JOCELYNE COUTURE AND JOACHIM LAMBEK

PHILOSOPHICAL REFLECTIONS ON THE FOUNDATIONS OF MATHEMATICS

ABSTRACT. This article was written jointly by a philosopher and a mathematician. It has two aims: to acquaint mathematicians with some of the philosophical questions at the foundations of their subject and to familiarize philosophers with some of the answers to these questions which have recently been obtained by mathematicians. In particular, we argue that, if these recent findings are borne in mind, four different basic philosophical positions, logicism, formalism, platonism and intuitionism, if stated with some moderation, are in fact reconcilable, although with some reservations in the case of logicism, provided one adopts a nominalistic interpretation of Plato's ideal objects. This eclectic view has been asserted by Lambek and Scott (LS 1986) on fairly technical grounds, but the present argument is meant to be accessible to a wider audience and to provide some new insights.

0. INTRODUCTION

The following conversation between two professors was overheard in the lounge of the mathematics building.

A. Did you know that at least two of the following three statements are equivalent: Fermat's so-called Last Theorem, the Riemann Hypothesis and the Continuum Hypothesis?

B. Which two?

A. That I don't know.

B. You can't say that two statements are equivalent unless you are able to prove it.

A. All I am saying is that I can prove the disjunction $(p \Leftrightarrow q) \vee (q \Leftrightarrow r) \vee (r \Leftrightarrow p)$, but I don't claim that I can prove any of the three alternatives.

B. How do you prove the disjunction?

A. Each of p, q or r can have two truth values: true or false. Therefore, at least two of them must have the same truth value.

B. Come on, we all know that there are propositions which are neither true nor false.

Without realizing it, B took an intuitionist stance; he believed that we cannot assume of any proposition p that it is true or false, that is, that

Erkenntnis **34**: 187–209, 1991.

$p \vee \neg p$ holds. An intuitionist would certainly have objected to A's claim to be able to assert a disjunction without being able to assert any of the alternatives.

Most mathematicians are quite happy to do mathematics without bothering about the foundations, not to speak of the philosophy of mathematics. Nonetheless, whether they know it or not, they do have some kind of philosophy. For example, B in the above dialogue, a distinguished analyst, was clearly defending an intuitionist position, although, when questioned, he denied being an intuitionist.

Among the many philosophical positions one can take, there are four that have been discussed widely: logicism, formalism, platonism and intuitionism. Each of these positions is based on valuable insights into what mathematics is or should be; but the answers they offer to these questions, when taken in their most radical form, appear to be incompatible. Our contention is that one can keep the best part of these insights and consistently defend an eclectic philosophy which combines mild forms of these four positions. However, if the reader should feel that, in order to achieve this synthesis, too much of the traditional positions has been sacrificed, particularly in the case of logicism, we hope he will at least accept our own position under a new name, let us call it "constructive nominalism".

We shall begin by discussing each of these positions separately, pointing out its characteristic conception of mathematics and indicating the directions in which we intend to weaken it. We shall then show how the interpretation of intuitionistic type theory in a particular model, call it "the real world of mathematics", allows for their synthesis. For the reader who wishes to know more about these four positions than he will find here, we recommend the books by van Heijenoort (1967) and Benacerraf and Putnam (1984).

1. LOGICISM

Logicism is the oldest of these schools and may be viewed as part of a reductionist program, originating with Pythagoras and pervading all science. We shall briefly discuss this wider perspective.

Pythagoras had said that all things are numbers. By "numbers" he meant primarily the positive integers, but he also admitted ratios of such, as when he said that friendship is as 284 is to 220. What he had in mind presumably was that all science reduces to mathematics, a word which he invented, and that mathematics reduces to arithmetic.

The reduction of the sciences to mathematics was incorporated into the program of the nineteenth century positivist Comte, who predicted the ultimate reduction of physics to mathematics, of chemistry to physics, of biology to chemistry, etc. Modern science has gone a good way to carry out the positivist program.

The reduction of mathematics to arithmetic ran into unexpected difficulties when it was discovered by the Pythagoreans that the ratio of the diagonal to the side of a square is irrational. (Curiously, the proof of the irrationality of $\sqrt{2}$ does not appear in the original version of Euclid's Elements but in Aristotle's Prior Analytics.) This crisis was overcome by Eudoxus, who salvaged the Pythagorean program, anticipating Dedekind, by showing that the ratio a/b of two geometric quantities was completely determined if one knew all pairs of positive integers (p, q) *such that* $pb > qa$ and $pb < qa$. (In fact, rational approximations p/q to $\sqrt{2}$ had already been calculated by the Pythagoreans by constructing solutions of the diophantine equation $p^2 - 2q^2 = \pm 1$.)

Further progress in the Pythagorean program depended on Descartes' reduction of geometry to algebra. Its ultimate success in the 19th century was summarized by Kronecker's assertion that God created the (positive) integers and everything else was created by man. (A recent spoof of this celebrated statement has it that man created the integers and that everything else is Dieudonné.)

Modern logicism asserts the possibility of reducing mathematics to logic. In view of the reduction of mathematics to arithmetic just discussed, this amounts to reducing arithmetic to logic. This program may be resolved into two separate tasks: to reduce set theory to logic and to define the natural numbers in the language of set theory. We shall first discuss the reduction of set theory to logic. This presumes that by "logic" one understands "higher order logic", in which it is possible to express membership, if only as a kind of functional application, to start with. The formal language which we have adopted (sketched in Section 7) not only contains symbols for membership and set formation, it defines the usual connectives and quantifiers in terms of these. Although one might argue that logic is thus absorbed by set theory rather than vice versa, we feel that most logicists would not object.

It was Leibniz who first dreamt of a universal language in which all of mathematics, and indeed all of science, could be expressed. Such a language was not developed until the end of the nineteenth century, when Frege invented quantifiers, leading to a formalization of logic, which was supposed to include the language of set theory. While Russell soon discovered that Frege's first formulation of the so-called comprehension scheme was inconsistent, he and Whitehead suggested a way out of the difficulty with the help of their theory of types.

There are other ways of circumventing the inconsistency inherent in the unrestricted comprehension scheme. On the whole, mathematicians, if they pay attention to foundations at all, favour the theory of Gödel-Bernays, which distinguishes between sets and classes. On the other hand, logicians interested in set theory usually favour the rival theory of Zermelo-Fraenkel, which allows comprehension only for elements of a given set. However, we believe that a form of type theory (or higher order logic) is more natural for presenting mathematics. Type theory follows natural language in refusing to lump cabbages and kings together, as when in English we are forced to distinguish between "something" and "somebody". It avoids such counterintuitive constructions as {0, {0}}, where 0 and {0} have different types. It is less prone than the first order set theories to lend itself to such meaningless questions as what is the intersection of the number π and the algebra of quaternions, both of which are usually defined as sets. It avoids Russell's and similar paradoxes from the start and not by ad hoc devices. It may be of some interest to note that some modern grammarians have adopted formal type theory from mathematical logic into so-called "categorial grammars" of natural languages (see e.g., Buszkowski et al., 1988). Even if the original type theory of Russell and Whitehead is too cumbersome for most people's taste, there are simplifications available in the literature, for instance the theory discussed by Church and Henkin (see Hintikka 1969) and more recently that used in (LS 1986).

Without becoming too technical at this stage, we just mention that, according to the last mentioned version, there should be given two basic types, a type N of individuals and a type Ω of truth values (also sometimes called propositions). From these other types can be built up by taking finite products and by the so-called power-set operation: if A is any type, there is a type PA; the entities of type PA are supposed to be sets of entities of type A.

To give a presentation of type theory, we have to describe terms of various types. For some details, see Section 7. At the moment let us only mention that terms of type Ω are called "formulas" and that among the terms of any type A there are countably many variables. Rather than introducing different styles for variables of different types, we usually write $x \in A$ (in the metalanguage) to indicate that x is a variable of type A. Quantifiers range over variables of specified type, thus we write $\forall_{x \in A} \varphi(x)$ and $\exists_{x \in A}(x)$.

If a is a term of type A and α a term of type PA, we are permitted to write down the formula $a \in \alpha$, but we are not allowed to write $a \in a$ for example.[1] We may also think of α as denoting a function which, when applied to the entity denoted by a, produces the truth-value denoted by $a \in \alpha$. The type theoretic comprehension scheme then becomes a kind of functional abstraction.

We may thus assume that the first part of the logicist program has been carried out: set theory has been reduced to logic by finding an appropriate axiomatization for higher order logic or type theory. Let us now turn to the second part of the logicist program and attempt to define the natural numbers in set theoretical notation.

What is the number 2? As far as we know, nobody takes the naive view that 2 consists of two platinum balls which are kept at room temperature somewhere in Paris. A more sophisticated view has it that 2 is the class of all unordered pairs of things, let us say that 2 is the set of all pairs $\{x, y\}$, where $x \neq y$ are individuals, that is, entities of type N. Note that $\{x, y\}$ is a subset of the set of individuals, hence an entity of type PN, hence 2 has been viewed as an entity of type $P(PN)$.

In this way we can define $0, 1, 2, 3, \ldots$ as entities of type $P(PN)$; but, of course, we must assume that there are at least n entities of type N for each natural number n. In other words, we must postulate an "axiom of infinity", which assures that there are arbitrarily many entities of type N. One way of doing this is to postulate a one-to-one correspondence between the set of individuals and a proper subset of it. Thus, we associate with each entity n of type N another entity Sn of the same type and postulate:

(P1) $\quad \forall_{x \in N} \forall_{y \in N}(Sx = Sy \Rightarrow x = y)$.

Moreover, we stipulate that a certain entity 0 of type N does not lie in the image of S, by postulating:

(P2) $\quad \forall_{x \in N}(Sx \neq 0)$.

However, once we have postulated P1 and P2, we may as well abandon our original attempt to define the natural numbers as entities of type $P(PN)$ and take them more simply as entities of type N, as was anticipated by the choice of the symbol "0". We now define 1 as $S0$, 2 as $S1$, etc. The set of natural numbers is then the smallest set u of individuals such that $0 \in u$ and $\forall_{x \in N}(x \in u \Rightarrow Sx \in u)$. To simplify matters even further, we may as well stipulate that this subset consists of all individuals, so we postulate:

$$\text{(P3)} \qquad \forall_{u \in PN}((0 \in u \wedge \forall_{x \in u}(x \in u \Rightarrow Sx \in u)) \Rightarrow \forall_{x \in N} x \in u).$$

P1, P2 and P3 are the three axioms suggested by Peano for a formal presentation of the natural numbers.

We have seen that, in order to introduce the natural numbers into the language of set theory, we must postulate an axiom of infinity or, equivalently, Peano's three axioms. Only if this is permitted, may the logicist program be said to have succeeded, albeit with some reservation. Indeed, any theorem Q of arithmetic may be obtained from a theorem in pure logic, namely

$$(P1 \wedge P2 \wedge P3) \Rightarrow Q.$$

We shall return later to an apparent objection against the logicist program raised by Gödel's incompleteness theorem.

Since writing the above words, we have become aware of a recent development in theoretical computer science, which seems to indicate that the ad hoc axiom of infinity can be avoided after all, provided one is willing to enlarge the theory of types by admitting variable types. It appears that the set of natural numbers can then be proved to exist as the solution of a certain fixpoint problem. However, the last word on this development has not yet been said.

2. FORMALISM

Formalism is the position according to which mathematics is essentially the study of a formal system, namely one in which mathematical statements can be proved, the ideal here being to do for mathematics as a whole what Euclid had attempted to do for geometry. Thus formalists are led to deny the importance attached by logicists to the logical status of arithmetical and other mathematical entities and rather confine logic to its role as a deductive apparatus for proving mathematical statements.

The early formalists had hoped that the semantical notion of mathematical truth could be captured by the syntactical notion of theorem. In attempting to show this, they were soon diverted from mathematics to metamathematics. Ironically, metamathematics turns out to be part of arithmetic, in view of the possibility of encoding a string of symbols into a single positive integer, its so-called Gödel number.

Originally, Hilbert and others (see, in particular, Kleene 1952) would permit only first order number theory into metamathematics; but, by now, all the machinery of higher order arithmetic has been admitted. The early formalists' hope that the notion of mathematical truth would be captured by the notion of theorem was dashed for classical mathematics by Gödel's incompleteness theorem, which asserts that there are closed formulas which can be neither proved nor disproved. As we shall see later, this discovery need not however discourage an intuitionist.

Hilbert had also proposed that mathematical tools be used to prove the consistency of the formal language of mathematics, that is, to show that $0 = 1$ is not a theorem. At first sight, this appears to be a rather pointless endeavour; for, if mathematics is inconsistent, then every proposition can be proved, including the proposition which asserts the consistency of mathematics. This was probably the reason why Hilbert wanted to restrict metamathematics to elementary number theory, whose consistency was not really in doubt. Unfortunately, Gödel showed, as a consequence of his incompleteness theorem, that, if elementary number theory is consistent, then its consistency cannot be proved within elementary number theory (and similarly for any formal system adequate for arithmetic).

Our treatment of formalism here has been rather cursory, as it is not meant to be a scholarly account of the history of this concept, but an attempt to compare our own position with other views, which we have taken the liberty of lumping together. In particular, it has been pointed out to us that, in addition to the *finitistic formalism* of Hilbert and the *constructive nominalism* proposed by us, one should take note of the *radical formalism* of those who claim that mathematical assertions have no content, that there are no mathematical entities and that mathematics is merely a game, and the *moderate formalism* of those who merely emphasize the importance of formal systems and formal deductions.

3. Platonism

Mathematicians who believe that mathematical entities exist independently of us and that mathematical truths are there to be discovered are called "platonists" or "realists".[2] A platonist would say, for example, that the symbol "2" denotes an entity in the real world, the number 2, and that the compound expression "1 + 1" denotes the same entity.

But what is the real world the platonists talk about? For sure, platonists are not empiricists and will not argue that 2 is just a pair of platinum balls in Paris; at any rate, there is not enough platinum in the world for all the numbers. According to the present state of physical theory, it is not at all certain that the material universe is infinite, so numbers cannot be material objects. We shall ignore here the ultrafinitists, who say that very large numbers, say larger than $10^{(10^{10})}$, do not exist.

Plato himself never identified the real world with the material universe; for him it was occupied by certain entities called "forms". Unfortunately, the Greek word for "form" developed into two distinct English words, namely "idea" and "ideal", and this has given rise to a lot of confusion. What Plato thought inhabited the real world were surely not mental ideas but ideal objects. Somehow the number 2 is an ideal object of which all concrete realizations, e.g., a pair of platinum balls, are but imperfect copies. What are these ideal entities?

Already in ancient times there were disciples of Plato who thought that his so-called forms were just concepts or words, a view which we call "the nominalistic interpretation of platonism". In modern parlance, we would say "equivalence classes of words", since the words "2" and "1 + 1" are supposed to denote the same entity. Because this interpretation of Plato's views was fashionable at the time the New Testament was written, we find that the gospel of St. John begins: "in the beginning was the word . . .".

After these historical asides, let us return to the modern formalist who is searching for a meaning of his formal expressions and who shares our predilection for such a nominalistic interpretation. He would be tempted to say that the number 2 is the equivalence class of all terms of type N which are provably equal to the term "$S(S0)$", that is to say, the set of all terms α of type N for which there is a proof of the formula "$\alpha = S(S0)$". This will account for the fact that "1 + 1" and "$S(S0)$" denote the same number. Why is this not the accepted view?

Using classical logic, it is fairly easy to prove the following formula:

$$\exists!_{x\in N}((x = 0 \wedge p) \vee (x = S0 \wedge \neg p)),$$

where p is some given formula. Let us abbreviate this as $\exists!_{x\in N}\varphi(x)$. Most mathematicians would now be willing to introduce a number which is 0 if p is true and 1 if p is false. Such a number can be represented by a closed term in our formal language if we suppose that it contains a description operator à la Russell or only a minimization operator $\mu_{x\in N}$. Even if the language does not contain such an operator, a conservative extension of the language will. We then obtain the closed term

$$\alpha \equiv \mu_{x\in N}\varphi(x),$$

meaning "the least (and only) $x \in N$ such that $\phi(x)$", so that $\phi(\alpha)$ is a theorem. The equivalence class $[\alpha]$ of α should be a number inhabiting our platonic universe.

Can we assume that $[\alpha] = [S^m0]$, that is, that $\alpha = S^m0$ is a theorem? If so, we have the theorem:

$$(S^m0 = 0 \wedge p) \vee (S^m0 = S0 \wedge \neg p).$$

In case $m = 0$, we can prove $S^m0 \neq S0$, hence p. In case $m \neq 0$, we can prove $S^m0 \neq 0$, hence $\neg p$. Thus our assumption that $[\alpha]$ is of the form $[S^m0]$ leads to the conclusion that p can be either proved or disproved. Unfortunately, Gödel has shown that, as long as our formal system is adequate to handle arithmetic and consistent, that is, $0 = 1$ cannot be proved, there will be a closed formula p which can be neither proved nor disproved. (This is Rosser's version.) We must therefore discard the assumption that $[\alpha]$ is one of the numbers 0, 1, 2, . . . and infer that our tentative platonic universe contains numbers other than those we learned about in kindergarten.

While a classical platonist might see Gödel's incompleteness theorem as a deathblow to the formalist program, a classical formalist would see it as a deathblow to nominalistic platonism!

Our argument involved numbers, that is, entities of type N. We could equally well have argued about truth values, that is, entities of type Ω, assuming these to be equivalence classes of closed formulas. If p is provable, its truth value is $\top$(true) and, if p is disprovable, its truth value is $\bot$(false). If p is neither provable nor disprovable, its truth value must be something else. This would contradict an ancient and, until recently, sacred dogma, first enunciated by Aristotle: every statement is true or false. If this so-called principle of the excluded third is taken as an axiom, say in the form $\forall_{x\in\Omega}(x \vee \neg x)$, it follows that $p \vee \neg p$ is provable for any formula p. Incidentally, the same principle was used in the proof of $\exists!_{x\in N}\varphi(x)$ mentioned above.

To sum up, the naive nominalistic attempt to reconcile formalism and platonism does not work as long as we subscribe to the Aristotelian doctrine that every proposition is true or false.

4. INTUITIONISM

As illustrated in our opening dialogue, an intuitionist does not believe in the principle of the excluded third, namely that every statement is either true or false. Thus, he does not accept as a principle of reasoning that, for all propositions p, $p \vee \neg p$ or, equivalently, $\neg\neg p \Rightarrow p$.

Why should one reject a principle which seems to be confirmed repeatedly by everyday experience, at least as far as declarative sentences are concerned?

Brouwer appears to have been the first to seriously challenge this venerable principle, which had been stated by Aristotle and which has been used by mathematicians all along. While Brouwer would have admitted that this principle can be used safely when reasoning about the finite quantities one meets in the material world, he claimed that it breaks down when one deals with the infinite quantities of mathematics.

For example, according to Aristotle's principle, a real number must be either rational or irrational. Brouwer would argue that we cannot know this by looking at any finite part of its decimal expansion, but only when we have either expressed it in the form p/q, with integers p and q, or given a proof of irrationality, such as the famous classical proof of the irrationality of $\sqrt{2}$ quoted by Aristotle.

Intuitionism had a precursor in the constructivism of Kronecker. To illustrate the kind of argument a constructivist would object to, we present the following example due to van Dalen: there exist two irrational numbers α and β such that α^{β} is rational. We argue that $\sqrt{2}^{\sqrt{2}}$ is either rational or irrational. In the first case take $\alpha = \sqrt{2}$ and $\beta = \sqrt{2}$, in the second case take $\alpha = \sqrt{2}^{\sqrt{2}}$ and $\beta = \sqrt{2}$, so that $\alpha^{\beta} = \sqrt{2}^{2} = 2$.

Let p be the statement: $\sqrt{2}^{\sqrt{2}}$ is rational. Then α is the least (and unique) $x \in N$ such that

$$(p \wedge x = \sqrt{2}) \vee (\neg p \wedge x = \sqrt{2}^{\sqrt{2}}).$$

The existence of such an x is assured in classical mathematics by the assumption that $p \vee \neg p$, but this is precisely what is in question.

Actually, we happen to know that $\sqrt{2}^{\sqrt{2}}$ is irrational by a deep result due to Gelfand, but this is irrelevant to our argument. It also so happens that there is an easy constructive proof of the same theorem anyway: take $\alpha = \sqrt{2}$ and $\beta = 2 \log_2 3$, so that $\alpha^\beta = 3$; but this is irrelevant also. What matters is that the proof suggested by van Dalen is non-constructive because it is based on the principle of the excluded third.

It is a surprising fact, discovered by Brouwer, that even without the assumption $p \vee \neg p$, for all propositions p, or the equivalent assumption that $\neg\neg p \Rightarrow p$ for all p, also due to Aristotle, one can do most of mathematics, namely the constructive part of mathematics. To this we may add that, if we list $\forall_{x \in \Omega}(x \vee \neg x)$ as a premiss to intuitionistic logic, we can of course do all of classical elementary mathematics, at least as long as we stay away from the axiom of choice. Thus, taking $p \equiv \forall_{x \in \Omega}(x \vee \neg x)$, a mathematical statement q is a theorem classically if and only if $p \Rightarrow q$ is a theorem intuitionistically.

Most of Brouwer's arguments were epistemic, concerning the nature of mathematical knowledge. They were directed against a platonistic conception of mathematical truth which presupposes a universe in which the truth or falsity of mathematical statements is fixed once and for all. But according to Brouwer, the truth of such a statement results only from a constructive process, namely that of finding and displaying a proof, apparently a product of the human mind. Unfortunately, the accident of discovery would make truth dependent on time; thus Fermat's theorem, of doubtful status today, might only become true in the next century. Few present day intuitionists would go as far, except metaphorically for purposes of illustration or to motivate Kripke models.

Brouwer's epistemic arguments also seemed to attack formalism: mathematics as a constructive process is neither a linguistic activity nor a formal game. However, Brouwer himself later felt that formalism was not incompatible with intuitionism. The formalization of intuitionistic logic was worked out by Heyting and Kleene. In fact, most present members of the Dutch intuitionist school make liberal use of formalist tools to justify their principles.

To illustrate how intuitionistic principles can be proved formally, let us sketch a proof of the disjunction principle: if $p \vee q$ is provable then either p is provable or q is provable. So suppose $p \vee q$ has been proved, then one can also prove:

$$\exists_{x \in N}((p \wedge x = 0) \vee (q \wedge x = 1)).$$

Let us abbreviate this as $\exists_{x \in N}\varphi(x)$ and consider the term $\mu_{x \in N}\varphi(x)$ of type N. Now it can be shown that, in pure intuitionistic type theory, every term of type N is provably equal to a standard numeral of the form S^n0. (For the proof of this, see e.g., LS 1986.) Therefore, one can prove $\mu_{x \in N}\varphi(x) = S^n0$, hence $\varphi(S^n0)$, that is,

$$(p \wedge S^n0 = 0) \vee (q \wedge S^n0 = S0).$$

In case $n = 0$ we obtain a proof of p, in case $n \neq 0$ we have a proof of q.

Why should a mathematician who believes in the principle of the excluded third study intuitionism? He might be induced to suspend his belief for the purpose of argument, if he could be persuaded that this exercise would pay off, aside from the fact that, as a formalist, he should have no objection to investigating the consequences of an axiomatic system which omits the principle of the excluded third.

Mathematicians are interested in certain categories called toposes, which first made their appearance in algebraic geometry. We shall make an attempt to convey the idea of what a topos is in Section 8 below. Prominent among toposes is the topos of sheaves on a topological space. While the concept of sheaf may be too technical for the non-specialist, a sheaf may be viewed as a variable set (Lawvere 1975). It was recently shown by Barr (1986) and Bénabou (unpublished) that the notion of sheaf also subsumes the popular concept of fuzzy set, provided equality is also allowed to be fuzzy. These toposes possess an internal language which, in general, turns out to be intuitionistic and obeys the principle of the excluded third only in some special cases. One important consequence of this is that any theorem about sets which has been proved using only intuitionistic logic must hold about the objects of any topos, for example about sheaves. In other words, a theorem proved intuitionistically about sets holds more generally than stated and may be interpreted as a new classical theorem about sheaves.

However, the most profound effect which intuitionism has had on the general mathematical community has been to direct attention to constructive proofs.

5. CONSTRUCTIVE NOMINALISM

Let us recall our earlier nominalistic attempt to reconcile formalism and platonism by saying that entities in the real world are provable equivalence classes of terms. Actually, it is not necessary to invoke the notion of equivalence class, for example, to explain the number 2. All we are doing is to work with terms such as $SS0$ and $S0 + S0$, but with a new equality relation between them: two (in the old sense) terms are to be considered equal (in the new sense) if they are provably equal in the formal language.[3]

While this attempted reconciliation did not work in the framework of classical logic, we would now like to defend the position that it does work if one adopts a moderate intuitionist position. In fact, the two arguments presented above to demolish the attempted reconciliation depended on the principle of the excluded third, the provability of $p \vee \neg p$, which is precisely what the intuitionist questions.

Of course, the mere fact that a counterargument fails does not prove that the attempt works. The proof that it does is somewhat tedious and we shall only give a sketch here. (For further details see LS 1986.) It depends on the notion of model for intuitionistic higher order logic or, as we shall call it here, the interpretation of intuitionistic type theory in a possible world. We are talking here about the possible worlds of mathematics; this does not include the material universe, as long as we do not know that it is infinite.[4]

The possible worlds in which we wish to interpret our formal language are certain toposes (which are assumed to have natural numbers objects), namely those toposes whose internal language satisfies the following three conditions:

(1) not every proposition is true;
(2) if $p \vee q$ is true then either p is true or q is true;
(3) if $\exists_{x \in A} \varphi(x)$ is true then $\varphi(a)$ is true for some entity a of type A.

One has a Gödel-Henkin style completeness theorem: a formula of any intuitionistic type theory (not just the pure one, e.g., the internal language of a topos) is provable if and only if it is true under every interpretation in a possible world.

We shall try to show that among all the possible worlds there is one that stands out from the others; it is constructed linguistically. To begin with, we must make sure that pure intuitionistic type theory has been formulated in such a way as to contain sufficiently many terms of each type. By this we mean that, whenever one can prove $\exists!_{x\in A}\varphi(x)$, there is a term a for which one can prove $\varphi(a)$. One way of doing this is to admit a Russellian description operator $1_{x\in A}$ so that we can write $a \equiv 1_{x\in A}\varphi(x)$. However, for pure intuitionistic type theory, this turns out to be unnecessary, provided we admit enough special instances of the description operator, for example $\{x \in A \mid \varphi(x)\}$ of type PA for

$$1_{u\in PA}\forall_{x\in A}(x \in u \Leftrightarrow \varphi(x)).$$

In particular, one does not have to adjoin the minimization scheme $\mu_{x\in N}\varphi(x)$ when $\exists_{x\in N}\varphi(x)$ is provable, as one gets it for free. Even so all closed terms of type N have the form S^n0 for some $n \geq 0$.

Given a presentation of pure intuitionistic type theory with enough terms, one can construct a topos by identifying terms of the same type which are provably equal. To be precise, its objects, called *sets*, are provable equivalence classes of terms of type PA, for any type A, and its arrows, say between a set α of type PA and a set β of type PB, are provable equivalence classes of terms ρ of type $P(A \times B)$ which can be proved to satisfy the usual conditions for a *function* from α to β:

$$\vdash \rho \subseteq \alpha \times \beta \wedge \forall_{x\in A}(x \in \alpha \Rightarrow \exists!_{y\in B}\langle x, y\rangle \in \rho).$$

For details of this construction we must refer the reader to (LS 1986); where it is also shown that this linguistic topos is a possible world, in the sense of satisfying (1) to (3) above. It is usually called the *free topos*. (It may also be characterized abstractly as the initial object in the category of all toposes.)

It turns out that in the free topos all entities of type N are "standard": they have the form S^n0 for some natural number n. Unfortunately, this is not true in all possible worlds, as follows from the proof of Gödel's incompleteness theorem.

It also follows from Gödel's theorem that this linguistic construction of a possible world won't work in classical mathematics, as we pointed out earlier. While one can construct the free "Boolean" topos (that is, a topos satisfying Aristotle's law of the excluded third) in the same manner, it is not a possible world to satisfy a classical mathematician, nor an intuitionist for that matter. For example, the free Boolean topos contains a proposition p for which $p \vee \neg p = \top$ but neither $p = \top$ nor

$\neg p = \top$. Moreover, it contains entities of type N which are not of the form S^n0 for any natural number n.

On the other hand, the free topos (constructed from pure intuitionistic type theory) seems to be acceptable to moderate intuitionists as a possible world. It has one property which they expect the real world to have: "true in the free topos" means "provable intuitionistically". We therefore suggest that the linguistically constructed free topos be accepted as the *real* (= ideal) *world of mathematics*, at least of elementary mathematics. To put this proposal to the test, we shall take another look at Gödel's incompleteness theorem.

The semantic version of this theorem asserts that there is a formula g which is true but not provable, thus apparently confounding the formalist and the logicist at the same time. This interpretation of what Gödel showed is all very well from the standpoint of a classical platonist, which Gödel was, as it leaves open the question "true where?"

If one examines Gödel's argument closely (see Couture 1985), one realizes that what he actually showed is that g is true in any model with the following property: if $\varphi(S^n0)$ is true in the model for each natural number n, then $\forall_{x \in N} \varphi(x)$ is true in that model. By the completeness theorem mentioned above, we then merely infer that not every model has the assumed property. In fact, the free topos does not have it and Gödel's formula g is false in the free topos. Anyway, intuitionists would not accept the property in question, unless it is modified by specifying that the reason for asserting $\varphi(S^n0)$ is the same for each n.

Let us summarize once more the four current philosophies in somewhat oversimplified form in order to extract from each what we are willing to accept, culminating in our own nominalistic synthesis.

> Logicism: Mathematical entities can be defined in the language of symbolic logic.
> Formalism: Whether mathematical entities exist or not, what matters are the terms of a formal language which supposedly describe them. Whether mathematical statements are true in an absolute sense or not, what matters is whether they can be proved.
> Platonism: Mathematical entities exist independently of our way of viewing them; mathematical truths are there to be discovered.

> Intuitionism: Mathematical entities are mental constructs, mathematical truths are statements capable of being known.

With some reservation, we accept the logicist reduction of mathematics to logic via the theory of types, but only provided Peano's postulates for the natural numbers are included. We are willing to think like intuitionists, temporarily suspending our belief in the principle of the excluded third, in order to find constructive proofs and to gain greater generality for our results. On the other hand, we do not hesitate to investigate consequences of the hypothesis $\forall_{x \in \Omega}(x \vee \neg x)$, particularly in metamathematics. We admit that only such propositions are true which are knowable, that is, provable, but we believe that truth is independent of the historical accident that a proof has already been discovered. We are prepared to study the formal language of intuitionistic type theory and to make use of any mathematical tools to establish intuitionistic principles as metatheorems. We allow that pure intuitionistic type theory describes a "real" world whose entities are equivalence classes of terms of the formal language, two terms α and β being equivalent if the statement $\alpha = \beta$ is provable.

As an illustration of this nominalistic synthesis of the four basic philosophical positions, we shall look at the problem of substitutional interpretation of quantifiers which has lately been discussed by philosophers.

6. SUBSTITUTIONAL INTERPRETATION OF QUANTIFIERS

Much has been written by philosophers about the question of substitutional interpretation of quantifiers. In its simplest formulation, the question takes the form: is the truth of the quantified statements $\exists_{x \in A} \varphi(x)$ and $\forall_{x \in A} \varphi(x)$ dependent on the truth of the statement $\varphi(a)$ for some or all closed terms a of the language respectively? The problem is of course that not all entities the formal language is claimed to be about need have names in the language. Nonetheless, the question is easily answered in the affirmative for languages specially constructed in a hierarchic fashion (Kripke 1976) or for languages which permit one to argue inductively on the complexity of formulas, in which $\varphi(a)$ is necessarily simpler than $\exists_{x \in A} \varphi(x)$ or $\forall_{x \in A} \varphi(x)$, as in Russell's ramified type theory without reducibility (see Couture 1989) or in Martin-Löf's (1984) constructive type theory. But neither of these theories is adequate for mathematics, so what about pure intuitionistic type theory?

Since quantifiers can be interpreted in any possible world, they can be interpreted in the real world, namely the free topos. Thus, the formula $\exists_{x\in A}\varphi(x)$ of pure intuitionistic type theory is true in the real world if and only if $\varphi(a)$ is true for some entity a of type A in the real world. But since the real world was created linguistically, we can say that $\exists_{x\in A}\varphi(x)$ is true if and only if $\varphi(a)$ is true for some term a of type A in the language. Therefore, the substitutional interpretation of existential quantifiers is valid for pure intuitionistic type theory.

Strictly speaking, the above argument applies only to formulas $\varphi(x)$ which contain no free variables other than x. The question is still partly open if there are other free variables, for example a free variable of type $\Omega \times N$.

What about universal quantifiers? One might hope that the formula $\forall_{x\in A}\varphi(x)$ of pure intuitionistic type theory is true in the real world if and only if $\varphi(a)$ is true for all closed terms a of type A. Unfortunately, this is not so, as follows from Gödel's proof of the incompleteness theorem. For Gödel constructed a formula $\varphi(x)$, with x a variable of type N, such that $\varphi(S^n 0)$ is provable for all $n \geqslant 0$ but $\forall_{x\in N}\varphi(x)$ is not provable. His proof remains valid about intuitionistic type theory, where "provable" means "true". Moreover, in pure intuitionistic type theory, all closed terms of type N are provably equal to some standard numeral of the form $S^n 0$. This argument shows that the naive attempt to interpret universal quantifiers substitutionally by means of closed terms does not work. Of course, if we allow substitution by open terms, the substitutional interpretation is valid trivially, since $\forall_{x\in A}\varphi(x)$ is provable if and only if the open formula $\varphi(x)$ is provable.

What about pure classical type theory, which we may take to be pure intuitionistic type theory with an additional axiom, namely $\forall_{x\in\Omega}(x \vee \neg x)$ or $\forall_{x\in\Omega}(\neg\neg x \Rightarrow x)$? Then $\forall_{x\in A}\varphi(x)$ is equivalent to $\neg\exists_{x\in A}\neg\varphi(x)$, by De Morgan's Rule, so we only have to look at the interpretation of existential quantifiers. Unfortunately, the substitutional interpretation of existential quantifiers does not hold in pure classical type theory, as is seen from the following example.

Let $\Psi(u)$ be the formal statement which asserts that u of type $P(PN)$ is a non-principal ultrafilter of sets of natural numbers. Then one easily proves classically:

$$\exists_{v\in P(PN)}(\exists_{u\in P(PN)}\Psi(u) \Rightarrow \Psi(v)),$$

yet no one has ever found a description of a non-principal ultrafilter, that is, a closed term α of type $P(PN)$ so that $\Psi(\alpha)$ can be inferred from the assumption $\exists_{u \in P(PN)} \Psi(u)$, nor does such a term exist.

7. APPENDIX ON INTUITIONISTIC TYPE THEORY

For the reader's convenience we give a brief sketch of a recent formulation of pure intuitionistic type theory, which is adequate for elementary mathematics, including arithmetic and analysis, when treated constructively. As far as we know, the only theorems in these disciplines which are essentially non-constructive are those whose proof requires the axiom of choice. Even there it is not easy to find an example of a theorem for which no constructive proof can be shown to exist (see above). Constructive arithmetic has been treated in some fashion by Goodstein (1970) and the standard text on constructive analysis is by Bishop (1967).

From basic types 1, Ω and N one builds others by two processes: if A is a type so is PA; if A and B are types so is $A \times B$. Intuitively:

> 1 is the type of a specified single entity (introduced for convenience);
> Ω is the type of truth-values or propositions;
> N is the type of natural numbers;
> PA is the type of sets of entities of type A;
> $A \times B$ is the type of pairs of entities of types A and B respectively.

We allow arbitrarily many variables of each type and write $x \in A$ to say that x is a variable of type A. In addition, we construct terms of different types inductively as follows:

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

it being assumed that a and a' are terms of type A already constructed, α of type PA, n of type N, $\varphi(x)$ of type Ω and b of type B.

Logical symbols may be defined as follows:

$$\top \equiv * = *,$$
$$p \wedge q \equiv \langle p, q\rangle = \langle \top, \top\rangle,$$

$$p \Rightarrow q \equiv p \wedge q = p,$$
$$\forall_{x \in A} \varphi(x) \equiv \{x \in A \mid \varphi(x)\} = \{x \in A \mid \top\},$$

where it is understood that p, q and $\varphi(x)$ are terms of type Ω. From these symbols one may define others, taking care not to make implicit use of De Morgan's rules (Prawitz 1965):

$$\bot \equiv \forall_{t \in \Omega} t,$$
$$\neg p \equiv \forall_{t \in \Omega}(p \Rightarrow t),$$
$$p \vee q \equiv \forall_{t \in \Omega}(((p \Rightarrow t) \wedge (q \Rightarrow t)) \Rightarrow t),$$
$$\exists_{x \in A} \varphi(x) \equiv \forall_{t \in \Omega}((\forall_{x \in A}(\varphi(x) \Rightarrow t)) \Rightarrow t).$$

Other symbols, such as appear in $\exists!_{x \in A} \varphi(x)$, $\{a\}$, $\alpha \subseteq \beta$, $\alpha \times \beta$ etc are defined in the usual fashion.

Axioms and rules of inference are stated in terms of a deduction symbol $\vdash_X$, where X is a finite set of variables. The permissible deductions take the form

$$p_1, \ldots, p_n \vdash_X p_{n+1},$$

where the p_i are terms of type Ω and X contains all the variables which occur freely in the p_i. The axioms and rules of inference hold no surprises. Here are a few special cases for purpose of illustration:

$$p \vdash p;$$
$$\frac{\varphi(x) \vdash_x \psi(x)}{\varphi(a) \vdash \psi(a)},$$
$$\langle a, b\rangle = \langle c, d\rangle \vdash a = c;$$
$$\frac{\varphi(x) \vdash_x \varphi(Sx)}{\varphi(0) \vdash_x \varphi(x)}.$$

As already mentioned, the language sketched here provides adequate foundations for constructive arithmetic and analysis. However, for metamathematics, as practiced nowadays, and this includes category theory, more powerful languages are required. This is because we may need higher types than those contained in the hierarchy constructed above, for instance when we speak of the category of all toposes, and because we may require the axiom of choice, as for example in the proof of the completeness theorem. It also appears that the needs of computer scientists have not been met, who would wish proofs to be incorporated into the language on the same level as terms. Moreover, it has been suggested that, in order to benefit fields as varied as linguistics, computer science and quantum mechanics, one should replace intuitionistic logic by linear logic (Girard 1987).

8. APPENDIX ON THE NOTION OF A TOPOS

Following the appearance of certain concrete toposes in algebraic geometry, the abstract notion of a topos was conceived by Lawvere and elaborated by him and Tierney. Their definition may be carried out in the language of category theory, which itself may be expressed in a formal language, unfortunately one more powerful than that outlined in Section 7, as was already pointed out there.

How can one convey the idea of categories and toposes without going into technical details? In category theory the emphasis is on morphisms (arrows) between structures (objects) and the former are by no means subordinate to the latter, as Bell (1988) puts it. He continues: "So category theory is like a language in which the 'verbs' are on equal footing with the 'nouns'."

A *topos* is an abstractly described category in which it is possible to carry out a number of important constructions usually carried out in the category of sets. In this general context one says "objects" rather than "sets" and "arrows" rather than "functions". In particular, there is assumed to be a *terminal* object 1 corresponding to a typical one element set $\{*\}$. One can form the *cartesian product* $A \times B$ of any two objects corresponding to the set of all pairs of elements of A and B. One can form the *power object* PA corresponding to the set of all subsets of A. Given any two arrows $f: A \to B$ and $g: A \to B$, one can imitate the construction which for sets yields $\{a \in A \mid f(a) = g(a)\}$. One can characterize the subobjects of a given object by their *characteristic* arrow into the object of *truth values*, just as every subset of a set A has the form $\{a \in A \mid h(a) = \top\}$, where $h: A \to \{\top, \bot\}$ is the usual characteristic function. Finally, there is given an object N of *natural numbers* with arrows $1 \to N$ (zero) and $N \to N$ (successor) allowing one to imitate the following construction for sets: given a set A, an element a of A and a function $h: A \to A$, one can construct a uniquely determined function $f: N \to A$ such that, for each n in N, $f(n) = h^n(a)$.

In Section 5 above we presented a construction of the so-called free topos from *pure* intuitionistic type theory. By the same process one can associate a topos with an *applied* intuitionistic type theory as well. An applied theory differs from the pure one in the following respects: (a) there may be types other than those contained in the hierarchy described in Section 7 and there may be equations holding between them; (b) there may be terms other than those described in Section 7 and there may be additional axioms.

On the other hand, starting with any topos whatsoever, one can construct an applied intuitionistic type theory, its "internal language", which was already mentioned in Section 4. The idea here is that all objects of the topos are admitted as types and that arrows from 1 to A are taken as terms of type A. In particular, one exploits the abstract properties of $A \times B$ to construct an arrow $\langle a, b \rangle : 1 \to A \times B$ from given arrows $a : 1 \to A$ and $b : 1 \to B$ and views it as a term of type $A \times B$. Similarly, one exploits the abstract properties of PA to construct an arrow $a \in \alpha : 1 \to \Omega$ from given arrows $a : 1 \to A$ and $\alpha : 1 \to PA$ and views it as a term of type Ω. And so on. As in pure intuitionistic type theory, one defines the term $p \wedge q$, thus obtaining an arrow $p \wedge q : 1 \to \Omega$. One is now in the position to define the deduction symbol. For example, $p \vdash q$ is taken to mean that the arrows $p \wedge q$ and p from 1 to Ω coincide. When we said in Section 5 that a proposition p is *true* in the internal language, we meant nothing else than that the arrows p and $\top$ from 1 to Ω coincide.

More detailed expositions of topos theory are available in the books by Goldblatt (1979) and Bell (1988).

NOTES

[1] We have followed Curry (1963) and the usual mathematical practice and used symbols such as "$\in$" to denote themselves and concatenation to represent concatenation. Thus $a \in \alpha$ is meant to be the result of placing the term a before and the term α after the symbol "$\in$". Sometimes, a half-hearted use of ordinary quotation marks may creep into our presentation, as in this footnote. To be quite rigorous about the "use versus mention" distinction, one would have to use some device such as Quine's corners (Quine 1958), which we feel would be out of place here.

[2] As far as we know, the word "platonism" was introduced into mathematics by Bernays (see Benacerraf and Putnam 1984).

[3] If α and β are terms of type A, we consider α to be equal to β in the new sense if the formula $\alpha = \beta$ is provable, more precisely, if the result of putting α before and β after the equal sign is a provable formula. See also note 1.

[4] The expression "possible world", although borrowed from Leibniz, here has a precise meaning, namely that of a topos satisfying (1), (2) and (3). It is not to be confused with

its homonyms used in the interpretation of natural languages, of quantum mechanics or of modal logic.

NOTE ADDED IN PROOF

The authors wish to express their indebtedness to the referee, whose searching questions helped them to elaborate and improve their argument.

REFERENCES

Barr, M.: 1986, 'Fuzzy Set Theory and Topos Theory', *Can. Math. Bull.* **29**, 501–508.

Bell, J. L.: 1988, *Toposes and Local Set Theories*, Oxford Logic Guides 14, Clarendon Press, Oxford.

Benacerraf, P. and H. Putnam (eds.): 1984, *Philosophy of Mathematics: Selected Readings*, 2nd edn., Cambridge University Press, Cambridge Massachusetts.

Bishop, E.: 1967, *Foundations of Constructive Analysis*, McGraw-Hill, New York.

Buszkowski, W. et al. (eds.): 1988, *Categorial Grammar*, John Benjamins, Amsterdam.

Couture, J.: 1985, *Analyticité et complétude*, *Cahiers d'épistémologie*, Dépt. de philosophie de UQAM 8510, also in R. Nadeau (ed.), *Contingence et Raison*, Vrin-Bellarnin (to appear).

Couture, J.: 1989, 'Are Substitutional Quantifiers a Solution to the Problem of Elimination of Classes in Principia Mathematica?', in I. Winchester and K. Blackwell (eds.), *Antinomies and Paradoxes*, McMaster University Press, Hamilton.

Curry, H.: 1963, *Foundations of Mathematical Logic*, McGraw-Hill, New York.

Girard, J. Y.: 1987, 'Linear Logic', *J. Theoretical Computer Science* **50**, 1–102.

Goldblatt, R. 1979, *Topoi, the Categorial Analysis of Logic*, Studies in the Foundations of Mathematics 98, North-Holland, Amsterdam.

Goodstein, R. L.: 1970, *Recursive Number Theory*, Academic Press, New York.

Hintikka, J.: 1969, *The Philosophy of Mathematics*, Oxford University Press.

Kleene, S. C.: 1952, *Introduction to Metamathematics*, Van Nostrand Reinhold, New York.

Kripke, S.: 1976, 'Is There a Problem about Substitutional Quantification?', in Evans and MacDowell (eds.), *Truth and Meaning*, Clarendon Press, Oxford, pp. 325–419.

Lambek, J. and P. J. Scott: 1986, *Introduction to Higher Order Categorical Logic*, Cambridge University Press, Cambridge.

Lawvere, F. W.: 1975, 'Continuously Variable Sets: Algebraic Geometry = Algebraic Logic', in H. E. Rose and J. C. Sheperdson (eds.), *Proc. A.S.L. Logic Coll. Bristol*, 1974, North-Holland Publishers Amsterdam pp. 135–156.

Martin-Löf, P.: 1984, *Intuitionistic Type Theory*, Bibliopolis, Naples.

Prawitz, D.: 1965, *Natural Deduction*, Almquist and Wiksell, Stockholm.

Quine, W.: 1958, *Mathematical Logic*, Harvard University Press, Cambridge Mass.

van Heijenoort J. (ed.): 1967, From Frege to Gödel, Harvard University Press, Cambridge Massachusetts.

FURTHER READING

Goodman, N. D.: 1978, 'Mathematics as an Objective Science', *Amer. Math. Monthly* **86**, 540–555.

Kneale, W. and M. Kneale: 1971, *The Development of Logic*, Oxford University Press, Oxford.

Lambek, J.: 1989, 'On the Sheaf of Possible Worlds', in J. Adamek and S. Mac Lane (eds.), *Categorial Topology*, World Scientific, Singapore.

MacLane, S.: 1981, 'Mathematical Models: A Sketch for the Philosophy of Mathematics', *Amer. Math. Monthly* **88**, 462–472.

Russell, B and A. N. Whitehead: 1910–1913, *Principia Mathematica I-III*, Cambridge University Press, Cambridge.

Tait, W. W.: 1986, 'Truth and Proof: the Platonism of Mathematics', *Synthese* **69**, 341–370.

Manuscript submitted November 7, 1989
Final version received May 31, 1990

J. Couture
Département de Philosophie
Université du Québec à Montréal
Montréal
Canada

J. Lambek
Mathematics Department
McGill University
Montreal
Canada