{a}(P)$ thus defined is a sheaf.

The morphism $\eta : P \to \underline{a}(P)$ is described as follows. For an object $C \in \mathbf{C}$ and an element $x \in P(C)$, $\eta_C(x)$ is the equivalence class of the morphism $\tilde{x} : \underline{y}(C) \to P$ which corresponds to x by the Yoneda lemma (I. 1.2). The universal property of η as described above, is immediate.

Notice that $\eta : P \to \underline{a}(P)$ is a monomorphism, precisely because P is separated. The map η is an isomorphism $P \cong \underline{a}(P)$ iff P is already a sheaf.

§2. The Freyd topos.

Let $\mathbf{F}$ be the category whose objects are the finite non-empty ordered sets

$$\underline{n} = \{0, 1, 2, ..., n\} \quad (n \geq 0)$$

and whose morphisms are the retractions; that is, an arrow $f : \underline{n} \to \underline{m}$ in $\mathbf{F}$ is a function such that $f(i) = i$ for $i \leq m$. (In particular, there are no morphisms $\underline{n} \to \underline{m}$ unless $n \geq m$.) We consider the dense topology ($\neg\neg$-topology) on $\mathbf{F}$, cf. II. 1.7. So a sieve $S \hookrightarrow \underline{y}(\underline{m})$ covers $\underline{m}$ iff for any arrow $g : \underline{n} \to \underline{m}$ in $\mathbf{F}$ there exists an $h : \underline{k} \to \underline{n}$ with $gh \in S$. I shall write

$$\mathcal{F} = Sh(\mathsf{F}, \neg\neg)$$

and call $\mathcal{F}$ the Freyd topos. For this topos $\mathcal{F}$, I will now prove theorem 1.

Notation: I shall write $P_n = \underline{y}(\underline{n}) = \mathrm{Hom}_{\mathsf{F}}(-, n)$ for the representable presheaf corresponding to an object $\underline{n}$ of F, and $E_n = \underline{a}(P_n)$ for its associated sheaf. E_n can be constructed by (1) of III. 1.4 since P_n is separated, as stated in the lemma below.

2.1 Lemma. *The site* $(\mathsf{F}, \neg\neg)$ *has the following properties:*
(a) *Any covering sieve is non-empty.*
(b) *Any non-empty sieve on* $\underline{0}$ *is a cover.*
(c) *Each presheaf* P_n *is separated;* P_0 *is a sheaf (so* $P_0 = E_0 = 1$*).*
(d) *There are only two subsheaves of* E_0 *(viz. the zero-sheaf and* E_0 *itself).*

Proof. (a) If $S \hookrightarrow \underline{y}(\underline{m})$ is a $\neg\neg$-cover, then by taking for g above the identity, we find an arrow $h \in S$.
(b) Suppose S is a sieve on $\underline{1}$, and $s : \underline{k} \to \underline{1}$ is an arrow in S. Then for any other arrow $g : \underline{n} \to \underline{1}$, and for any arrows $h : \underline{n+k} \to \underline{n}$ and $\ell : \underline{n+k} \to \underline{k}$, we have $gh = s\ell \in S$ since S is a sieve. Thus S is a cover of $\underline{1}$.
(c) Consider a diagram

$$S \overset{i}{\hookrightarrow} \underline{y}(\underline{m}) \underset{\psi}{\overset{\varphi}{\rightrightarrows}} P_n$$

such that $\varphi \circ i = \psi \circ i$, where i is the inclusion of a covering sieve of $\underline{y}(\underline{m})$. We need to show that $\varphi = \psi$. Pick any arrow $g : \underline{k} \to \underline{m}$ in F, and consider the arrows $\varphi_{\underline{k}}(g)$ and $\psi_{\underline{k}}(g) : \underline{k} \to \underline{n}$. Since S is a cover, there exists an $h : \underline{\ell} \to \underline{k}$ with $gh \in S$. Thus

$$\begin{aligned}
\varphi_{\underline{k}}(g) \circ h &= \varphi_{\underline{\ell}}(gh) && \text{(natuarality of } \varphi) \\
&= \varphi_{\underline{\ell}}(i_{\underline{\ell}}(gh)) && (gh \in S) \\
&= \psi_{\underline{\ell}}(i_{\underline{\ell}}(gh)) && (\varphi i = \psi i) \\
&= \phi_{\underline{\ell}}(gh) && (i \text{ is inclusion}) \\
&= \psi_{\underline{k}}(g)h. && \text{(naturality of } \psi)
\end{aligned}$$

But h is a surjection, so $\varphi_{\underline{k}}(g) = \psi_{\underline{k}}(g)$. Since g is arbitrary, we conclude that $\varphi = \psi$.
(d) Let $U \subseteq 1 = P_1 = F_1$ be a subsheaf; that is, U is a closed subpresheaf (cf. I. 2.5). In particular, if U contains a covering sieve S on $\underline{1}$, then by closedness $U = 1$. By (b) above, it follows that $U = \emptyset$ or $U = 1$.

2.2 Proposition. *The unique arrow* $p : E_n \to 1$ *is an epimorphism, for each* $n \geq 0$.

Proof. Let $p(E_n) \subseteq 1$ be the image of p, as in 1.2 above. Then $p(E_n)$ is non-empty since E_n is non-empty. By 2.1(d), $p(E_n) \tilde{\to} 1$; thus $E_n \to 1$ is an epimorphism.

2.3 Proposition. *If $n > m$ then $E_n(\underline{m}) = \emptyset$.*

Proof. By §1.4 above, an element of $E_n(\underline{m})$ is an equivalence class of morphisms $\tau : S \to P_n$, where S is a covering sieve of $\underline{m}$. I claim that such a τ doesn't exist, when $n > m$. Suppose to the contrary that we have such a $\tau : S \to P_n$. Since $S \neq \emptyset$ (2.1(a)), we can pick an $s : \underline{k} \to \underline{m}$ in S. Consider

$$f = \tau_{\underline{k}}(s) \in P_n(\underline{k}), \text{ i.e. } f : \underline{k} \to \underline{n}.$$

Let $g, h : \underline{k+1} \to \underline{k}$ be the arrows in F with $g(k+1) = n$ and $h(k+1) = s(n) \leq m < n$. Then $sg = sh$. But by naturality of τ,

$$fg = \tau_{\underline{k}}(s)g = \tau_{\underline{k+1}}(sh) = \tau_{\underline{k}}(s)h = fh.$$

This contradicts the fact that $fg(k+1) = f(n) = n$ while since $s(n) < n$ we have $fh(k+1) = fs(n) = s(n)$.

2.4 Corollary. *$\prod_n E_n$ is the zero-sheaf.*

Proof. If $(\prod_n E_n)(\underline{m}) \neq \emptyset$, then so is $E_{m+1}(\underline{m})$, since there is a projection $\prod_n E_n \to E_{m+1}$. This contradicts 2.3.

2.5 Proposition. *For each natural number n there is a monomorphism $E_n \rightarrowtail \mathcal{P}(N)$.*

Proof. Fix $n \geq 0$. Recall that $N = \Delta(\mathsf{N})$ is the constant sheaf as described in I. 2.10, and that $\mathcal{P}(N)$ is the corresponding powersheaf. If we can construct a morphism $\mu : P_n \to \mathcal{P}(N)$ then it extends uniquely to a morphism $\bar{\mu} : E_n \to \mathcal{P}(N)$, by the universal property of $\eta : P_n \to \underline{a}(P_n) = E_n$, since $\mathcal{P}(N)$ is a sheaf. It is easy to see from the explicit description of $\underline{a}$ given in III. 1.4 that $\bar{\mu}$ is mono if μ is.

To construct a monomorpshism $\mu : P_n \to \mathcal{P}(N)$, first enumerate without repetitions all arrows in F with codomain $\underline{n}$, say as $\{g_k : k \in \mathsf{N}\}$. Each g_i generates a smallest closed subpresheaf C_i of P_n, as

$$(h : \underline{m} \to \underline{n}) \in C_i(\underline{m}) \text{ iff there is cover } S \in \neg\neg(\underline{n}) \text{ such that}$$
$$\forall s \in S \exists t : gt = hs.$$

Now for an object $\underline{m}$ of F, the elements of $\mathcal{P}(N)(\underline{m})$ are the closed subpresheaves of $\underline{y}(\underline{m}) \times N = P_m \times N = P_m \times \Delta(\mathsf{N})$. An element of $(P_m \times \Delta(\mathsf{N}))(\underline{k})$ is given by an arrow $h : \underline{k} \to \underline{m}$ and a disjoint family of sieves S_i on $\underline{k}$ such that $\cup_{i \in \mathsf{N}} S_i$ covers $\underline{k}$. Now define

$$\mu_{\underline{m}} : P_n(\underline{m}) \longrightarrow \mathcal{P}(N)(\underline{m}) = CSubp(P_m \times \Delta(\mathsf{N}))$$

as follows. For an arrow $f : \underline{m} \to \underline{n}$, $\mu_{\underline{m}}(f)$ is the subpresheaf of $P_m \times \Delta(\mathsf{N})$ given by

$(h,\{S_i\}) \in \mu_{\underline{m}}(f)(\underline{k})$ iff for each arrow $s \in S_i$, $fhs \in C_i$.

This is clearly a closed subpresheaf of $P_m \times \Delta(\mathbf{N})$. Two things remain to be shown: that μ is a natural transformation, and that μ is a monomorphism. Leaving the first statement for the reader to verify, we conclude by proving that each $\mu_{\underline{m}}$ is injective.

Suppose $f, f' : \underline{m} \to \underline{n}$ are distinct arrows such that $\mu_{\underline{m}}(f) = \mu_{\underline{m}}(f')$. These arrows f, f' occur in the enumeration, say as g_j and $g_{j'}$. Consider $\xi = (id, \{S_i\}) \in (P_m \times \Delta(\mathbf{N}))(\underline{m})$, where $S_i = \emptyset$ if $i \neq j$, while S_j is the maximal sieve $\underline{y}(\underline{m})$ on $\underline{m}$. Then clearly $\xi \in \mu_{\underline{m}}(f)(\underline{m})$. On the other hand, $\xi \in \mu_{\underline{m}}(f')(\underline{m})$ would mean that $f' \in C_j$, i.e. that $C_j \cap C'_j \neq \emptyset$. By 2.1(a), this implies that there is at least one commutative square in $\mathbf{F}$ of the form

$$\begin{array}{ccc} \underline{l} & \xrightarrow{s'} & \underline{m} \\ s \downarrow & & \downarrow f' = g_{j'} \\ \underline{n} & \xrightarrow{f = g_j} & \underline{n} \end{array}$$

But for $i \leq m$ we have $s(i) = i = s'(i)$. Hence by commutativity $f(i) = f(s(i)) = f'(s'(i)) = f'(i)$. This contradicts $f \neq f'$.

§3. Logical aspects of the construction.

I will now show that corollary 2 follows from theorem 1 as stated in the beginning of this lecture. Consider for each $n \in \mathbf{N}$ the monomorphism $\mu_m : E_n \rightarrowtail \mathcal{P}(N)$ of proposition 2.5. These can be summed up to give a monomorphism $\sum E_n \to \sum_n \mathcal{P}(N)$. Now by the distributive law of I. 2.9, $\sum_n \mathcal{P}(N) \cong \sum_n (1 \times \mathcal{P}(N)) \cong (\sum_n 1) \times \mathcal{P}(N) = N \times \mathcal{P}(N)$. Thus if we write $E = \sum E_n$ then the sum of the μ_n's is a monomorphism

$$E \rightarrowtail N \times \mathcal{P}(N). \tag{1}$$

In other words, E is (isomorphic to) a subsheaf of $N \times \mathcal{P}(N)$.

Since each $E_n \to 1$ is epi by 2.2, so is the induced map $E \to N$ (cf. 1.2); thus we have a diagram

$$\begin{array}{ccc} E & \xrightarrow{\text{mono}} & N \times \mathcal{P}(N) \\ & \text{epi} \searrow & \downarrow \text{epi} \\ & & N \end{array} \tag{2}$$

The sheaf E can be identified with an element of $\mathcal{P}(N \times \mathcal{P}(N))(1)$, and we claim that for this element,

$$\underline{1} \Vdash \forall n \in N \exists x \in \mathcal{P}N (n, x) \in E \tag{3}$$

$$\underline{1} \Vdash \neg(\exists f \in \mathcal{P}(N)^N \forall n \in N (n, f(n)) \in E). \tag{4}$$

These together yield that in the Freyd topos $\mathcal{F}$, the following instance of the axiom of choice fails:

$$\forall E \in \mathcal{P}(N \times \mathcal{P}(N))(\forall n \exists x (n, x) \in E \to \exists f \forall n (n, f(n)) \in E).$$

Proof of (4). Suppose to the contrary that for some object $\underline{k}$ of $\mathbf{F}$, and some morphism $F : \underline{y}(\underline{k}) \times N = P_k \times N \to \mathcal{P}(N)$, it holds that $\underline{k} \Vdash \forall n \in N(n, F(n)) \in E$. In particular, each natural number i gives an element $\hat{i} \in N(\underline{k})$, and $\underline{k} \Vdash (\hat{i}, F(\hat{i})) \in E$. This means that the map

$$P_k = P_k \times 1 \xrightarrow{1 \times i} P_k \times N \xrightarrow{(\pi_2, F)} N \times \mathcal{P}(N)$$

factors through E. Hence $F \circ (1 \times i)$ factors through $E_i \rightarrowtail \mathcal{P}(N)$ for each i, say as θ_i:

$$\begin{array}{ccc} P_k & \xrightarrow{F \circ (1 \times i)} & \mathcal{P}(N) \\ & \searrow^{\theta_i} & \downarrow \\ & & E_i \end{array}$$

These θ_i together give a map $P_k \to \Pi E_i$, i.e. an element in $(\Pi E_i)(\underline{k})$. This contradicts 2.4.

Proof of (3). Let $\underline{k}$ be any object of $\mathbf{F}$, and let $\alpha \in N(\underline{k})$, i.e. α is a map $\underline{y}(\underline{k}) = P_k \xrightarrow{i} N$. We need to show that $\underline{k} \Vdash \exists x \in \mathcal{P}(N)(\alpha, x) \in E$. Consider the pullback square

$$\begin{array}{ccc} P_k \times_N E & \to & E \\ \pi_1 \downarrow & & \downarrow \pi_2 \\ P_k & \xrightarrow{\alpha} & N \end{array}$$

By 1.2, the projection $\pi_1 : P_k \times_N E \to P_k$ is epi. By 1.2 again, this implies that $id \in P_k(\underline{k})$ is "locally" in the image of π_1. Thus there is a cover S of $\underline{k}$ such that for each $s : \underline{\ell} \to \underline{k}$ in S there is an $e \in E(\underline{\ell})$ such that $(s, e) \in (P_k \times_N E)(\underline{\ell})$; i.e. $\alpha_{\underline{\ell}}(s) = (\pi_2)_{\underline{\ell}}(e)$. But then e is of the form $e = (\alpha \circ \underline{y}(s), b)$ where $b \in \mathcal{P}(N)(\underline{\ell})$, and $\underline{\ell} \Vdash (\alpha \circ \underline{y}(s), b) \in E$. Since S is a cover, it follows that $\underline{k} \Vdash \exists x \in \mathcal{P}(N)(\alpha, x) \in E$, as was to be shown.

Notice that the logic of $\mathcal{F}$ is classical, as a general property of the $\neg\neg$ - topology (§II. 1.7). The theory of $\mathcal{F}$ is also complete, in the sense that for any formula φ without free variables, either $\mathcal{F} \models \varphi$ or $\mathcal{F} \models \neg\varphi$. Indeed, by 2.1 (d) the subsheaf $\{\cdot \mid \varphi\}$ of 1 is either empty or equal to 1 itself.

Lecture IV: All functions are continuous

Let R be the set of Dedekind reals, definable in HHA. The purpose of this lecture is to prove

Theorem. *The statement "all functions from R to R are continuous" is consistent with HHA.*

In fact we shall exhibit a large family of toposes in which "all functions $R \to R$ are continuous" is valid. Our proof will be completely constructive.

§1. Some topological sites.

I shall be interested in subcategories $\mathbf{T}$ of the category of topological spaces, which have some (or all) of the following properties:

(a) if $X \in \mathbf{T}$, and U is an open subspace of X, then U is an object of $\mathbf{T}$ and the inclusion $U \hookrightarrow X$ is an arrow in $\mathbf{T}$;

(b) $\mathbf{T}$ is closed under products;

(c) $\mathbf{T}$ is *full* subcategory of the category of topological spaces;

(d) $\mathbf{R} \in \mathbf{T}$ ($\mathbf{R}$ is the real line with the usual topology).

If $\mathbf{T}$ is a category satisfying (a) then we can define a Grothendieck topology J on $\mathbf{T}$, by setting for $T \in \mathbf{T}$ and $S \hookrightarrow \underline{y}(T)$,

$$S \in J(T) \text{ iff } T = \cup\{U \mid U \text{ is an open subset of } T, \text{ and } (U \hookrightarrow T) \in S\}.$$

For various such $\mathbf{T}$, we shall *only* consider this Grothendieck topology on $\mathbf{T}$, and denote the category of sheaves by $Sh(\mathbf{T})$.

If $\mathbf{T}$ satisfies (a) and X is *any* space (not necessarily an object from $\mathbf{T}$), then there is a sheaf $C(X)$ on $\mathbf{T}$, defined by

$$C(X)(T) = Cts(T, X), \quad \text{for } T \in \mathbf{T}. \tag{1}$$

Here $Cts(X,T)$ is the set of all continuous maps from T into X. This defines a functor

$$C : (\text{spaces}) \to Sh(\mathbf{T}). \tag{2}$$

Notice that if $X \in \mathbf{T}$ and $\mathbf{T}$ also satisfies (c), then

$$C(X) = \underline{y}(X) = \mathbf{T}(-, X). \tag{3}$$

We call $C(X)$ the *sheaf represented by* X.

The functor C preserves certain exponentials:

1.1 Lemma. *Suppose* $\mathbf{T}$ *satisfies* (a), (b), (c). *Let T and X be spaces such that T is locally compact and in* $\mathbf{T}$. *Then there is an isomorphism of sheaves*

$$C(X)^{C(T)} \cong C(X^T)$$

(here X^T is the space of continuous maps, with the compact-open topology).

Proof. This is a trivial application of the Yoneda lemma: for any other object U of $\mathbf{T}$ we have

$$\begin{aligned}
C(X)^{C(T)}(U) &= \mathrm{Hom}(\underline{y}(U), C(X)^{C(T)}) && \text{(Yoneda)}\\
&\cong \mathrm{Hom}\,(\underline{y}(U) \times C(T), C(X)) && \text{(univ. prop. of expo.)}\\
&\cong \mathrm{Hom}\,(\underline{y}(U) \times \underline{y}(T), C(X)) && \text{(by (3))}\\
&\cong \mathrm{Hom}\,(\underline{y}(U \times T), C(X)) &&\\
&\cong Cts(U \times T, X) && \text{(def. of } C)\\
&\cong Cts(U, X^T) && \text{(standard prop. of compact} - \text{open top.)}\\
&\cong C(X^T)(U) && \text{(def. of } C\text{).}
\end{aligned}$$

§2. Representability of Dedekind reals.

In this section, $\mathbf{T}$ is a category satisfying (a) of §1. Recall the standard interpretation $N = \Delta(\mathbf{N})$ of the natural numbers in any topos. From N, one can define the object Q of rationals in the usual way as equivalence classes of pairs of natural numbers (say as $< n, m >= n/m$ if m is even and $m \neq 0$, while $< n, m >= -n/m + 1$ if m is odd). Next, one can define the set of Dedekind reals

$$R \subseteq \mathcal{P}(Q) \times \mathcal{P}(Q)$$

as follows: for variables L, U of type $\mathcal{P}(Q)$, let $Cut(L, U)$ be the conjunction of the following formulas:

(i) $\neg\exists q \in Q(q \in U \wedge q \in L)$

(ii) $\exists r \in Q(r \in U)$, $\exists q \in Q(q \in L)$

(iii) $\forall q, r \in Q(q < r \wedge r \in L \Rightarrow q \in L)$ $\quad \forall q, r \in Q(r < q \wedge r \in U \Rightarrow q \in U)$

(iv) $\forall q \in Q(q \in L \Rightarrow \exists r \in Q(r \in L \wedge q < r))$ $\forall q \in Q(q \in U \Rightarrow \exists r \in Q(r \in U \wedge r < q))$

(v) $\forall q, r \in Q(q < r \Rightarrow (q \in L \vee r \in U))$.

Now define

$$R = \{(L, U) \in \mathcal{P}(Q) \times \mathcal{P}(Q) \mid Cut(L, U)\}.$$

2.1 Proposition. *For the standard interpretation in $Sh(\mathbf{T})$, N and Q are (interpreted as sheaves) isomorphic to the representable sheaves $C(\mathbf{N}_{\mathrm{dis}}), C(\mathbf{Q}_{\mathrm{dis}})$.*

Here $\mathbf{N}_{\mathrm{dis}}$ and $\mathbf{Q}_{\mathrm{dis}}$ denote the natural numbers and the rationals, both equipped with the discrete topology. This proposition is rather easy to verify, and I prefer to concentrate on the more interesting result concerning R.

Proposition. *For the standard interpretation in $Sh(\mathbf{T})$, the sheaf of Dedekind reals R is isomorphic to the representable sheaf $C(\mathbf{R})$.*

Proof. For an object W of the site $\mathbf{T}$, an element of $R(W)$ consists of a pair (L, U), where L and U are elements of $\mathcal{P}(Q)(W)$; that is, L and U are subsheaves of $\underline{y}(W) \times Q =$

$\mathsf{T}(-,W)\times C(\mathbf{Q}_{\text{dis}})$ (cf. prop. 1 above, and I. 2.5; notice that $\underline{y}(W)$ is a sheaf here). Thus L consists of pairs (α, p) where $\alpha : Y \to W$ is a map in T and $q : Y \to \mathbf{Q}$ is locally constant. To say that L is a subsheaf means, first that $(\alpha, p) \in L$ implies $(\alpha\beta, p\beta) \in L$ for any map $\beta : Z \to Y$ in T, and secondly that if $(\alpha \mid V_i, p \mid V_i) \in L$ for each V_i in some open cover $\{V_i\}$ of Y, then $(\alpha, p) \in L$. The same applies to U.

Now to say that (L, U) lies in the subsheaf $R \subseteq \mathcal{P}(Q) \times \mathcal{P}(Q)$ means that $W \Vdash Cut(L, U)$. We can spell this out by unwinding the forcing definitions. This gives:

(i') For any map $\beta : W' \to W$ and any locally constant function $q : W' \to \mathbf{Q}$, not both $(\beta, q) \in L(Q')$ and $(\beta, q) \in U(W')$.

(ii') There is an open cover $\{W_i\}$ of W such that for each i there are locally constant functions $q_i, r_i : W_i \to \mathbf{Q}$ with $(W_i \hookrightarrow W, q_i) \in L(W_i)$ and $(W_i \hookrightarrow W, r_i) \in U(W_i)$.

(iii') For any map $\beta : W' \to W$ in T and any locally constant functions $q, r : W' \to \mathbf{Q}$: if $q(x) < r(x)$ for all $x \in W'$ and $(\beta, r) \in L(W')$ then $(\beta, q) \in L(W')$; and if $r(x) < q(x)$ for all $x \in W'$ and $(\beta, r) \in U(W')$ then $(\beta, q) \in U(W)$.

(iv') For any map $\beta : W' \to W$ in T and any locally constant function $q : W' \to \mathbf{Q}$: if $(\beta, q) \in L(W')$ (resp. $(\beta, q) \in U(W')$) then there are an open cover $\{W_i'\}$ of W' and locally constant functions $r_i : W_i' \to \mathbf{Q}$ such that $r_i(x) > q(x)$ for all $x \in W_i'$ and $(\beta \mid W_i', r_i) \in L(W_i')$ (resp. $r_i(x) < q(x)$ for all $x \in W_i'$ and $(\beta \mid W_i', r_i) \in U(W_i')$).

(v') For any $\beta : W' \to W$ in T and any two locally constant functions $q, r : W' \to \mathbf{Q}$ such that $q(x) < r(x)$ for all $x \in W'$, there exists a cover $\{W_i'\}$ of W' such that for each i, either $(\beta \mid W_i', q \mid W_i') \in L(W_i')$ or $(\beta \mid W_i', r \mid W_i') \in U(W_i')$.

Write $\hat{q}$ for the constant function with value $q \in \mathbf{Q}$. Consider for a point $x \in W$ the sets of rationals

$$L_x = \{q \in \mathbf{Q} \mid \exists \text{ open } V \subseteq W : x \in V \text{ and } (V \hookrightarrow W, \hat{q} \mid V) \in L(V)\},$$

$$U_x = \{r \in \mathbf{Q} \mid \exists \text{ open } V \subseteq W : x \in V \text{ and } (V \hookrightarrow W, \hat{r} \mid V) \in U(V)\}.$$

It readily follows from (i') - (v') above that L_x and U_x form a Dedekind cut (in the category of sets). Therefore there is a unique real number $\sup L_x = \inf U_x$. We can thus define a function

$$f_{L,U} : W \to \mathbf{R} \quad f_{L,U}(x) = \sup L_x.$$

In other words, for rationals $q, r \in \mathbf{Q}$,

$$q < f_{L,U}(x) < r \text{ iff } q \in L_x \text{ and } r \in U_x.$$

It follows that f is continuous.

Conversely, suppose we start with a given continuous function $f : W \to \mathbf{R}$. Define subsheaves L_f and U_f of $\mathsf{T}(-,W) \times Q$ by setting for $\beta : W' \to W$ in T and $p : W' \to \mathbf{Q}$ locally constant,

$$p \in L_f(W') \text{ iff } \forall x \in W' \quad p(x) < f\beta(x)$$

$$p \in U_f(W') \text{ iff } \forall x \in W' \quad p(x) > f\beta(x).$$

Then $W \Vdash Cut(L_f, U_f)$; that is, (i') - (v') above hold for L_f and U_f.
It is now a straightforward matter to spell out the definitions and check that

$$R(W) \rightleftarrows C(\mathbf{R})(W) \begin{cases} (L,U) \mapsto f_{L,U} \\ L_f, U_f \leftarrow f \end{cases}$$

are mutually inverse operations, natural in W. This proves the proposition.

From lemma 1.1 and proposition 2.2. we obtain:

2.3 Corollary. *If* $\mathbf{T}$ *also satisfies conditions* (b) - (d), *then for the standard interpretation in* $Sh(\mathbf{T})$, *there is an isomorphism of sheaves*

$$R^R \cong C(\mathbf{R}^{\mathbf{R}}).$$

§3. Continuity of all functions.

In this section, $\mathbf{T}$ is a category satisfying conditions (a) - (d) from §1. I shall now prove:

3.1 Theorem. *In* $Sh(\mathbf{T})$, *the HHA-sentence "all functions* $R \to R$ *are continuous" is valid.*

Proof. A more explicit version of the theorem states that the sentence

$$\forall f \in R^R \ \forall x \in R \ \forall \epsilon \in R \ (\epsilon > 0 \Rightarrow \exists \delta \in R(\delta > 0 \wedge$$

$$\forall y \in R \ (x - \delta < y < x + \delta \to f(x) - \epsilon < f(y) < f(x) + \epsilon)))$$

is valid in $Sh(\mathbf{T})$. By 2.2 and 2.3, R is the sheaf $C(\mathbf{R})$ and R^R is $C(\mathbf{R}^{\mathbf{R}})$, up to isomorphism. So take an object $W \in \mathbf{T}$, and $f \in C(\mathbf{R}^{\mathbf{R}})(W)$, $a, \epsilon \in C(\mathbf{R})(W)$ such that $W \Vdash \epsilon > 0$. Thus $f : W \to \mathbf{R}^{\mathbf{R}}$ and $a, \epsilon : W \to \mathbf{R}$ are continuous maps such that $\epsilon(x) > 0$ for all $x \in W$. We have to show

$$(*) W \Vdash \exists \delta \in R \ (\delta > 0 \wedge \forall y \in R(a - \delta < y < a + \delta \Rightarrow f(a) - \epsilon < f(y) < f(a) + \epsilon)).$$

(Here $f(a)$ stands for the composition of $(a, f) : W \to \mathbf{R} \times \mathbf{R}^{\mathbf{R}}$ with the evaluation map $\mathbf{R} \times \mathbf{R}^{\mathbf{R}} \to \mathbf{R}$, as before).
Now f corresponds to a continuous map $\hat{f} : W \times \mathbf{R} \to \mathbf{R}$ via

$$f(x)(t) = \hat{f}(x, t) \quad (x \in W, t \in \mathbf{R}).$$

By continuity of $\hat{f}$ and ϵ, there exists for each $x \in W$ a neighbourhood $W_x \subseteq W$ of x and a $\delta_x > 0$ such that for all $\xi \in W_x$ and all $t \in (a(x) - \delta_x, a(x) + \delta_x)$, we have both

$$|a(\xi) - a(x)| < \frac{1}{2}\delta_x \tag{1}$$

and

$$|\hat{f}(\xi,t) - \hat{f}(x,a(x))| < \frac{1}{2}\epsilon(\xi). \tag{2}$$

I claim that it follows that

$$W_x \Vdash \forall y \in R(a - \frac{1}{2}\delta_x < y < a + \frac{1}{2}\delta_x \Rightarrow f(a) - \epsilon < f(y) < f(a) + \epsilon). \tag{3}$$

To see this, choose $\beta : V \to W_x$ in $\mathbf{T}$ and $b : V \to \mathbf{R}$ such that

$$V \Vdash a\beta - \frac{1}{2}\delta_x < b < a\beta + \frac{1}{2}\delta_x. \tag{4}$$

Thus for all $\zeta \in V$,

$$|a\beta(\zeta) - b(\zeta)| < \frac{1}{2}\delta_x,$$

and hence by (1),

$$|a(x) - b(\zeta)| < \delta_x.$$

Consequently, we can substitute $\beta\zeta$ for ξ and $b(\zeta)$ for t in (2), to obtain

$$|\hat{f}(\beta(\zeta), b(\zeta)) - \hat{f}(x,a(x))| < \frac{1}{2}\epsilon\beta(\zeta).$$

Also by (2),

$$|\hat{f}(\beta(\zeta), a\beta(\zeta)) - \hat{f}(x,a(x))| < \frac{1}{2}\epsilon\beta(\zeta).$$

Thus combining the last two inequalities

$$|\hat{f}(\beta(\zeta), b(\zeta)) - \hat{f}(\beta(\zeta), a\beta(\zeta))| < \epsilon\beta(\zeta).$$

But this means that

$$V \Vdash (f \cdot \beta)(a \cdot \beta) - \epsilon \cdot \beta < (f \cdot \beta)(b) < (f \cdot b)(a \cdot \beta) + \epsilon \cdot \beta.$$

Since this holds for all $\beta : V \to W_x$ and all $b : V \to \mathbf{R}$ such that (4), it follows that (3) holds. Since the open sets W_x, for all $x \in W$, form a cover of W, $(*)$ above follows, and the proof is complete.

Lecture V. Morphisms and definability.

In this lecture, I give simple sheaf-theoretic proofs of the following results.

Theorem 1. *Let $f(x,y)$ be a formula of HHA such that HHA $\vdash$ "f is a function $\mathbf{R} \to \mathbf{R}$" (where $\mathbf{R}$ is the set of Dedekind reals). Then HHA $\vdash$ "f is continuous".*

Theorem 2. *There are no definable non-principal ultrafilters on $\mathbf{N}$ (in ZFC, or classical type theory with choice).*

The first theorem roughly states that intuitionistically, every definable function is continuous. It is known as the continuity rule (for intuitionistic higher order logic). The proof that I will present here is due to A. Joyal (unpublished). The second theorem states that there is no formula $\varphi(x)$, where x is a variable of type $\mathcal{P}N$, for which $ZFC \vdash$ "$\{x \mid \varphi(x)\}$ is a non-principal ultrafilter". This should be contrasted with the consequence of the axiom of choice that there are many non-principal ultrafilters on $\mathbf{N}$. The proof I give of theorem 2 is taken from unpublished joint work with A. Joyal. (Theorem 2 is perhaps the simplest example of a series of "there exists no definable...") - results which can be proved in essentially the same way.)

§1. Morphisms.

An important aspect of sheaf theory is the study of morphisms of sites $(\mathbf{C}, J) \to (\mathbf{C}', J')$, and how and when these give rise to various functors between the topoi $Sh(\mathbf{C}, J)$ and $Sh(\mathbf{C}', J')$. I will not develop the general theory here. Instead, I consider only some very special cases, which, however, suffice for the applications stated above.

1.1 Open embeddings. Let U be an open subspace of a topological space X, and let $i : U \to X$ be the inclusion. There is an obvious "restriction" functor

$$i^* : Sh(X) \to Sh(U),$$

defined for a sheaf F on X and an open subset $V \subseteq U$ by $i^*(F)(V) = F(V)$. This functor clearly preserves "everything", i.e. products, exponentials, powersheaves, sums,

$$i^*(F \times G) \cong i^*(F) \times i^*(G),$$

$$i^*(G^F) \cong i^*(G)^{i^*(F)},$$

$$i^*(\mathcal{P}F) \cong \mathcal{P}(i^*(F)),$$

$$i^*(\sum_\alpha F_\alpha) \cong \sum_\alpha (i^* F_\alpha),$$

etc., etc. Notice that in particular, i^* preserves the standard interpretation of HHA. (One sometimes says that i^* is a *logical functor.*)

1.2 Homeomorphisms. Let $f : Y \to X$ be a continuous map of topological spaces. There is an induced functor

$$f_* : Sh(Y) \to Sh(X),$$

defined for a sheaf F on Y and an open subset U of X by $f_*(F)(U) = F(f^{-1}U)$. In general, f_* preserves hardly anything (only limits of sheaves are preserved). Here however, I am interested in the case where f is a homeomorphism. Clearly in that case f_* commutes with products, exponentials, powersheaves, etc., as in

$$f_*(F \times G) \cong f_*(F) \times f_*(G) \ , \ f_*(G^F) \cong f_*(G)^{(f_*(F))} \ , \ f_*(\mathcal{P}F) = \mathcal{P}(f_*F),$$

and f_* preserves the natural numbers in the sense that $f_*N = N$. It follows that f_* preserves the standard interpretation of HHA. In particular, f_* preserves the sheaf of Dedekind reals: writing R_X for the sheaf of Dedekind reals on X (and similarly for Y), there is a natural isomorphism

$$f_*(R_Y) \xrightarrow{\sim} R_X. \tag{1}$$

By proposition IV. 2.2, the sheaf R_X can be identified with the sheaf $C(\mathbf{R})$ of continuous real valued functions on X (and similarly for Y). With this identification, the isomorphism (1) is given by composition with f^{-1}, as

$$Cts(f^{-1}U, \mathbf{R}) \to Cts(U, \mathbf{R}) \ , \ \alpha \mapsto \alpha \circ f^{-1} \tag{2}$$

for any open $U \subseteq X$.

For a homeomorphism $f : Y \to X$, we write

$$f^* = (f^{-1})_*.$$

So for example for the sheaf of Dedekind reals, the isomorphism

$$f^*(R_X) \xrightarrow{\sim} R_Y \tag{3}$$

is now given by composition with f.

1.3 General remarks. Any continuous map $f : Y \to X$ of topological spaces induces adjoint functors

$$f_* : Sh(Y) \leftrightarrows Sh(X) : f^*,$$

where f^* is left adjoint to f_*. It is the functor f^* which is harder to describe but more important for logic. f^* always preserves so-called (first order) geometric logic. If f is an open map, then f^* preserves all first order logic. If in addition f has locally connected fibers, then f^* preserves exponentials (for a more precise statement, see Johnstone (1984)), while if f is an etale map then f^* is a logical functor. More generally, there is a fully developed theory of morphisms between sites which induce such adjoint functors between the sheaf categories, and their behaviour at the level of logic. This leads to the theory of classifying topoi, or universal models for geometric logic. See e.g. Makkai-Reyes (1977).

§2. The continuity rule.

We now present Joyal's proof of theorem 1. Let $\varphi(x, y)$ be a formula of HHA such that HHA proves $\forall x \in \mathbf{R} \exists! y \in \mathbf{R} \varphi(x, y)$. Then by the soundness theorem, the interpretation of φ yields for every topos $\mathcal{E}$ a morphism of sheaves

$$F_{\mathcal{E}} : R_{\mathcal{E}} \to R_{\mathcal{E}},$$

where $R_{\mathcal{E}}$ is the sheaf of Dedekind reals in $\mathcal{E}$.

In particular, for the case of a topological space X there is a morphism of sheaves

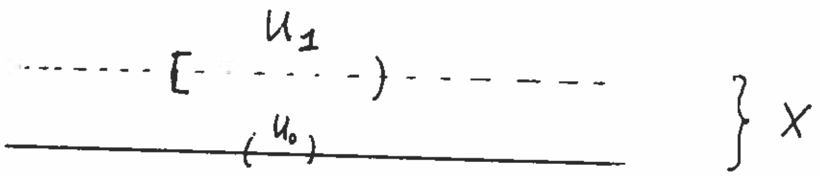
$$F_X : R_X \to R_X \tag{1}$$

where R_X is the sheaf of continuous real-valued functions on X (cf. prop. IV. 2.2). If $U \subseteq X$ is an open subspace, then since the restriction functor $i^* : Sh(X) \to Sh(U)$ of 1.1 is logical, i^* "commutes" with the calculation of the interpretation of the formula φ, so

$$F_U = (F_X) \mid U. \tag{2}$$

Let f be the "real" calculation of φ, in our own universe of sets. So f is a function $\mathbf{R} \to \mathbf{R}$, and our problem is to show that f is continuous.

Consider the space X consisting of two copies $\mathbf{R} \times \{0\}$ and $\mathbf{R} \times \{1\}$ of the real line, with the following topology: $U \subseteq \mathbf{R} \times \{0,1\}$ is open in X iff $U \cap \mathbf{R} \times \{0\}$ is open in $\mathbf{R}$ for the usual euclidean topology, and for any $x \in \mathbf{R}$: if $(x,0) \in U$ then also $(x,1) \in U$. In other words, an open subset of X is given by two subsets U_0 and U_1 of $\mathbf{R}$ where U_0 is open in $\mathbf{R}$ and U_1 is any subset which contains U_0. You should think of X as the line $\mathbf{R}$ with a copy of the discrete reals glued on top of it.



(3)

Consider $F_X : R_X \to R_X$, and write $\gamma : X \to \mathbf{R}$ for the projection. Clearly γ is continuous, so γ is an element of R_X. Therefore $F_X(\gamma)$ is again a continuous map $X \to \mathbf{R}$. In particular, its restriction to the bottom copy of $\mathbf{R}$ in X,

$$g : \mathbf{R} \to \mathbf{R}\ ,\ g(\alpha) = F_X(\gamma(\alpha, 0)) \tag{4}$$

is continuous. As for the 'discrete" top-copy $\mathbf{R} \times \{1\} \subseteq X$, we observe that

$$F_X(\gamma)(\alpha, 1) = f(\alpha) \tag{5}$$

(recall that the function f is the "real" interpretation of the formula φ). Indeed, of $\alpha \in \mathbf{R}$ then $(\alpha, 1)$ is an isolated point of X, so that the inclusion of this one-point open subset $i : \{(\alpha,1)\} = U \hookrightarrow X$ yields a logical morphism by restriction

$$Sh(X) \to Sh(U),$$

as in 1.1. Hence by (2) above,

$$F_X(\gamma) \mid U = F_U(\gamma \mid U). \tag{6}$$

But since U is a one-point space, $Sh(U)$ is the category of sets, and F_U is "'the real f", while $\gamma \mid U$ is the real number α. Thus (6) simply says that $F_X(\gamma)(\alpha, 1) = f(\alpha)$, as asserted in (5).

To see that f is continuous. it now suffices to observe that $f = g$. But more generally,

Lemma. *Let $h : X \to \mathbf{R}$ be continuous. Then for any $\alpha \in \mathbf{R}$, $h(\alpha, 0) = h(\alpha, 1)$.*

Proof. Suppose $\beta = h(\alpha, 0) \neq h(\alpha, 1)$. Then $h^{-1}(\mathbf{R} - \{\beta\})$ is an open set which contains $(\alpha, 0)$ but not $(\alpha, 1)$. There are no such open sets in X.

We have now proved that f is continuous — i.e. that our formula φ describes a continuous function in the category of sets. To obtain the desired conclusion, viz. that it is *provable* (in HHA) that φ defines a continuous function, we simply observe that the preceding argument is completely *constructive* and *explicit*, and can therefore be formalized in HHA. (A category-theorist would prefer to say: the argument can be performed in the free topos.)

§3. Locales and spaces.

I wish to come back for a moment to the Heyting- and Boolean-valued models mentioned in §II. 2, and introduce locales as generalized spaces.

The most useful notion of a morphism $F : B \to A$ if cHa's is that of a function which commutes with finite meets and arbitrary sups: that is

$$F(1) = 1, \quad F(0) = 0, \quad F(b \wedge b') = F(b) \wedge F(b'), \quad F(\bigvee b_i) = \bigvee F(b_i).$$

This defines a category *cHa* of complete Heyting algebras and such morphisms. The category of *locales* is by definition the opposite if this category. There is an obvious functor

$$\text{(spaces)} \to \text{(locales)}, \tag{1}$$

which sends a space X to the complete Heyting algebra $\mathcal{O}(X)$ of open sets in X, and a continuous map $f : X \to Y$ to the cHa morphism $f^{-1} : \mathcal{O}(Y) \to \mathcal{O}(X)$, which is thus a morphism of locales in the other direction.

One should think of locales as *generalized spaces*. In developing the theory, it is extremely useful to employ a suggestive notation: One often denotes the objects of the category of locales by $X, Y, \ldots$ and the morphisms by $f : X \to Y$ etc. For a locale X, the corresponding cHa is denoted $\mathcal{O}(X)$, while $f^{-1} : \mathcal{O}(Y) \to \mathcal{O}(X)$ denotes the cHa-morphism corresponding to a locale morphism $f : X \to Y$. One sometimes even goes as far as calling the elements of $\mathcal{O}(X)$ *opens* of the locale X, although they are not open subsets of X in any sense. Note that for a locale X, what is *given* is really a cHa $\mathcal{O}(X)$; the notation X is just a formal device to emphasize that we wish to regard $\mathcal{O}(X)$ not in the category of cHa's, but in that of locales.

For locales X and Y, I shall write

$$Cts(X, Y) \tag{2}$$

for the set of "continuous maps" $X \to Y$, that is for the set of cHa-morphisms $\mathcal{O}(Y) \to \mathcal{O}(X)$. If for instance Y is a topological space, then $Cts(X,Y)$ of course denotes the set of locale-maps $X \to Y$, where we identify Y with its image under the functor (1).

A morphism $j : X \to Y$ of locales is called an *embedding* if the corresponding cHa-morphism $j^{-1} : \mathcal{O}(Y) \to \mathcal{O}(X)$ is surjective; X is then called a sublocale of Y. Two examples of such are important for us. First, if $U \in \mathcal{O}(X)$, then the set $\{V \in \mathcal{O}(X) \mid V \leq U\}$ is again a cHa, hence defines a locale which we denote again by U. So

$$\mathcal{O}(U) = \{V \in \mathcal{O}(X) \mid V \leq U\} \tag{3}$$

(where on the right U denotes an element of $\mathcal{O}(X)$, and on the left U is a formal symbol to denote the locale whose cHa of opens is described by the right-hand side of (3).) The embedding

$$j : U \to X$$

of locales "is" the cHa-morphism

$$j^{-1} : \mathcal{O}(X) \to \mathcal{O}(U)\ ,\ j^{-1}(V) = V \wedge U.$$

Sublocales of this form are called *open sublocales.*

Another example is the embedding

$$i : X_{\neg\neg} \to X.$$

where the locale $X_{\neg\neg}$ is defined by the cHa

$$\mathcal{O}(X_{\neg\neg}) = \{V \in \mathcal{O}(X) \mid \neg\neg V = V\}.$$

(Here $\neg$ is the pseudo-complement of the cHa $\mathcal{O}(X)$: $\neg U = \bigvee\{W \in \mathcal{O}(X) \mid W \wedge U = 0\}$.) The embedding $i : X_{\neg\neg} \to X$ is given by the surjective cHa-morphism

$$i^{-1} : \mathcal{O}(X) \to \mathcal{O}(X_{\neg\neg}), i^{-1}(U) = \neg\neg U.$$

Notice that in the cHa $\mathcal{O}(X_{\neg\neg})$, meets are computed as in $\mathcal{O}(X)$, but sups are not. More precisely, if we write $i_* : \mathcal{O}(X_{\neg\neg}) \to \mathcal{O}(X)$ for the inclusion function, then $i_*(U \wedge V) = i_*U \wedge i_*V$, but $i_*(\bigvee U_i) = \neg\neg \bigvee U_i$. It is well-known and easy to check that $\mathcal{O}(X_{\neg\neg})$ is in fact a *complete Boolean algebra.*

For a locale X, we shall write $Sh(X)$ for the category of sheaves on X; that is, for the category of sheaves on the cHa $\mathcal{O}(X)$ with the "supremum" Grothendieck topology. An example of such a sheaf is the sheaf R_X of real-valued continuous maps: for $U \in \mathcal{O}(X)$,

$$R_X(U) = Cts(U, \mathbf{R}),$$

where on the right U stands for the open sublocale U of X as just defined, and $Cts(U,\mathbf{R})$ is the set of locale-maps $f : U \to \mathbf{R}$. It is clear that R_X is a sheaf.

The proof of proposition IV. 2.2 can now be "translated" from the language of topological spaces to that of locales, and gives

3.1 Proposition. *For the standard interpretation in $Sh(X)$, the sheaf of Dedekind reals is (isomorphic to) the sheaf R_X.*

The Cantor space $2^{\mathbf{N}}$ is a subspace of $\mathbf{R}$, and this definition of $2^{\mathbf{N}}$ can be "copied" in HHA so as to give an isomorphism of 2^N with a subset of the Dedekind reals R. Proposition 3.1 then yields

3.2 Corollary. *For the standard interpretation in $Sh(X)$, the exponential 2^N is (isomorphic to) the sheaf $Cts(-,2^{\mathbf{N}})$ of continuous map into the Cantor space.*

In this corollary, 2 is the constant sheaf $\Delta(\{0,1\})$, and N is the constant sheaf $\Delta\mathbf{N}$ (cf. I. 2.10). It is easy to prove 3.2 directly (without using 3.1) from the universal property of the exponential 2^N and the adjunction between Δ and Γ of I. 2.10.

Notice that the remarks concerning morphisms in 1.1 and 1.2 generalize immediately from open embeddings and homeomorphisms of spaces to open embeddings and isomorphisms of locales.

There is one other well-known result which I shall use, viz.

3.3 Proposition. *The axiom of choice is valid in $Sh(X_{\neg\neg})$, for any locale X.*

Proof. (sketch) I shall prove validity of the formula

$$\forall f \in B^A[\forall y \in B \exists x \in A f(x) = y \to \exists s \in A^B \forall y \in B f s(y) = y] \quad (*)$$

for any two sheaves A and B on $X_{\neg\neg}$. Let $U \in \mathcal{O}(X_{\neg\neg})$, and let $f \in B^A(U)$ (that is, $f : A \mid U \to B \mid U$ is a morphism of sheaves on the open sublocale U of $X_{\neg\neg}$), such that $U \Vdash \forall y \in B \exists x \in A f(x) = y$. This means that $f : A \mid U \to B \mid U$ is an epimorphism (see III. 1.2). By Zorn's lemma, there is a maximal section $s : M \to A \mid U$ (a morphism in $Sh(U)$), where M is a subsheaf of $B \mid U$ and $f \circ s = id$. Suppose $M \neq B \mid U$. Then there is a $V \leq U$ in $\mathcal{O}(X_{\neg\neg})$ and a $b \in B(V)$ with $b \notin M(V)$. Let $W_b = \bigvee\{W \subseteq V \mid b \cdot (W \hookrightarrow V) \in M(W)\}$. Then $W_b \in \mathcal{O}(X_{\neg\neg})$, and since M is a subsheaf $b \in M(W_b)$. Therefore $W_b \neq V$, and hence $0 \neq (V \wedge \neg W_b) \in \mathcal{O}(X_{\neg\neg})$. Since f is epi, there is a non-zero $P \leq V \wedge \neg W_b$ in $\mathcal{O}(X_{\neg\neg})$ such that $b \cdot (P \hookrightarrow V) = f(a)$ for some $a \in A(P)$. Define M' to be the subsheaf of B generated by M and $b \cdot (P \hookrightarrow V)$, and let $s' = M' \to A$ be the extension of s defined by $s'(b \cdot (P \hookrightarrow V)) = a$. Then s' is a proper extension of s, contradicting the maximality of s. Thus we must have $M = B \mid U$, i.e. $s : B \mid U \to A \mid U$. A fortiori $U \Vdash \exists s \in A^B \forall y \in B f s(y) = y$. This proves the validity of $(*)$.

There is much more to say about locales. The interested reader might look at the first chapters of Joyal-Tierney (1984), or at Johnstone (1982).

§4. No definable ultrafilters.

In this section I shall prove theorem 2. I shall work in classical type theory with choice; that is, the system obtained from HHA by adding the axioms $\forall p \in \mathcal{P}(1)(\exists x(x \in p) \vee \neg\exists x(x \in p))$ for classical logic and $\forall f \in A^B(\forall a \in A \exists b \in B f(b) = a \rightarrow \exists s \in B^A \forall a f s(a) = a)$ for choice. The same argument will apply to ZFC, see II. 1.6.)

Let A be the countable product of copies of the ring $\mathbb{Z}/2\mathbb{Z}$,

$$A = \prod_{n \geq 0} \mathbb{Z}/2\mathbb{Z}.$$

A is a topological ring (with the product topology). As a set, A is the exponential $2^{\mathbb{N}}$. So if X is a locale then the interpretation of A in X is the sheaf $C(A)$ of locale maps into A (see 3.2). Consider also the subring

$$D = \oplus_{n \geq 0} \mathbb{Z}/2\mathbb{Z}$$

of A, given by those sequences of zeros and ones which are eventually zero. D is a discrete dense subring of A.

A non-principal ultrafilter on $\mathbb{N}$ is the same thing as a homomorphism

$$T : A \rightarrow \mathbb{Z}/2\mathbb{Z}$$

which vanishes on $D \subseteq A$. Let $\theta(x, y)$ be a formula for which it is provable that θ defines a homomorphism $A \rightarrow \mathbb{Z}/2\mathbb{Z}$ which vanishes on D. (As usual, I write $\theta(x) = y$ for $\theta(x, y)$.)

Now let X be any space, and let $j : X_{\neg\neg} \rightarrow X$ be the embedding as in §3. Let T be the interpretation of θ in $Sh(X_{\neg\neg})$. So T is a morphism of sheaves $C(A) \rightarrow \Delta(\mathbb{Z}/2\mathbb{Z})$. If $\tau : X \rightarrow X$ is a homeomorphism, then τ restricts to an isomorphism $\tau_1 : X_{\neg\neg} \rightarrow X_{\neg\neg}$ of locales. Since τ_1^* is logical and T is definable, we have

$$\tau_1^*(T) = T. \tag{1}$$

If $\alpha : X \rightarrow A$ is any continuous map, then $\alpha \circ j : X_{\neg\neg} \rightarrow A$ is an element of the interpretation $C(A)$ of the ring A in $Sh(X_{\neg\neg})$. Thus

$$\begin{aligned} \tau_1^*(T(\alpha \circ j)) &= \tau_1^*(T)(\tau_1^*(\alpha \circ j)) \\ &= T(\tau_1^*(\alpha \circ j)) \\ &= T(\alpha \circ j \circ \tau_1) \\ &= T(\alpha \circ \tau \circ j). \end{aligned} \tag{2}$$

I now wish to apply this to the special case where $X = A$ and α is the identity. Consider for each $d \in D$ the translation homeomorphism

$$\tau(d) : A \xrightarrow{\sim} A, \quad \tau(d)(a) = d + a,$$

together with its restriction

$$\tau(d)_1 : A_{\neg\neg} \to A_{\neg\neg}.$$

Lemma. *Let $F : A_{\neg\neg} \to Y$ be any continuous map into some topological space Y. If $F \circ \tau(d)_1 = F$ for all $d \in D$, then F is constant.*

Proof. Suppose $F \circ \tau(d)_1 = F$ for all $d \in D$. I wish to show that for some point $y \in Y$ it holds that for any open neighbourhood U of y, $F^{-1}(U) = A_{\neg\neg}$ (= the top element 1 of $\mathcal{O}(A_{\neg\neg})$). Clearly any non-empty open subset of A which is invariant under translations $\tau(d)$ in dense (since D is dense). It follows that if $0 \neq V \in \mathcal{O}(A_{\neg\neg})$ is invariant under $\tau(d)_1$ for all $d \in D$, then V must be the maximal element of $A_{\neg\neg}$. Since $F^{-1}(U)$ is invariant, it thus follows that it suffices to find a point $y \in Y$ such that $F^{-1}(U) \neq 0$ for any neighbourhood U of y. Suppose no such y exists. Then any $y \in Y$ has a neighbourhood U_y with $F^{-1}(U_y) = 0$. Hence $A_{\neg\neg} = F^{-1}(Y) = F^{-1}(\bigvee_{y \in Y} U_y) = \bigvee_{y\in Y} F^{-1}(U_y) = 0$, contradiction.

Now consider the inclusion $j : A_{\neg\neg} \to A$, which is an element of the interpretation of the ring A inside $Sh(A_{\neg\neg})$. Pick any $d \in D$, and write $\hat{d} : A_{\neg\neg} \to A$ for the constant map with value d. Thus $\hat{d}$ can be viewed as an element of the interpretation of $D \subseteq A$, and hence $T(\hat{d}) = 0$. Therefore

$$\begin{aligned}
\tau(d)_1^*(T(j)) &= T(\tau(d) \circ j) \quad \text{(see (2))}\\
&= T(\hat{d} + j)\\
&= T(\hat{d}) + T(j)\\
&= T(j).
\end{aligned}$$

By the lemma (applied to the case where Y is the discrete space $\mathbf{Z}/2\mathbf{Z}$), $T(j)$ is constant.

On the other hand, consider the homeomorphism

$$\rho : A \to A$$

sending an element a to its inverse $\rho(a)$, together with the restriction

$$\rho_1 : A_{\neg\neg} \to A_{\neg\neg}.$$

Also, write $\hat{1} : A_{\neg\neg} \to A$ for the constant map with value 1. Since $\hat{1}$ is the unit of the ring $C(A)$,

$$T(\hat{1}) = 1. \tag{3}$$

On the other hand, $j + \hat{1} = -j = \rho_1^*(j)$. Therefore

$$\begin{aligned}
T(j) + T(\hat{1}) &= T(j + \hat{1})\\
&= \rho_1^*(T)(\rho_1^*(j))\\
&= \rho_1^*(T(j))\\
&= T(j) \circ \rho_1\\
&= T(j),
\end{aligned}$$

the latter identity since $T(j)$ is constant. Thus $T(\hat{1}) = 0$, contradiction.

This proves that in $Sh(A_{\neg\neg})$, the formula " $\theta(x, y)$ defines a non-principal ultrafilter" is not valid.

References

In the text, I have referred to Grothendieck et al. (1972) as SGA 4, and to Mac Lane (1971) as CWM.

J.L. Bell - Toposes and Local Set Theories, Oxford U.P. 1988.

E.W. Beth - Semantic construction of intuitionistic logic, K.N.A.W. n.s. 19/11 (1956), p. 357-388.

A. Blass, A. Scedrov - Freyd's models for the independence of the axiom of choice, Memoirs AMS 404 (1989).

A. Boileau, A. Joyal - La logique de topos, Journal Symb. Logic 46 (1981), p. 6-16.

P.J. Cohen - Set Theory and the Continuum Hypothesis, Benjamin 1966.

M. Dummett - Elements of Intuitionism, Oxford U.P. 1977.

M.P. Fourman - Sheaf models for set theory, Journal Pure & Appl. Alg. 19 (1980), p. 91-101.

M.P. Fourman, D.S. Scott - Sheaves and logic, in: Springer Lecture Notes in Maths 753 (1979), p. 302-401.

P.J. Freyd - All topoi are localic, or: Why permutation models prevail, unpublished, 1979.

P.J. Freyd - The axiom of choice, J. Pure & Appl. Alg. 19 (1980), p. 103-125.

R.J. Grayson - Heyting-valued models for set theory, in: Springer Lecture Notes in Math. 753 (1979), p. 402-414.

R.J. Grayson - Concepts of general topology in constructive mathematics and in sheaves, Annals of Math. Logic 20 (1981), p. 1-41.

R.J. Grayson - Idem, part II, Annals of Math. Logic 23 (1982), p. 55-98.

A. Grothendieck et al. - Théorie des topos et cohomologie étale des schémas, Springer Lecture Notes in Maths. 269, 270 (1972) ("SGA4").

J.M.E. Hyland - The effective topos, in: The L.E.J. Brouwer Centenary Symposium, North Holland (1982), p. 165-216.

J.M.E. Hyland - A small complete category, Ann Pure & Appl. Logic 40 (1988), p. 135-165.

A. Joyal, M. Tierney - An Extension of the Galois Theory of Grothendieck, Memoirs of the A.M.S. 309 (1984).

A. Joyal, I. Moerdijk - Toposes as homotopy groupoids, Advances in Math. 80 (1990), p. 22-38.

P.T. Johnstone - Topos Theory, Academic Press, 1977.

P.T. Johnstone - Stone Spaces, Cambridge U.P. 1982.

P.T. Johnstone - Open locales and exponentiation, in: Mathematical Applications of Category Theory, Contemp. Mathematics vol. 30 (A.M.S. 1984), p. 84-116.

S.C. Kleene - Introduction to Metamathematics, Wolters-Noordhoff 1952.

A. Kock - Synthetic Differential Geometry, Cambridge U.P. 1981.

S.A. Kripke - Semantical analysis of intuitionistic logic, in: Crossley, Dummett (eds.) Formal Systems and Recursive Functions (North-Holland 1965), p. 92-130.

J. Lambek, P. Scott - Introduction to Higher Order Categorial Logic, Cambridge U.P. 1986.

S. Mac Lane - Categories for the Working Mathematician, Springer Verlag 1971.

S. Mac Lane, I. Moerdijk - A First Introduction to Topos Theory, Springer Verlag (to appear).

M. Makkai, G. Reyes - First Order Categorical Logic, Springer Lecture Notes in Maths. 611 (1977).

I. Moerdijk - The classifying topos of a continuous groupoid I, Transactions of the A.M.S. 310 (1988), p.629-668.

I. Moerdijk, G. Reyes - A smooth version of the Zariski topos, Advances in Maths. 65 (1987), p. 229-253.

D.S. Scott - Extending the topological interpretation to intuitionistic analysis I., Comp. Math. 20 (1986), p. 194-210.

D.S. Scott - Idem, II, in: Myhill et. al. (eds.), Intuitionism and Proof Theory, North-Holland 1970, p. 235-255.