Reflections on the Categorical Foundations of Mathematics

Joachim Lambek and Philip J. Scott

1 Introduction

Most practicing mathematicians see no need for the foundations of their subject. But those who wish to place it on a solid ground usually pick set theory, an axiomatic treatment of the membership relation expressed in first order logic. Some of us feel that higher order logic is more appropriate and, since Russell and Whitehead's *Principia Mathematica*, such a system has been known as type theory (more precisely, classical *impredicative* type theory with Peano's axioms). Although type theory has been greatly simplified by works of Alonzo Church, Leon Henkin, and others, and despite its naturalness for expressing mathematics, it was unjustly neglected until quite recently.

An apparently different approach to foundations is via category theory, a subject that was introduced by Samuel Eilenberg and Saunders Mac Lane in 1945. In 1964, F. W. Lawvere proposed to found mathematics on the category of categories (Lawvere, 1966). When he lectured on this at an international conference in Jerusalem, Alfred Tarski objected: "But what is a category if not a set of objects together with a set of morphisms?" Lawvere replied by pointing out that set theory axiomatized the binary relation of membership, while category theory axiomatized the ternary relation of composition.

Later Lawvere returned from the category of categories to the category of sets. Trying to axiomatize the latter (e.g. Lawvere, 1964), he ended up with the notion of an elementary topos, which made its first public appearance in joint work with Myles Tierney (Lawvere, 1970; Tierney, 1972). Elementary toposes have the advantage of describing not only sets, but also sheaves (called "variable sets" by Lawvere). Quoting Lawvere (1972):

> This is the development on the basis of elementary (first-order) axioms of a theory of "toposes" just good enough to be applicable not only to sheaf theory, algebraic spaces, global spectrum etc. as originally envisaged by Grothendieck, Giraud, Verdier, and Hakim but also to Kripke semantics, abstract proof theory, and the Cohen-Scott-Solovay method for obtaining independence results in set theory.

G. Sommaruga (ed.), *Foundational Theories of Classical and Constructive Mathematics*, The Western Ontario Series in Philosophy of Science 76,
DOI 10.1007/978-94-007-0431-2_9,

Indeed, it was soon realized that an elementary topos has an associated "internal" logic which is essentially a version of (intuitionistic) type theory. In the second part of our book *Introduction to higher order categorical logic* (Lambek and Scott, 1986), we tried to exploit the close connections between higher order logic (better called "higher order arithmetic") and topos theory.

2 Type Theory

By a *type theory* (or *higher order arithmetic*) we understand a formulation of higher order logic with Peano's axioms. We shall follow our book (1986) and consider type theories based on equality. Thus the language contains both *types* and *terms* (of the indicated types) as follows:

Types:	1	Ω	N	Ω^A	$A \times B$
Terms:	$*$	$a = a'$	0	$\{x : A \mid \varphi(x)\}$	$\langle a, b\rangle$
		$a \in \alpha$	Sn		

It is assumed that $1, N, \Omega$ are types, and that the types are closed under the operations Ω^A and $A \times B$ for given types A and B. Here 1 denotes a one-point type, Ω the type of propositions (or truth values), N the type of natural numbers, Ω^A the "powerset" of A, and $A \times B$ the cartesian product of types A and B. Among the terms (not indicated above) there are infinitely many variables of each type. We assume that $*$ is a term of type 1, 0 is a term of type N, and that the terms are closed under the operations $=, \in, S, \{- \mid -\}$, and $\langle -, -\rangle$, as indicated above, where it is understood that a, a' are of the same type A, b is of type B, α is of type Ω^A, n is of type N and $\varphi(x)$ is of type Ω. We adopt the current convention of writing $t : A$ for "t is a term of type A".

We also assume that there is a collection of theorems which include the usual axioms and which is closed under the usual rules of inference (i.e. equality, pairing, comprehension, extensionality, and Peano's axioms). The familiar logical symbols are now definable as follows (see also Lawvere, 1972)

$$
\begin{aligned}
&\top && := * = * \\
&p \wedge q && := \langle p, q\rangle = \langle \top, \top\rangle \qquad \text{where } p, q : \Omega \\
&p \Rightarrow q && := p \wedge q = p \\
&\forall_{x:A}\phi(x) && := \{x : A \mid \phi(x)\} = \{x : A \mid \top\} \qquad \text{where } \phi(x) : \Omega \\
&\bot && := \forall_{x:\Omega} x \\
&\neg p && := \forall_{x:\Omega}(p \Rightarrow x) \\
&p \vee q && := \forall_{x:\Omega}(((p \Rightarrow x) \wedge (q \Rightarrow x)) \Rightarrow x) \\
&\exists_{x:A}\varphi(x) && := \forall_{y:\Omega}(\forall_{x:A}((\phi(x) \Rightarrow y) \Rightarrow y))
\end{aligned}
$$

The usual properties of these logical connectives can now be proved (see Lambek and Scott, 1986).

We will call a type theory *analytic* if it contains no types and terms other than the ones it must contain according to the above definition. Thus, an analytic type theory does not contain the type of humans or the type of vegetables, nor does it contain terms denoting the binary relations of loving or eating. Even the internal language of a topos (see below) is not analytic, since it admits as types all sets (a *set* in a topos being a morphism $1 \to \Omega^A$, for some object A).

Pure type theory $\mathcal{L}_0$ is the analytic type theory containing no theorems other than those following from the above inductive definition. Every analytic type theory has the form $\mathcal{L}_0/\theta$, where θ is a set of propositions (i.e. terms of type Ω) now considered as additional *nonlogical axioms*. We may even take θ to be the set of all theorems.

3 Elementary Toposes

A *topos*, according to Lawvere, is a cartesian closed category (ccc) with pullbacks, a subobject classifier Ω and a natural numbers object N. By a ccc we mean a category with a terminal object 1, cartesian products $A \times B$ and exponentiation C^B, together with a canonical bijection between arrows $(A \times B) \to C$ and arrows $A \to C^B$. As Lawvere himself pointed out (Lawvere, 1969), the prime example of a ccc is the proof theory of the positive intuitionistic propositional calculus, with

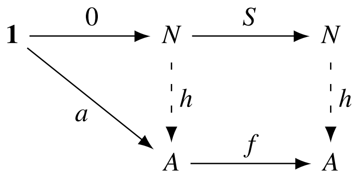
$$1 = \top\,, \quad A \times B = A \wedge B\,, \quad C^B = B \Rightarrow C$$

According to the so-called Curry-Howard isomorphism, the associated proof theory can also be described by the typed lambda calculus (with surjective pairing); hence it is quite natural that ccc's, typed lambda calculi, and the proof theory of positive intuitionistic propositional calculi turn out to be equivalent (see Lambek and Scott, 1986).

A *subobject classifier* in a ccc with pullbacks is an object Ω together with a canonical (monic) arrow $\top : 1 \to \Omega$ and a canonical bijection between subobjects B of A and their characteristic morphisms $\chi_B : A \to \Omega$.[1] This generalizes the familiar set-theoretic bijection between subsets of a set A and characteristic functions $A \to \Omega$, where Ω is a two-element set. Viewed as usual classical sets, Ω and the powerset Ω^A are Boolean algebras, whereas in toposes, Ω and Ω^A carry the more general structure of a Heyting algebra. It is therefore not surprising that the "internal logic" of a topos is in general intuitionistic.

According to Lawvere, a *Natural Numbers Object* (NNO) in a ccc is an object N, together with arrows $0 : 1 \to N$ and $S : N \to N$ such that, given arrows $a : 1 \to A$ and $f : N \to N$, there is a unique arrow $h : N \to A$ making the following diagram commute:

[1] This bijection is induced by pulling back the arrow $\top : 1 \to \Omega$ along $\chi_B : A \to \Omega$.



In the case of a topos, this yields Lawvere's categorical formulation of the well-known Peano axioms for set theory (Lawvere, 1964), which is seen here by putting $h(n) = f^n(a)$. In the case of cartesian closed categories (and their equivalent typed lambda calculi), this leads to notions of higher-type "iteration" arising in proof theory, recursive function theory, and theoretical computer science.

4 Comparing Type Theories and Toposes

In Lambek and Scott (1986) we compared two categories: the category of type theories, by which we mean intuitionistic type theories with axiom of infinity or, equivalently, Peano's axioms, and the category of toposes, which we understand to be elementary toposes with natural numbers object. As morphisms in the former we took "translations" between type theories, and in the latter, so-called "logical morphisms" between toposes (we ignored the alternative "geometric morphisms" arising from the Grothendieck tradition; for that, see Mac Lane and Moerdijk, 1992; Makkai and Reyes, 1977). We introduced functors between the two categories as follows. One functor L assigns to any topos $\mathcal{T}$ its "internal language" $L(\mathcal{T})$ (an intuitionistic type theory); the other functor T assigns to any type theory $\mathcal{L}$, the topos $T(\mathcal{L})$ "generated" by it, a kind of Lindenbaum-Tarski category constructed from the language. Let us briefly recall these two constructions.

The types of $L(\mathcal{T})$ are the objects of $\mathcal{T}$ and the closed terms of type A in $L(\mathcal{T})$ are the arrows $a : 1 \to A$ in $\mathcal{T}$. In particular, propositions of $L(\mathcal{T})$ are the arrows $p : 1 \to \Omega$ in $\mathcal{T}$. We say that *p holds in* $\mathcal{T}$, *p is true* in $\mathcal{T}$, or *p is a theorem* of the type theory $L(\mathcal{T})$, if and only if $p = \top$; that is, if p equals the distinguished arrow $\top : 1 \to \Omega$. Thus $L(\mathcal{T})$ has a "semantic" definition of theorem; it differs from logicians' more familiar (freely generated) type theories, in which terms are defined inductively from a small set of primitives, and in which "theorems" are introduced with the help of a recursive proof predicate. The internal language of a topos has some interesting properties. For example, $L(\mathcal{T})$ satisfies the *unique existence property*: if $\exists!_{x:A}\phi(x)$ holds in $\mathcal{T}$, then there is a closed term of type A, namely an arrow $a : 1 \to A$ in $\mathcal{T}$, such that $\phi(a)$ holds in $\mathcal{T}$. As Bertand Russell would have said: "a is the unique $x : A$ such that $\phi(x)$". We sometimes denote such a unique a by $\imath_{x:A}.\phi(x)$.

The topos $T(\mathcal{L})$ generated by the type theory $\mathcal{L}$ has as objects closed terms α of type Ω^A (modulo provable equality), and as morphisms $\alpha \to \beta$, where $\alpha : \Omega^A$ and $\beta : \Omega^B$, we choose those binary "relations" (closed terms) $\phi : \Omega^{A\times B}$ (again, modulo provable equality) such that the following is provable in $\mathcal{L}$:

$$\forall_{x:A}(x \in \alpha \Rightarrow \exists!_{y:B}(y \in \beta \wedge (x, y) \in \phi))$$

Intuitively, $T(\mathcal{L})$ is the category of "sets" and "functions" formally definable within the higher-order logic $\mathcal{L}$: its objects are the "sets" $\alpha, \beta, \ldots$ in $\mathcal{L}$ and its morphisms are the "provably functional relations" ϕ in $\mathcal{L}$ between such objects, all modulo provable equality.

We proved quite formally that there are two natural transformations $\varepsilon : LT \to id$ and $\eta : id \to TL$ rendering the functor T to be left adjoint to L. Moreover, we showed that η was an isomorphism, so that every topos is equivalent to the topos generated by its internal language. We pointed out in an exercise that a slight tightening of the definition of translation would also make ε an isomorphism; this was carried out by Lavendhomme and Lucas (1989). However, returning to our more natural notion of translation, we showed that, for any type theory $\mathcal{L}$, the translation $\mathcal{L} \to LT(\mathcal{L})$ is a conservative extension.

A type theory $\mathcal{L}$ may be interpreted in a topos $\mathcal{T}$ by means of a translation of languages $\mathcal{L} \to L(\mathcal{T})$ or, equivalently, by a logical morphism $T(\mathcal{L}) \to \mathcal{T}$, recalling that T is left adjoint to L. In some sense, every such interpretation may be viewed as a "model" of $\mathcal{L}$ in $\mathcal{T}$. By abuse of language, one often refers to $\mathcal{T}$ itself as the model. In particular, this view is justified for models of *pure type theory* $\mathcal{L}_0$, the initial object in the category of type theories and translations. For in this case, there is a unique translation from $\mathcal{L}_0$ to any type theory. In particular, for any topos $\mathcal{T}$, there is a unique translation $\mathcal{L}_o \to L(\mathcal{T})$, thus a unique logical morphism $T(\mathcal{L}_0) \to \mathcal{T}$. $\mathcal{F} = T(\mathcal{L}_0)$ is thus initial in the category of toposes and logical morphisms and is known as the *free topos*. Hence any elementary topos (with Natural Numbers Object) serves as a model of $\mathcal{L}_0$.

5 Models and Completeness

Following Leon Henkin's presentation of classical type theory (Henkin, 1950), we adopt a more restrictive notion of model. A *model* of a type theory $\mathcal{L}$ is a topos $\mathcal{T}$ satisfying three properties (for formulas in $L(\mathcal{T})$):

(a) **consistency**: $\perp$ is not true ;
(b) **disjunction property**: if $p \vee q$ is true in $\mathcal{T}$, then so is p or q;
(c) **existence property**: if $\exists_{x:A}\phi(x)$ is true in $\mathcal{T}$, then so is $\phi(a)$ for some closed term a of type A in $L(\mathcal{T})$, that is, for some morphism $a : 1 \to A$ in $\mathcal{T}$.

Following Alexander Grothendieck, we now call such a topos a *local topos*.

Peter Freyd observed that the above three linguistic properties can be expressed categorically as follows:

(a) the terminal object 1 is not initial;
(b) 1 is indecomposable;
(c) 1 is projective.

Local toposes of interest also have another property:

(d) all numerals are *standard*; that is, all the arrows $1 \to N$ have the form $S^n 0$ for some natural number n.

As we mentioned earlier, Russell and Whitehead (as well as Gödel and Henkin) dealt with *classical* type theory. This theory differs from intuitionistic type theory by the addition of a single axiom β (the *law of excluded middle*), which we may write as:

$$\forall_{x:\Omega}(\neg\neg x \Rightarrow x)$$

or, equivalently,

$$\forall_{x:\Omega}(x \vee \neg x)$$

A topos is said to be *Boolean* if its internal language is classical. In particular, this implies that $\Omega \cong 1 + 1$ (the coproduct). Boolean local toposes may be characterized as follows (see Seldin and Hindley, 1980):

Proposition 5.1 *A topos $\mathcal{T}$ is Boolean local iff it satisfies*

I *Consistency:* $\mathsf{T} \neq \mathsf{F} : 1 \to \Omega$.

II *Universal Property: If $\phi(x)$ is a formula in $L(\mathcal{T})$ such that $\phi(a)$ holds in $\mathcal{T}$ for all closed terms $a : A$ in $L(\mathcal{T})$, then $\forall_{x:A}\phi(x)$ holds in $\mathcal{T}$.*

The second property can be expressed in categorical language by saying that the terminal object 1 of $\mathcal{T}$ is a *generator : if $f, g : A \to B$ and $fa = ga$ for all $a : 1 \to A$, then $f = g$.*

Proof Assuming that $\mathcal{T}$ is Boolean and local, the universal property follows from the existence property, using negation.

Conversely, assume that $\mathcal{T}$ satisfies properties (i) and (ii) above. Among the subobjects of 1 are the (isomorphism classes of) monomorphisms $0 \hookrightarrow 1$ and $1 \hookrightarrow 1$, with characteristic morphisms $\mathsf{F} : 1 \to \Omega$ and $\mathsf{T} : 1 \to \Omega$, respectively. Here 0 is the initial object of $\mathcal{T}$. By (i), these are distinct subobjects of 1. We claim there are no others. For let $m : A \hookrightarrow 1$ be any subobject of 1. If there is an arrow $a : 1 \to A$, clearly $ma = 1_1$, hence $mam = m1_A$, so $am = 1_A$ and m is an isomorphism $A \cong 1$. If there is no arrow $1 \to A$, we claim $A \cong 0$. For trivially $\phi(a)$ then holds in $\mathcal{T}$ for all closed terms a of type A; hence $\forall_{x:A}\phi(x)$ holds in $\mathcal{T}$, whatever formula $\phi(x)$ we take. In particular, for any object B, let $\phi(x)$ be the formula $\exists!_{y:B}\psi(x, y)$, where $\psi(x, y)$ is, for example, $x \neq x \wedge y = y$. Then clearly ψ defines an arrow $A \to B$. Since 1 is a generator, there is at most one such arrow $A \to B$, and thus A is an initial object.

Therefore 1 has exactly two subobjects, and so there are exactly two arrows $\mathsf{T}, \mathsf{F} : 1 \to \Omega$. Thus the topos $\mathcal{T}$ is *two-valued*. Hence for all arrows $p : 1 \to \Omega$, $\neg\neg p = p$, hence $\neg\neg = 1_\Omega$, since 1 is a generator. Hence the formula $\forall_{x:A}\neg\neg x = x$ holds in $\mathcal{T}$, so $\mathcal{T}$ is Boolean.

Once $\mathcal{T}$ is Boolean, the universal property gives rise to the existence property (by negation). Similarly the conjunction property (which holds in any topos) gives rise to the disjunction property by de Morgan's law. Thus $\mathcal{T}$ is local.

Remark 5.2 In general in toposes, Boolean does not imply 2-valued; however it does in the presence of the disjunction property. Conversely, 2-valued does not imply Boolean, but it does if 1 is a generator.

Gödel's completeness theorem was originally enunciated for classical first order logic, but was extended by Henkin to higher order as follows (in our terminology):

> A proposition holds in $T(\mathcal{L})$, the topos generated by a classical type theory $\mathcal{L}$, if and only if it is true in all models of $\mathcal{L}$, i.e. in all Boolean local toposes.

Of course, if $\mathcal{L}$ is inductively generated, such propositions are usually called *provable*, and the Completeness Theorem asserts the equivalence between provability in $\mathcal{L}$ and truth in all models.

What about Gödel's more famous Incompleteness Theorem, which he himself had originally stated for classical type theory? An examination of its proof in our setting (carried out in the next section) shows it actually asserts the following:

> In a consistent analytic type theory $\mathcal{L}$ whose theorems are recursively enumerable, in order to characterize provability in $\mathcal{L}$, it is *not* sufficient to look only at local toposes which also satisfy the *ω-property:* if $\phi(S^n 0)$ is true for all natural numbers n, then $\forall_{x:N}\phi(x)$ is also true in the model.

The crucial role of the ω-property was first pointed out by David Hilbert. Classically, though not intuitionistically, it is equivalent to what we call the ω^* *property*: if $\exists_{x:N}\phi(x)$ is true in the model, then so is $\phi(S^n 0)$ for some natural number n. For a local topos, the ω^* property follows from the existence property, provided we assume that all numerals are standard. In fact, for a local topos, the ω^*-property is equivalent to the condition that all numerals are standard.

6 Gödel's Incompleteness Theorem

In any analytic type theory $\mathcal{L}_0/\theta$, we may effectively enumerate all terms of a given type. This may be done with the help of the well-known method of Gödel numbering, or even just by putting the terms into alphabetical order. In particular, let p_n be the nth proposition (closed term of type Ω) and α_n be the nth numerical predicate (closed term of type Ω^N).

The analytic type theories we are usually interested in also possess a recursive proof predicate, ensuring that the set of theorems is recursively enumerable.[2] If θ contains all the axioms and is closed under the rules of inference, θ is the set of theorems of $\mathcal{L}_0/\theta$ and is recursively enumerable by a primitive recursive function g. Thus $p_{g(n)}$ denotes the nth theorem of $\mathcal{L}_0/\theta$. Note that the set of numerical predicates in the internal language of the "usual" category of sets cannot be enumerated, recursively or otherwise, as follows from Cantor's theorem. This serves as an inspiration for the theorems of Gödel and Tarski, as we shall see.

Here is our formulation of Gödel's incompleteness theorem, which includes both the classical and intuitionist cases.

[2] The set of theorems (of an analytic type theory) is the set of propositions formally provable from the logical and nonlogical axioms, using the rules of inference of $\mathcal{L}_0$.

Theorem 6.1 *In a consistent analytic type theory $\mathcal{L}$ whose theorems are recursively enumerable, there is a proposition q which does not hold in any model in which all numerals are standard, yet its negation $\neg q$ is not provable. Thus, $\neg q$ must hold in every Boolean model in which all numerals are standard. Hence, if $\mathcal{L}$ has at least one model in which the numerals are standard, neither q nor $\neg q$ is a theorem.*

Proof For a type theory $\mathcal{L}$, we write $\vdash_{\mathcal{L}}$ to denote provability in $\mathcal{L}$. Recall that any primitive recursive function f can be *numeralwise represented* by a formula $\phi(x, y)$ in $\mathcal{L}_0$ such that

$$\text{for all } m \in \mathbb{N}, \qquad \vdash_{\mathcal{L}_0} S^{f(m)}0 = \imath_{y:N}.\phi(S^m 0, y)$$

where $\imath_{y:N}.\phi(x, y)$ denotes "the unique $y : N$ such that $\phi(x, y)$" (Russell's definite description operator, which we can introduce as an abbreviation in $\mathcal{L}_0$). Recall that provability in $\mathcal{L}_0$ implies provability in any type theory. The representability of the primitive recursive functions in $\mathcal{L}_0$ is shown in our book (Lambek and Scott, 1986, Remark 3.6, p. 266).

Consider the two primitive recursive functions f and g, represented by ϕ and ψ, respectively, where f enumerates the propositions $S^m 0 \notin \alpha_m$ (already considered by Cantor) and g enumerates the theorems of $\mathcal{L}$. Thus, for any $m \in \mathbb{N}$,

$$\vdash_{\mathcal{L}} S^m 0 \notin \alpha_m \quad \text{iff} \quad \text{for some } n \in \mathbb{N}, \;\; f(m) = g(n) \quad . \tag{1}$$

Putting $\chi = \phi \wedge \psi$, we may write the RHS of (1) as:

$$\text{for some } n \in \mathbb{N}, \quad \vdash_{\mathcal{L}} \chi(S^m 0, S^n 0), \tag{2}$$

which implies

$$\vdash_{\mathcal{L}} \; \exists_{y:N}\chi(S^m 0, y), \tag{3}$$

that is,

$$\vdash_{\mathcal{L}} S^m 0 \in \alpha_k, \;\; \text{where} \;\; \alpha_k := \{x : N \mid \exists_{y:N}\chi(x, y)\}. \tag{4}$$

Therefore

$$\vdash_{\mathcal{L}} S^m 0 \notin \alpha_m \quad \text{implies} \quad \vdash_{\mathcal{L}} S^m 0 \in \alpha_k$$

Putting $m = k$, we infer by consistency that $S^k 0 \notin \alpha_k$ is not a theorem of $\mathcal{L}$.

Let us try to reverse the above reasoning. Clearly (4)⇒(3). The implication (2)⇒(3) may be reversed if we pass to the internal language $\mathcal{L}'$ of any model of $\mathcal{L}$ in which all numerals are standard, which thus inherits the existence and disjunction properties. We thus obtain

$$\vdash_{\mathcal{L}'} S^m 0 \in \alpha_k \text{ implies } \vdash_{\mathcal{L}'} \exists_{y:N}\chi(S^m 0, y)$$

$$\text{implies for some } n \in \mathbb{N} \quad \vdash_{\mathcal{L}'} \chi(S^m 0, S^n 0)$$
$$\text{by the Existence Property in } \mathcal{L}'$$

$$\text{implies for some } n \in \mathbb{N} \ \ f(m) = g(n),$$

$$\text{hence } \vdash_{\mathcal{L}} S^m 0 \notin \alpha_m \qquad \text{by (1).}$$

Again, putting $m = k$ and recalling that $S^k 0 \notin \alpha_k$ is not a theorem of $\mathcal{L}$, we infer that not $\vdash_{\mathcal{L}'} S^k 0 \in \alpha_k$, hence $S^k 0 \in \alpha_k$ does not hold in any model where the numerals are standard.

The theorem follows if we take q to be $(S^k 0 \in \alpha_k)$.

Corollary 6.2 *Assuming that the "usual" category of sets $\mathcal{S}$ is a Boolean local topos in which all numerals are standard, the set of propositions of $\mathcal{L}_0$ which hold in $\mathcal{S}$ is not recursively enumerable. Hence $\mathcal{S}$ cannot be construed as the topos generated by an analytic type theory whose theorems are recursively enumerable.*

Remark 6.3 The assumption that all numerals are standard is redundant, if we define "standard numerals" to be the arrows $1 \to N$ in the "usual" category $\mathcal{S}$ of sets. Thus Gödel's proposition $\neg q$ is true in $\mathcal{S}$ but not provable.

7 Reconciling Foundations

7.1 Constructive Nominalism

Gödel's incompleteness theorem seemed to show that Formalism and Platonism are mutually incompatible philosophies of mathematics. Indeed, this is what Gödel himself had in mind. He believed that the ω-property must hold in the Platonic universe of mathematics, later to be called "the model in the sky" by William Tait (1986). The contradiction disappears if one abandons classical mathematics for a moderate form of Intuitionism. According to the Brouwer-Heyting-Kolmogorov interpretation of formal intuitionistic arithmetic, the validity of a universal statement $\forall_{x:N}\phi(x)$ does not follow from the collection of its numerical instances $\phi(S^n 0)$, for each $n \in \mathbb{N}$, unless the validity of all these instances has been established in a *uniform* way. For all we know, a proof of $\phi(S^n 0)$ may increase in length and complexity with n. No such objection applies to the ω^*- property.

Although the formulation of Gödel's incompleteness theorem remains valid for intuitionistic higher order logic, this is no longer the case if the ω-property is replaced by the ω^*-property. In fact, a statement in pure intuitionistic higher order logic is provable if and only if it holds in $\mathcal{F} = T(\mathcal{L}_0)$, the *free topos*. Recall, this is the initial object in the category of all toposes and logical morphisms, and is constructed linguistically as the topos generated from pure intuitionistic type theory

$\mathcal{L}_0$. As has been pointed out repeatedly (Lambek, 2004; Lambek and Scott, 1986), the free topos should satisfy moderate adherents of various traditional philosophical schools in the foundations of mathematics:

- Platonists, because as an initial object it is unique up to isomorphism;
- Formalists, or even nominalists, because of its linguistic construction;
- Constructivists, or moderate intuitionists, because the underlying type theory is intuitionistic;
- Logicists, because this type theory is a form of higher order logic, although it must be complemented by an axiom of infinity, say in the form of Peano's axioms.

This eclectic point of view has been called "constructive nominalism" in Couture and Lambek (1991).

Proofs that the free topos is local have been obtained by Boileau and Joyal (Boileau, 1975; Boileau and Joyal, 1981), and by us (Lambek and Scott, 1980, 1986). Our ultimate proof was based on what is called the *Freyd Cover*, obtained by "glueing" $\mathcal{F}$ into the "usual" category of sets. Freyd showed that every locally small topos $\mathcal{T}$ gives rise to a local topos $\widehat{\mathcal{T}}$ in which all numerals are standard, together with a logical functor $G : \widehat{\mathcal{T}} \to \mathcal{T}$. The condition that $\mathcal{T}$ is *locally small* ensures that each set of arrows $Hom_{\mathcal{T}}(A, B)$ lives (as an object) in this category $\mathcal{S}$ of sets; the latter is presumed to be local and such that all numerals are standard. But what is this "usual" category of sets? We shall return to this question; for now, the reader may have to use her intuition to identify $\mathcal{S}$.

Returning to Freyd's argument (Freyd, 1978), let $\mathcal{T} = \mathcal{F}$, the free topos. Then, by initiality, there is a unique logical functor $F : \mathcal{F} \to \widehat{\mathcal{F}}$. Thus we obtain a logical functor $GF : \mathcal{F} \to \mathcal{F}$, which must equal $id_{\mathcal{F}}$, again by initiality. It follows that $\mathcal{F}$ inherits (from $\widehat{\mathcal{F}}$, hence from $\mathcal{S}$) the properties of being local and that all numerals are standard.

7.2 *What Is the Category of Sets?*

We saw above that we were able to construct a local topos in which all numerals are standard, which should satisfy moderate intuitionists. Unfortunately, Freyd's proof assumed the existence of the "usual" category of sets $\mathcal{S}$, which is itself assumed to be a local topos in which all numerals are standard. The category $\mathcal{S}$ may be said to live, if not in the world of mathematics, then in the world of metamathematics. If the metamathematician is herself an intuitionist, she might believe that this category of sets could be the free topos itself. But then we reach a circularity: to prove the free topos is a model in which numerals are standard, we must assume an ambient category of sets $\mathcal{S}$ which itself has that property, and of course we cannot then postulate that to be the free topos. One is reminded of Lewis Carroll (1895).

What if the metamathematician believes in classical logic? In that case, she must assume the existence of a model topos $\mathcal{S}$ in which the terminal object is a generator, and in which all numerals are standard, the "usual" category of sets. While the existence of such model toposes can be shown with the help of the axiom of choice (Lambek and Scott, 1986), can even a single one be "constructed"? For example, consider *classical* type theory $\mathcal{L}_1 = \mathcal{L}_0/\beta$, where the formula $\beta = \forall_{x:\Omega}(\neg\neg x \Rightarrow x)$ is added to $\mathcal{L}_0$ as a new axiom. The Boolean topos $T(\mathcal{L}_1)$ generated by $\mathcal{L}_1$ is the initial object in the category of all Boolean toposes. Is $T(\mathcal{L}_1)$ a model? Unfortunately it is not local, by the Incompleteness Theorem for $\mathcal{L}_1$. Indeed, the disjunction property fails for any undecidable sentence q, since we can prove in $\mathcal{L}_1$ that $\vdash q \vee \neg q$. Indeed, we conjecture that no such classical model topos can be constructed, at least if we require it to satisfy reasonable properties.

8 What Is Truth?

What is truth? This question, once raised by Pilate, received different answers from different mathematicians.

Hilbert famously proposed the problem of showing that mathematical statements are true if and only if they can be proved. Like all of us, he assumed the set of proofs to be recursive.

Brouwer once asserted that mathematical statements are true if and only if they are known. In retrospect, he should have said "can be known", if truth is to be independent of time.

Gödel believed that a (classical) mathematical statement is true if and only if it holds in some kind of Platonic universe, which we take to be a Boolean local topos in which numerals are standard.

It follows from Gödel's incompleteness theorem that Hilbert's position is incompatible with the assumption that the Platonic universe is classical. However, if we assume that this universe is intuitionistic (the free topos), there is no contradiction. Moreover, Brouwer's modified position is vindicated if we interpret "knowable" as "provable".

Tarski defines truth differently. He said "p is true" instead of asserting p. By abuse of language, we often abbreviate *p is true* by p, ignoring quotation marks, like most mathematicians. Tarski then said that a numerical predicate τ *defines* truth (for a language $\mathcal{L}$) provided

$$\text{for all } n \in \mathbb{N}, \quad \vdash_{\mathcal{L}} (p_n \Leftrightarrow S^n 0 \in \tau).$$

where we use the same conventions of Gödel numbering as in Gödel's theorem. Here is our formulation of Tarski's undefinability theorem.

Theorem 8.1 *In any consistent analytic type theory, truth (in Tarski's sense) is not definable by a numerical predicate.*

Proof As in the proof of Gödel's theorem, suppose there were such a τ, and let

$$p_{f(n)} := S^n 0 \notin \alpha_n.$$

Then

$$\vdash_{\mathcal{L}} (S^{f(n)}0 = \imath_{y:N}.\phi(S^n 0, y))$$

where ϕ represents f. Put

$$\alpha_k := \{x : N \mid \exists_{y:N}(y \in \tau \wedge \phi(x, y))\}$$

Then we have the following provable equivalences:

$$\begin{aligned} \vdash_{\mathcal{L}} (S^k 0 \in \alpha_k &\Leftrightarrow S^{f(k)}0 \in \tau \\ &\Leftrightarrow p_{f(k)} \\ &\Leftrightarrow S^k 0 \notin \alpha_k) \end{aligned}$$

which contradicts consistency.

We will attempt to briefly compare the different notions of truth. Let $\vdash_{\mathcal{L}_0}$ stand for provability in $\mathcal{L}_0$, hence for truth in the free topos $\mathcal{F}$. We would like to interpret this as truth in Brouwer's sense. Comparing this with Tarski's notion of meta-truth, we would believe that, for all propositions p, $(\vdash_{\mathcal{L}_0} p) \Leftrightarrow p$. In particular, soundness corresponds to the entailment ($(\vdash_{\mathcal{L}_0} p) \Rightarrow p$).

Most post-Gödel mathematicians still believe in soundness. However soundness already implies consistency; for by soundness

$$(\vdash_{\mathcal{L}_0} p \wedge \ \vdash_{\mathcal{L}_0} \neg p) \Rightarrow (p \wedge \neg p) \ .$$

Yet, Gödel's second incompleteness theorem (not treated here) shows that consistency of $\mathcal{L}_0$ cannot be proved in $\mathcal{L}_0$; hence (using the encoding methods of the second Gödel theorem) soundness in the above sense cannot be formally proved either.

We may also ask whether Gödel's notion of truth (in a classical Platonic universe $\mathcal{S}$) implies Tarski's notion of meta-truth, i.e. whether ($(\mathcal{S} \models p) \Rightarrow p$). Now this implies ($(\vdash_{\mathcal{L}_1} p) \Rightarrow p$), where $\mathcal{L}_1 = \mathcal{L}_0/\beta$ is pure classical type theory, which is initial among all classical type theories (including the internal language of $\mathcal{S}$). Thus Tarski's notion of meta-truth implies soundness of $\mathcal{L}_1$, which again cannot be proved in $\mathcal{L}_1$ (when suitably encoded), by the same argument as above.

9 Continuously Variable Sets

It would appear that metamathematics is an attempt by mathematicians to lift themselves up by their own bootstraps. This had already been noted by Lewis Carroll (1895), in connection with the rule of modus ponens. It is also evident to anyone who looks at Gentzen style deductive systems, which derive the meaning of logical connectives from that of the meta-logical ones.

If we cannot single out a distinguished Boolean local topos as a candidate for the classical category of sets, we may be forced to look at the totality of all such models.

From an algebraic point of view, Gödel's completeness theorem asserts

Every topos is a subtopos of a direct product of local toposes.

This is analogous to the familiar assertion:

Every commutative ring is a subring of a direct product of local rings.

However, the latter statement can be improved to one that plays a crucial rôle in modern algebraic geometry (see Grothendieck and Dieudonné, 1960)

Every commutative ring is the ring of continuous global sections of a sheaf of local rings.

It has been realized for some time that Gödel's completeness theorem can be improved analogously:

Every topos is equivalent to the topos of global sections of a sheaf of local toposes .

It had also been clear that the models of any type theory, including those of the internal language of a topos, are the points of a topological space and that the truth of a proposition varies continuously from point to point. With any proposition in $\mathcal{L}_0$ one associates a *basic open set* consisting of all models in which the proposition is true. After various starts towards the sheaf representation of toposes, the result was ultimately established by Awodey (2000).[3]

10 Some Intuitionistic Principles

The fact that the free topos is local (and has only standard numerals) may be exploited to prove a number of intuitionistic principles for pure intuitionistic type theory $\mathcal{L}_0$, as we showed in our book (Lambek and Scott, 1986):

Consistency: not ($\vdash \bot$) .
Disjunction Property: If $\vdash_{\mathcal{L}_0} p \vee q$, then $\vdash_{\mathcal{L}_0} p$ or $\vdash_{\mathcal{L}_0} q$.
Existence Property: If $\vdash_{\mathcal{L}_0} \exists_{x:A}\phi(x)$ then $\vdash_{\mathcal{L}_0} \phi(a)$ for some closed term a of type A.

[3] Having observed that the truth of a proposition varies continuously from point to point, one of the present authors was led to announce the sheaf representation at conferences in Sussex and Amsterdam, but he made a bad choice of the basic open sets and used a definition of "local" which employed only the disjunction property. The first fault was rectified in a joint paper with Moerdijk (Lambek, 1994), further expository development occurred in our book (Lambek and Scott, 1986), and the second fault was rectified in (Lambek and Moerdijk, 1982), in which the author introduced a large number of "Henkin constants" to witness existential statements. This was shown to be unnecessary in a more recent article of Awodey (2000), who replaced the earlier logical proofs, based on definition by cases, by a purely categorical one. Similar ideas had also been pursued by Peter Freyd.

Troelstra's Uniformity principle for $A = \Omega^C$: If $\vdash_{\mathcal{L}_0} \forall_{x:A} \exists_{y:N} \phi(x, y)$ then $\vdash_{\mathcal{L}_0} \exists_{y:N} \forall_{x:A} \phi(x, y)$.
In the free topos $\mathcal{F}$, the uniformity principle says the arrows $\Omega^C \to N$ are constant (i.e. factor through some standard numeral).

Independence of premisses: If $\vdash_{\mathcal{L}_0} \neg p \Rightarrow \exists_{x:A} \phi(x)$ then $\vdash_{\mathcal{L}_0} \exists_{x:A}(\neg p \Rightarrow \phi(x))$.

Markov's Rule: If $\vdash_{\mathcal{L}_0} \forall_{x:A}(\phi(x) \vee \neg\phi(x))$ and $\vdash_{\mathcal{L}_0} \neg\forall_{x:A} \neg\phi(x)$, then $\vdash_{\mathcal{L}_0} \exists_{x:A} \phi(x)$.
This says: if in pure intuitionist type theory $\mathcal{L}_0$ we have that ϕ is provably decidable, and if there is a proof of $\neg\forall_{x:A} \neg\phi(x)$, then there must also be a constructive proof of $\exists_{x:A} \phi(x)$, i.e. (by the existence property) a proof in $\mathcal{L}_0$ of $\phi(a)$, for some closed term a of type A.

The Existence Property with a parameter of type $A = \Omega^C$: If $\vdash_{\mathcal{L}_0} \forall_{x:A} \exists_{y:B} \phi(x, y)$ then $\vdash_{\mathcal{L}_0} \forall_{x:A} \phi(x, \psi(x))$, where $\psi(x)$ is some term of type B.

A similar statement for the disjunction property with a parameter of type A is also provable. The disjunction property and already the unique existence property fail for parameters of type N, but hold in the internal language $L(\mathcal{F}(x))$, where $\mathcal{F}(x)$ is the free topos with an indeterminate $x : 1 \to N$ adjoined. The existence property in this case amounts to showing that the slice topos $\mathcal{F}/N$ is local, hence that N is projective in $\mathcal{F}$. This is equivalent to closure of the logical system under a rule of countable choice. For second order arithmetic, there is a proof due to A. Troelstra (1973, Theorem 4.5.12) based on methods of S. Hayashi (1977) which is proof-theoretic in nature. There is apparently not yet a clean categorical proof of such results.

11 Concluding Remarks

Aside from the historical discussion of our categorical approach to the foundations of mathematics, our formulation of the proof of Gödel's Incompleteness Theorem exploits the struggle between two primitive recursive functions. One enumerates all theorems and the other enumerates the Cantorian formulas which exclude the nth numeral from the nth numerical predicate. In our view, Gödel's theorem does not assert that provability fails to capture the notion of absolute truth in *the* Platonic universe. Rather, it asserts that other models of set theory are required than those which resemble the alleged Platonic universe. In fact, we have some doubts about the constructive existence of a Platonic universe, except in the context of intuitionistic (higher-order) arithmetic. Even there, the proof that our candidate, the free topos, is a model depends on the metamathematical assumption that a model of set theory exists.

References

Awodey, S. (2000) Sheaf Representation for Topoi, *Journal of Pure and Applied Algebra* **145**, 107–121.

Barwise, J. ed. (1977) *Handbook of Mathematical Logic*, Amsterdam: North-Holland, Elsevier.

Boileau, A. (1975) Types vs. Topos, Phd thesis, University of Montreal.

Boileau, A. and Joyal, A. (1981) La logique des topos, *Journal of Symbolic Logic* **46**, 6–16.

Carroll, L. (1895) What The Tortoise Said To Achilles, *Mind* **4**, No. 14, 278–280.

Church, A. (1940) A Foundation for the Simple Theory of Types, *Journal of Symbolic Logic* **5**, 56–68.

Couture, J. (1997) Analyse logique et analyticité: de Carnap à Gödel, *Dialectica* **51**(2), 95–117.

Couture, J. and Lambek, J. (1991) Philosophical Reflections on the Foundations of Mathematics, *Erkenntnis* **34**, 187–209.

Eilenberg, S. and Mac Lane, S. (1945) General Theory of Natural Equivalences, *Transactions of the American Mathematical Society* **58**, 231–294.

Feferman, S. (1977) Theories of Finite Type Related to Mathematical Practice, in: Barwise (1977), 913–971.

Freyd, P. (1978) On proving that 1 is an indecomposible projective in various free categories, manuscript.

Gödel, K. (1931) Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematische Physik* **38**, 173–198.

Gödel, K. (1958) Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica* **12**, 280–287.

Grothendieck, A. and Dieudonné, J. (1960) *Eléments de géométrie algébrique, tome I, le langage des schémas*, I.H.E.S. Publ. Math **4**, Paris, 1960. Second Edition published in *Die Grundlehren der math. Wissenschaften*, **166** (1971) Springer-Verlag.

Hayashi, S. (1977) On Derived Rules of Intuitionistic Second Order Arithmetic, *Commentarii Mathematici Universitatis Sancti Pauli* **26–1**, 77–103.

Henkin, L.A. (1950) Completeness in the Theory of Types, *Journal of Symbolic Logic* **15**, 81–91.

Howard, W.A. (1980) The Formulas-as-types Notion of Construction, in Seldin and Hindley eds. (1980), 479–490.

Lambek, J. (1994) Are the Traditional Philosophies of Mathematics Really Incompatible? *The Mathematical Intelligencer* **16**, 56–62.

Lambek, J. (2004) What is the World of Mathematics? *Annals Pure and Applied Logic* **126**, 149–158.

Lambek, J. and Moerdijk, I. (1982) Two Sheaf Representations of Elementary Toposes, in Troelstra, A.S. and van Dalen, D., eds. *The L. E. J. Brouwer Centenary Symposium* Vol 110, Studies in Logic and the Foundations of Mathematics, Amsterdam: North-Holland, pp. 275–295.

Lambek, J. and Scott, P.J. (1980) Intuitionist Type Theory and the Free Topos, *Journal of Pure and Applied Algebra* **19**, 215–257.

Lambek, J. and Scott P.J. (1986) *Introduction to Higher Order Categorical Logic*, Cambridge Studies in Advanced Mathematics 7, Cambridge: Cambridge University Press.

Lavendhomme, R. and Lucas, T. (1989) Toposes and Intuitionistic Set Theories, *Journal of Pure and Applied Algebra* **57**, 141–157.

Lawvere, F.W. (1964) An Elementary Theory of the Category of Sets. *Proceedings of the National Academy of Sciences USA* **50**, 869–872.

Lawvere, F.W. (1966) The Category of Categories as a Foundation for Mathematics, in Eilenberg, S. et. al., eds. *Proceedings of the Conference on Categorical Algebra, La Jolla 1965*, Berlin, Heidelberg and New York: Springer, pp. 1–20.

Lawvere, F.W. (1969) Adjointness in Foundations, *Dialectica* **23**, 281–296.

Lawvere, F.W. (1970) Quantifiers and Sheaves, in *Actes du Congrès International des Mathématiques*, Nice 1970, tome I, Paris: Gauthier-Villars, pp. 329–334.

Lawvere, F.W. (1972) Introduction *to Toposes, Algebraic Geometry, and Logic LNM* **274**, New York: Springer, pp. 1–12.

Mac Lane, S. and Moerdijk, I. (1992) *Sheaves in Geometry and Logic*, Universitext, New York: Springer.

Makkai, M. and Reyes, G. (1977) *First Order Categorical Logic LNM* **611**, New York: Springer.

McLarty, C. (1992) *Elementary Categories, Elementary Toposes*, Oxford: Clarendon Press.

Seldin, J.P. and Hindley, J.R. eds. (1980) *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism*, London: Academic Press.

Smorynski, C. (1977) The Incompleteness Theorems, in Barwise (1977), 821–865.

Tait, W.W. (1986) Truth and Proof: The Platonism of Mathematics, *Synthèse* **69**, 341–370.

Tarski, A. (1996) *Logic, Semantics, Metamathematics*, Oxford: Clarendon Press.

Tierney, M. (1972) Sheaf Theory and the Continuum Hypothesis, in *Toposes, Algebraic Geometry, and Logic LNM* **274**, Berlin: Springer, pp.13–42.

Troelstra, A. S. (ed) (1973) *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis* New York: Springer *LNM* **344**. (For the correction of Theorem 4.5.12 in the text, based on Hayashi's work (1977)) see Corrections to the volume, on the webpage *http://staff.science.uva.nl/~ anne*