SOLOMON FEFERMAN

CATEGORICAL FOUNDATIONS AND FOUNDATIONS OF CATEGORY THEORY*

ABSTRACT. This paper is divided into two parts. Part I deals briefly with the thesis that category theory (or something like it) should provide the proper foundations of mathematics, in preference to current foundational schemes. An opposite view is argued here on the grounds that the notions of operation and collection are prior to all structural notions. However, no position is taken as to whether such are to be conceived extensionally or intensionally.

Part II describes work by the author on a new non-extensional type-free theory $\hat{T}$ of operations and classifications.[1] Its interest in the present connection is that much of 'naive' or 'unrestricted' category theory can be given a direct account within $\hat{T}$. This work illustrates requirements and possibilities for a foundation of unrestricted category theory.

I

The reader need have no more than a general idea of the nature of category theory to appreciate most of the issues discussed below. MacLane [15] gives a succinct account which is particularly related to these questions. Two views are intermixed in [15] as to current set-theoretical foundations, namely that (i) they are *inappropriate* for mathematics as practised, and (ii) they are *inadequate* for the full needs of category theory. The latter is taken up in Part II below. The view (i) evidently derives from the increasingly dominant conception (by mathematicians) of mathematics as the study of abstract structures. This view has been favored particularly by workers in category theory because of its successes in organizing substantial portions of algebra, topology, and analysis. It is perhaps best expressed by Lawvere [13]: "In the mathematical development of recent decades one sees clearly the rise of the conviction that the relevant properties of mathematical objects are those which can be stated in terms of their abstract structure rather than in terms of the elements which the objects were thought to be made of. The question thus naturally arises whether one can give a foundation for mathematics which expresses wholeheartedly this conviction concerning what mathematics is about and in particular

Butts and Hintikka (eds.), Logic, Foundations of Mathematics and Computability Theory, 149–169.

in which classes and membership in classes do not play any role." Further: "A foundation of the sort we have in mind would seemingly be much more natural and readily usable than the classical one ..." Lawvere went on in [13] to formulate a (first-order) theory whose objects are conceived to be arbitrary categories and functors between them.[2] Each object in Lawvere's theory is thus part of a highly structured situation which must be prescribed axiomatically by the elementary theory of categories.

There are several objections to such a view and program,[3] among which are arguments for what is achieved in set-theoretical foundations that is not achieved in other schemes (present or projected). I wish to stress here instead a very simple objection to it which is otherwise neutral on the question of 'proper foundations' for mathematics: the argument given is itself not novel.[4]

The point is simply that *when explaining* the general notion of structure and of particular kinds of structures such as groups, rings, categories, etc. we implicitly *presume as understood* the ideas of *operation* and *collection*; e.g. we say that a group consists of a collection of objects together with a binary operation satisfying such and such conditions. Next, when explaining the notion of *homomorphism* for groups or *functor* for categories, etc., we must again understand the concept of operation. Then to follow category theory beyond the basic definitions, we must deal with questions of *completeness*, which are formulated in terms of collections of morphisms. Further to verify completeness in concrete categories, we must be able to form the operation of *Cartesian product* over collections of its structures. Thus at each step we must make use of the unstructured notions of operation and collection to explain the structural notions to be studied. The *logical* and *psychological priority* if not primacy of the notions of operation and collection is thus evident.

It follows that a theory whose objects are supposed to be highly structured and which does not explicitly reveal assumptions about operations and collections cannot claim to constitute a foundation for mathematics, simply because those assumptions are unexamined. It is evidently begging the question to treat collections (and operations between them) as a category which is supposed to be one of the objects of the universe of the theory to be formulated.

The foundations of mathematics must still be pursued in a direct examination of the notions of operation and collection. There are at present only two (more or less) coherent and comprehensive approaches to these, based respectively on the Platonist and the constructivist viewpoints. Only the first of these has been fully elaborated, taking as basis the conception of sets in the cumulative hierarchy. It is distinctive of this approach that it is *extensional*, i.e., collections are considered independent of any means of definition. Further, operations are identified with their graphs.

On the other hand, it is distinctive of the constructive point of view that the basic notions are conceived to be *intensional*, i.e. operations are supposed to be *given by rules* and collections are supposed to be *given by defining properties*.[5] Theories of such are still undergoing development and have not yet settled down to an agreed core comparable to the set theories of Zermelo or Zermelo-Fraenkel. Nevertheless, a number of common features appear to be emerging, for example, in the systems proposed by Scott [20], Martin-Löf [16], and myself [5].

Since neither the realist nor constructivist point of view encompasses the other, there cannot be any present claim to a *universal foundation* for mathematics, unless one takes the line of rejecting all that lies outside the favored scheme. Indeed, *multiple foundations* in this sense may be necessary, in analogy to the use of both wave and particle conceptions in physics. Moreover, it is conceivable that still other kinds of theories of operations and collections will be developed as a result of further experience and reflection. I believe that none of these considerations affects the counter-thesis of this part, namely that foundations for structural mathematics are to be sought in theories of operations and collections (if they are to be sought at all).

In correspondence concerning the preceding Professor MacLane has raised several criticisms which I shall try to summarize and respond to here. Though I had tried to keep the argument as simple as possible, it now seems to me that elaboration on these points will help clarify certain issues; I am thus indebted to Professor MacLane for his timely rejoinder.

First, he says that the program of categorical foundations has made considerable progress (since the papers [13], [15], and [14]) via work on elementary theories of topoi by Lawvere, Tierney, Mitchell, Cole, Osius, and others. He believes that this makes the discussion above out of date and beside the point.

Second, MacLane thinks that questions of psychological priority are 'exceedingly fuzzy' and subjective. Further, mathematicians are well known to have very different intuitions, and these may be strongly affected by training.

It is necessary to indicate the nature of the work on topoi before going into these points. A good introductory survey is to be found in MacLane [21]. Elementary topoi are special kinds of categories $\mathbb{C}$ which have features strongly suggested by the category of sets. (At the same time there are various other examples of mathematical interest such as in categories of sheaves.) In place of membership, one deals with morphisms $f: 1 \to X$ where 1 is a terminal object of $\mathbb{C}$. The requirements to be a topos include closure under such constructions on objects X, Y of $\mathbb{C}$ as product $X \times Y$ and exponentiation X^Y. The principal new feature is that $\mathbb{C}$ is required to have a 'subobject classifier' Ω whereby every subobject $S \rightarrowtail X$ of an object X corresponds uniquely to a 'member' of Ω^X. Thus Ω generalizes the role of the set of truth-values $\{0, 1\}$ and Ω^X generalizes the role of the power set of X for sets X. By means of additional axioms on topoi one may reflect more and more of the particularities of the category of sets. Indeed, (following Mitchell and Cole) Osius [22] gives two extensions ETS(Z) and ETS(ZF) of the elementary theory of topoi which are equivalent (by translation), respectively, to the theories of sets Z and ZF. (Incidentally, this work clarifies the relationship of Lawvere's theory in [13] to ZF.)

According to MacLane, the technical development just described shows that set-theoretical foundations and categorical foundations are entirely equivalent and hence that one cannot assign any logical priority to the former.

In response:

(i) My use of 'logical priority' refers not to relative strength of formal theories but to order of definition of concepts, in the cases where certain of these *must* be defined before others. For example, the concept of vector space is logically prior to that of linear transformation; closer to home, the (or rather some) notions of set and function are logically prior to the concept of cardinal equivalence. (By contrast, there are cases where there is no priority, e.g., as between Boolean algebras and Boolean rings.)

(ii) On the other hand, 'psychological priority' has to do with natural order of understanding. This is admittedly 'fuzzy' but not always 'exceedingly' so. Thus one cannot understand abstract mathematics unless one has understood the use of the logical particles 'and', 'implies', 'for all', etc. and understood the conception of the positive integers. Moreover, in these cases formal systems do not serve to explain what is not already understood since these concepts are implicitly involved in understanding the workings of the systems themselves. This is not to deny that formal systems as well as informal discussions can serve to clarify meanings when there is ambiguity, e.g. as between classical and constructive usage.

(iii) My claim above is that the general concepts of operation and collection have logical priority with respect to structural notions (such as 'group', 'category', etc.) because the latter are defined in terms of the former but not conversely. At the same time, I believe our experience demonstrates their psychological priority. I realize that workers in category theory are so at home in their subject that they find it more natural to think in categorical rather than set-theoretical terms, but I would liken this to not needing to hear, once one has learned to compose music.

(iv) The preceding has to do with an order between concepts, according to which some of them appear to be more basic than others. There is in consequence an order between theories of these concepts. Namely, we choose certain systems *first* because they reflect our understanding (as well as one can formulate it) of basic conceptions; other systems may be chosen *later* because they are useful and reducible to the former. For example, axioms about real numbers reflect some sort of basic understanding, while axioms for non-standard real numbers are only justified by a relative consistency proof. This is in opposition to the (formalist) idea that equivalence of formal systems makes one just as good a foundation for a certain part of mathematics as another; it is contrary to our actual experience in the development and choice of such systems.

(v) My neglect of the work on topoi was not due to ignorance but rather based on the conclusion that it was irrelevant to my argument. Indeed, since topoi are just special kinds of categories the objections here to a program for the categorical foundations of mathematics apply all the more to foundations via theories of topoi. It should be added that the axioms of ETS(Z) and ETS(ZF) were clearly obtained by tracing out just what was needed to secure the translations of Z and ZF in the language of topoi. This applies particularly to the replacement scheme and bears out my contention of the priority of set-theoretical concepts.

(vi) To avoid misunderstanding, let me repeat that I am *not* arguing for accepting current set-theoretical foundations of mathematics. Rather, it is that on the platonist view of mathematics something like present systems of set theory must be prior to any categorical foundations. More generally, on any view of abstract mathematics priority must lie with notions of operation and collection.

II

1. *Defects of Present Foundations for Category Theory*

Current set-theoretical foundations do not permit us to meet the following two requirements:

(R1) *form the category of all structures of a given kind, e.g. the category* $\mathbb{G}$ *of all groups, the category* $\mathbb{T}$ *of all topological spaces, the category* $\mathbb{C}$ *of all categories, etc.;*

(R2) *form the category* $\mathbb{B}^{\mathbb{A}}$ *of all functors from* $\mathbb{A}$ *to* $\mathbb{B}$ *when* $\mathbb{A}, \mathbb{B}$ *are any given categories.*

This is the main reason that MacLane [15] argues that set-theoretical foundations are inadequate. As described in [15] there are two means at present to reformulate (R1), (R2), and other constructions of category theory in set-theoretically acceptable terms. Briefly, these are:

(i) The Grothendieck method of 'universes'. A *universe U* is a set of sets satisfying strong closure conditions, including closure under exponentiation and under Cartesian product more generally. (The sets of rank $<\alpha$ form a universe when α is inaccessible.) It is assumed that for every universe U there is another U' which contains U as member. With each universe and notion of a particular kind of structure is associated the category of all such structures in U; this is a member of U' when $U \in U'$. For example, we can speak of the category $\mathbb{G}_U$ of all groups in U; $\mathbb{G}_U$ is then an element of U' when $U \in U'$. Thus (R1) is satisfied only in a relative way. For any U and categories $\mathbb{A}, \mathbb{B} \in U$ the category $\mathbb{B}^{\mathbb{A}}$ also belongs to U. The method of universes provides a reduction of category theory to ZF+'there exist infinitely many inaccessibles'.[6]

(ii) The Eilenberg-MacLane reduction to the BG theory of sets and classes, via the distinctions between 'small' and 'large'. A category is said to be *small* if it is a set, otherwise *large*. For example, the category $\mathbb{G}$ of all sets which are groups is large. The functor category $\mathbb{B}^{\mathbb{A}}$ exists only under the hypothesis that $\mathbb{A}$ is small, so (R2) is not satisfied. Also (R1) may be said to be satisfied only partially since, e.g. there is no category of all large groups.[7]

In addition to the inadequacies just explained there is dissatisfaction with the two schemes in that the restrictions employed seem mathematically unnatural and irrelevant. Though bordering on the territory of the paradoxes, it is felt that the notions and constructions involved in (R1), (R2) have evolved naturally from ordinary mathematics and do not have the contrived look of the paradoxes. Thus it may be hoped to find a way which gives them a more direct account.

It must be said that there is no urgent or compelling reason to pursue foundations of unrestricted category theory, since the schemes (i), (ii) (or their variants and refinements) serve to secure all practical purposes. Speaking in logical terms, one can be sure that any statement of set theory proved using general category theory has a set-theoretical proof (to be more precise, a proof in ZF if we follow scheme (ii), since BG is conservative over ZF). The aim in seeking new foundations is mainly as a problem of logical interest motivated largely by aesthetic considerations (or rather by the inaesthetic character of the present solutions).

2. *The Search for Foundations of Unrestricted Category Theory*

The situation here is analogous to several classical problems of foundations in mathematics, when one employed objects conceived beyond ordinary experience, such as infinitesimals, imaginary numbers, and points at infinity. The restrictions by set-theoretical distinctions of size in category theory are analogous to the employment of the "ε, δ" language for the foundations of the calculus. On the other hand, the consistency proofs for the complex number system and for projective geometry justified a direct treatment of the latter ideas.[8] We have in these cases *extensions* of familiar systems, namely of the real number system and Euclidean geometry, resp. But it is to be noted that in each case a price was paid: certain properties of the familiar objects were sacrificed in the process. When passing from the reals to the complex numbers one must give up the ordering properties; when passing from Euclidean to the projective space, metric properties must be given up.

In the case of category theory the idea would be to formulate a system in which one could obtain familiar collections and carry out familiar operations, such as the following:

(R3) *form the set N of natural numbers, form ordered pairs* (a, b), *form* $A \cup B$, $A \cap B$, $A - B$, $A \times B$, B^A, $\mathcal{P}(A)$, $\bigcup_{x \in A} B_x$, $\bigcap_{x \in A} B_x$, $\prod_{x \in A} B_x$, *etc.*

Since new 'unlimited' collections would have to be objects of the theory and we would have *self-application* by (R1) we could not expect to carry over all familiar laws. The problem then would be to select an appropriate *part of* (R3) and such familar laws as *extensionality*, to see which should be extended. Here there is no clear criterion for selection, but one wants to preserve ordinary mathematics as much as possible.

In the past few years I have experimented with several formal systems to achieve this purpose. One system based on an extension of Quine's stratification was reported in [4] and a write-up of the work was informally distributed. However, the paper has not been sent for publication for two reasons: (i) the system introduced turned out to be very similar to one that had been discovered independently by Oberschelp [19], and (ii) it was not completely successful for the intended purposes. In particular, though one could carry out (R1), (R2), and (R3) to a certain extent (e.g. to form $A \cup B$, $A \cap B$, $A - B$, $A \times B$, B^A) one could not carry out $\prod_{x \in A} B_x$. However, Cartesian product is a very important operation used to build structures, to verify completeness of concrete categories, etc., and so ought not to be given up.

Last year I found another theory which does somewhat better and also has a more intrinsic plausibility. This was suggested by my earlier work [5] on constructive theories of operations and classifications (collections). The point in the preceding work had been to deal with those notions conceived of as given *intensionally* by presentations of rules respectively defining properties, i.e. in both cases as certain kinds of syntactic expressions. In a universe which contains objects which 'code' syntactic expressions, we can have self-application. Of course, this may be accidental, depending on the coding. The problem is to arrange instead mathematically significant instances of self-application. Where [5] had used *partial operations* and *total classifications*, the new theory [6] achieved its aims by also using *partial classifications*. Intuitively, suppose c is a classification, i.e. an object given by a property φ_c. In testing whether or not $\varphi_c(x)$ holds we are in some cases led into a circle. Write $(x\eta c)$ if it can be *verified* (in some sense) that φ_c holds of x and $(x\bar{\eta}c)$ if it can be *verified* that φ_c does *not* hold of x. Here we conceive of verification as a possibly transfinite process, e.g. $\forall y \varphi(y)$ is verified if for each b, $\varphi(b)$ is verified. Not every attempt to verify leads to a conclusion, e.g. $\varphi_c(c)$ cannot be verified if $\varphi_c(x)$ is $x\eta x$ or $(x\bar{\eta}x)$. Write $\Box\varphi$ if φ is verified. Thus $x\bar{\eta}c \leftrightarrow \Box\neg(x\eta c)$. Pursuing these ideas leads one to a comprehension scheme of the following form:

$$(*) \qquad \exists c\, \forall x [x\eta c \leftrightarrow \Box\varphi(x)] \wedge [x\bar{\eta}c \leftrightarrow \Box\neg\varphi(x)],$$

where no restrictions are placed on φ in (*). Such a scheme was first shown to be consistent by Fitch [7, 8] and a closely related scheme in ordinary predicate calculus was shown consistent by Gilmore [9]. However, to arrange (R1), (R2) and as much as possible of (R3), it appears that somewhat more must be built in. Namely, there must be a theory of operations in terms of which one defines such classifications as b^a; further, these operations should be extensive enough to take both operations and classifications as arguments or values. In the next section I shall describe a theory resulting from this combination of ideas.

3. *A Non-extensional Theory of Partial Operations and Classifications*

3.1. Let T be any theory in any logic including the classical first order predicate calculus, and which contains a symbol 0 and operation symbols ′ and (,) and which proves:

A1. $x' \neq 0$.

A2. $x' = y' \rightarrow x = y$.

A3. $(x_1, y_1) = (x_2, y_2) \rightarrow x_1 = x_2 \wedge y_1 = y_2$.

We write $\mathscr{L}$ for the language of T.

Our first step is to extend this to a theory of partial operations. We adjoin a three-place predicate symbol App (x, y, z) which is also written $xy \simeq z$; App is then denoted $\simeq$. The intuitive interpretation is that the operation x is defined at y with value z. We write $\mathscr{L}(\simeq)$ for the language extended by the symbol $\simeq$. Note that expressions compounded by xy are not terms of $\mathscr{L}(\simeq)$. Extend the language by a binary operation symbol for application; by a *pseudo-term* t we mean a term of the extended language. The meaning of $t \simeq z$ for any pseudo-term t is defined inductively as a formula of $\mathscr{L}(\simeq)$. Then $(t\downarrow)$ is written for $\exists z(t \simeq z)$ and $t_1 \simeq t_2$ for $\forall z[t_1 \simeq z \leftrightarrow t_2 \simeq z]$. Finally, $xy_1 \cdots y_n$ is written for $(\ldots(xy_1)\ldots)y_n$.

The theory $T_{\simeq}$ contains the following axioms:

A4. (*Unicity*). $xy \simeq z_1 \wedge xy \simeq z_2 \rightarrow z_1 = z_2$.

A5. (*Explicit definition*). For each pseudo-term t with variables $x, y_1, \ldots, y_n$

$$\exists f \forall y_1, \ldots, y_n, x[fy_1 \cdots y_n\downarrow \wedge fy_1 \cdots y_n x \simeq t].$$

In addition, $T_{\simeq}$ contains axioms guaranteeing the existence of p_1, p_2, d, and e satisfying:

A6. (*Projections*) $p_1(x, y) \simeq x \wedge p_2(x, y) \simeq y$.

A7. (*Definition by cases*)

$$(x = y \rightarrow dxyab \simeq a) \wedge (x \neq y \rightarrow dxyab \simeq b).$$

A8. (*Quantification*)

$$[ex \simeq 0 \leftrightarrow \exists y(xy \simeq o)] \wedge [ex \simeq 1 \leftrightarrow \forall y \exists z(xy \simeq z \wedge z \neq 0)].$$

If we fix any f associated with t by (5), we write $\lambda x \cdot t$ or $\lambda x \cdot t[x]$ for $fy_1 \ldots y_n$. The idea of (5) is that $(\lambda x \cdot t)$ exists because it names a rule, whether or not $t[x]$ is defined. Thus we have

$$(\lambda x \cdot t[x])\downarrow \wedge (\lambda x \cdot t[x])x \simeq t.$$

As special cases of A5 we get existence of the *identity* i satisfying: $ix \simeq x$; the *combinators* k, s satisfying: $kxy \simeq x$, $sxy\downarrow \wedge sxyz \simeq xz(yz)$; the *successor operation* s_1 satisfying: $s_1 x \simeq x'$; and the *pairing operator* p satisfying $pxy \simeq (x, y)$. A single scheme of explicit definition by certain quantified formulas may also be given which implies A5–A8 and in such a way as to insure the following.

THEOREM 1. $T_{\simeq}$ *is a conservative extension of* T.

A proof of this for the given axioms A5–A8 may be obtained by associating with any model $\mathfrak{M} = (M, \ldots)$ of T a generalization of recursion theory, essentially *prime computability in* $\exists^{(M)}$ defined by Moschovakis [17]. The interpretation of $xy \simeq z$ is $\{x\}(y) \simeq z$ in this generalization.

We can already define partial classifications in T, as those induced by partial characteristic functions: say, $xpc \leftrightarrow cx \simeq 0$ and $x\bar{p}c \leftrightarrow cx \simeq 1$. Let $cx \simeq 0$ for all x; then $\forall x(xpc)$, i.e. c is a universal classification. Write $x: u \rightarrow w$ for $\forall y[ypu \rightarrow (xy\downarrow) \wedge (xy)pw]$. It is easily proved that

$$\rightarrow \exists d \forall x[xpd \leftrightarrow x: c \rightarrow c],$$

i.e. there is no classification which consists exactly of the total operations. For this reason we now pass to a richer theory of classifications.

3.2. The second step is to expand the language $\mathscr{L}(\simeq)$ by a pair of new binary relation symbols η and $\bar{\eta}$; we further adjoin a new unary propositional operator $\Box$; $x\eta c$, $x\bar{\eta}c$ and $\Box\varphi$ may be read informally as suggested at the end of Section 2. The extended language is denoted $\mathscr{L}(\simeq, \eta, \bar{\eta})$.

The logic is that of $\mathscr{L}$ augmented by the modal system S4+BF with rules and axioms as follows (cf. [10] for more details):

B1 from φ infer $\Box\varphi$.

B2. $\quad \Box\varphi \to \varphi$.

B3. $\quad \Box\varphi \to \Box\Box\varphi$.

B4. $\quad \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$.

B5. $\quad \forall x \Box\varphi(x) \to \Box\forall x\varphi(x)$.

We have also the following axioms for atomic formulas special to our situation.

B6. $\quad \varphi \to \Box\varphi$ for each atomic formula.

B7. $\quad \neg\varphi \to \Box\neg\varphi$ for each atomic formula of $\mathscr{L}(=)$.

B8. $\quad x\bar{\eta}c \leftrightarrow \Box\neg(x\eta c)$.

We write $S \vdash \varphi$ if for some $\psi_1, \ldots, \psi_n \in S$ we can derive $\psi_1 \wedge \cdots \wedge \psi_n \to \varphi$ in this logic. Note that if $\psi \in S$ we cannot in general apply B1 to get $S \vdash \Box\psi$. Using B6, B7, we may derive from these axioms:

$$\varphi \leftrightarrow \Box\varphi \text{ for each formula } \varphi \text{ of } \mathscr{L}(=).$$

Now the axioms of the main theory $\hat{\mathrm{T}}$ are those of $\mathrm{T}_=$ together with the following:

D. (*Disjointness*) $\neg(x\eta a \wedge x\bar{\eta}a)$.
C. (*Comprehension*) For each formula φ of $\mathscr{L}(=, \eta, \bar{\eta})$ with free variables among x, a_i $(i \leqslant n)$,

$$\exists f \forall a_1, \ldots, a_n \exists c[fa_1 \ldots a_n = c \wedge \forall x(x\eta c \leftrightarrow \Box\varphi) \wedge \forall x(x\bar{\eta}c \leftrightarrow \Box\neg\varphi].$$

THEOREM 2. $\hat{\mathrm{T}}$ *is a conservative extension of* $\mathrm{T}_=$, *hence of* T.

The idea of the proof is to consider any model $\mathfrak{M} = (M, \ldots, =)$ of $\mathrm{T}_=$ and to apply Kripke semantics to the collection $\mathscr{K}(\mathfrak{M})$ consisting

of all $(\mathfrak{M}, R, \bar{R})$ where R, $\bar{R}$ are disjoint binary relations on $\mathfrak{M}$, ordered by $(\mathfrak{M}, R, \bar{R}) \leqslant (\mathfrak{M}, S, \bar{S}) \Leftrightarrow R \subseteq \bar{R}$ and $S \subseteq \bar{S}$. In particular, $(\mathfrak{M}, R, \bar{R}) \models \Box\varphi(\mathbf{a}) \Leftrightarrow$ for all $(\mathfrak{M}, S, \bar{S}) \geqslant (\mathfrak{M}, R, \bar{R})$ we have $(\mathfrak{M}, S, \bar{S}) \models \varphi(\mathbf{a})$. Next use 0, ′, and (,) to set up a coding in M of arbitrary formulas so that $\ulcorner\varphi\urcorner \in M$ represents φ. Then put $f_\varphi y_1 \cdots y_n = (\ulcorner\varphi\urcorner, y_1, \ldots, y_n)$. Define η_α, $\bar{\eta}_\alpha$ by transfinite recursion $\eta_0 = \bar{\eta}_0 =$ empty. We take $x\eta_{\alpha+1}c \Leftrightarrow$ for some $\varphi(x, \mathbf{y})$, $c = (\ulcorner\varphi\urcorner, \mathbf{y})$ and $(\mathfrak{M}, \eta_\alpha, \bar{\eta}_\alpha) \models \Box\varphi(x, \mathbf{y})$; $x\bar{\eta}_{\alpha+1}c \Leftrightarrow$ for some $\varphi(x, \mathbf{y})$, $c = (\ulcorner\varphi\urcorner, \mathbf{y})$ and $(\mathfrak{M}, \eta_\alpha, \bar{\eta}_\alpha) \models \Box\neg\varphi(x, \mathbf{y})$. Let $\eta_\lambda = \bigcup_{\alpha<\lambda} \eta_\alpha$, $\bar{\eta}_\alpha = \bigcup_{\alpha<\lambda} \bar{\eta}_\alpha$ for λ a limit number. Finally take $\eta = \eta_\alpha$ and $\bar{\eta} = \bar{\eta}_\alpha$ where $\eta_\alpha = \eta_{\alpha+1}$, $\bar{\eta}_\alpha = \bar{\eta}_{\alpha+1}$. It may be shown that $(\mathfrak{M}, \eta, \bar{\eta})$ is a model of $\hat{\mathrm{T}}$ (in the extended logic applied to $\mathscr{K}(\mathfrak{M})$).

It should be noted that extensionality can actually be disproved for total functions in $\mathrm{T}_=$ and for total classifications in $\hat{\mathrm{T}}$; this goes by a diagonal argument (cf. [6] 3.6 and 4.7).

3.3. The special case of T = ZF in the language of $\in$, $=$ (with 0, ′, (,) defined as usual) is useful to consider in comparison with present set-theoretical foundations of category theory. Here it is natural to strengthen $\hat{\mathrm{T}}$ by the scheme

$$S \ (\textit{Separation}) \quad \exists b \forall x[x \in b \leftrightarrow x \in a \wedge \varphi],$$

for each formula φ of $\mathscr{L}(=, \eta, \bar{\eta})$. Let $\mathrm{T}^\# = \hat{\mathrm{T}} + S$.

THEOREM 3. $\mathrm{T}^\#$ *is a conservative extension of* $\mathrm{T}_=$, *hence of* T.

The proof uses standard models for any fragment of $\mathrm{T}^\#$ and the reflection principle for ZF.

3.4. We now draw some basic consequences in these theories. Call φ *persistent* if $(\varphi \to \Box\varphi)$ is provable, and *invariant* if both φ and $\neg\varphi$ are persistent. It is easily shown that if both η and $\bar{\eta}$ have only positive occurrences in φ then φ is persistent. With each φ in the language of 1st order predicate calculus is associated a pair of formulas φ^+ and φ^- which are both positive w.r. to η and $\bar{\eta}$ and which are approximations in a certain sense to φ and $\neg\varphi$, resp. For example, when φ is in prenex disjunctive normal form, to obtain φ^+ we replace each negated η by $\bar{\eta}$ and each negated $\bar{\eta}$ by η. Then $\vdash(\varphi^+ \to \varphi)$ and $\vdash(\varphi^- \to \neg\varphi)$. Under suitable circumstances we actually have $\varphi^+ \leftrightarrow \varphi$ and $\varphi^- \leftrightarrow \neg\varphi$. For example, suppose $\varphi(x, a_1, a_2)$ is $x\eta a_1 \wedge \neg(x\eta a_2)$. Then φ^+ is $x\eta a_1 \wedge x\bar{\eta} a_2$, and φ^- is $x\bar{\eta} a_1 \vee x\eta a_2$. Now if a_1, a_2 are total, i.e. $x\bar{\eta} a_i \leftrightarrow \neg(x\eta a_i)$ we have $\varphi^+ \leftrightarrow \varphi$ and $\varphi^- \leftrightarrow \neg\varphi$. In these cases $(\varphi \leftrightarrow \Box\varphi)$ and $(\neg\varphi \leftrightarrow \Box \neg \varphi)$.

For each $\varphi(x, a_1, ..., a_n)$, take any f satisfying the comprehension scheme C of $\hat{\mathrm{T}}$. We write $\hat{x}\varphi(x, a_1, ..., a_n)$ or $\hat{x}\varphi(a, \mathbf{a})$ for $fa_1 \cdots a_n$ and $\lambda\mathbf{a} \cdot \hat{x}\varphi(x, \mathbf{a})$ for f. Hence

(1) $[x\eta\hat{x}\varphi(x, \mathbf{a}) \leftrightarrow \Box\varphi(x, \mathbf{a})]$ and
$[x\bar{\eta}\hat{x}\varphi(x, \mathbf{a}) \leftrightarrow \Box \neg \varphi(x, \mathbf{a})]$

are provable in $\hat{\mathrm{T}}$. Note we can write $[x\eta\hat{x}\varphi(x, \mathbf{a}) \leftrightarrow \varphi(x, \mathbf{a})]$ only when $\mathbf{a}$ is such that $\varphi(x, \mathbf{a})$ is persistent. One must be cautious to observe whether this is the case.

To begin with, define

(2) $CL = \hat{x}\forall z[z\eta x \vee z\bar{\eta} x]$.

Then $x\eta CL$ iff x is a *total classification*. CL is itself not total (by Russell's argument). Next we proceed to define

(3) (i) $V = \hat{x}(x = x)$, $\Lambda = \hat{x}(x \neq x)$.
(ii) $\{a, b\} = \hat{x}(x = a \vee x = b)$.
(iii) $a \cup b = \hat{x}(x\eta a \vee x\eta b)$.
(iv) $a \cap b = \hat{x}(x\eta a \wedge x\eta b)$.
(v) $-a = \hat{x} \neg (x\eta a)$.
(vi) $a \times b = \hat{x}\exists x_1, x_2(x = (x_1, x_2) \wedge x_1\eta a \wedge x_2\eta b)$.
(vii) $b^a = \hat{x}(x: a \to b) = \hat{x}\forall u[u\eta a \to \exists v(xu \simeq v \wedge v\eta b)]$.

It may be seen, e.g. that $x\eta(a \cap b) \leftrightarrow x\eta a \wedge x\eta b$ because of the positivity of the defining condition; further if $a\eta CL$, $b\eta CL$ then $(a \cap b)\eta CL$. On the other hand, $x\eta(-a) \leftrightarrow x\bar{\eta}a$ by (1).

In the case of the function classification b^a, it is seen that $x\eta b^a \leftrightarrow (x: a \to b)$ whenever a is total; moreover, $a\eta CL \wedge b\eta CL \to (b^a\eta CL)$.

Note that for each of these operations we actually have an element that represents it, e.g. $f = \lambda a, b(a \cap b)$ gives $fab = a \cap b$.

We may also define extended operations of union, intersection, sum and product as follows:

(4) (i) $\bigcup a = \hat{x}\exists z(x\eta z \wedge z\eta a)$.
(ii) $\bigcap a = \hat{x}\forall z(z\bar{\eta} a \vee x\eta z)$.
(iii) $\sum_{z\eta a} (gz) = \hat{x}\exists z, w, v[x = (z, w) \wedge z\eta a \wedge gz \simeq v \wedge w\eta v]$.
(iv) $\prod_{z\eta a} (gz) = \hat{x}\forall z[z\bar{\eta} a \vee \exists w, v(xz \simeq w \wedge gz \simeq v \wedge w\eta v)]$.

It is clear that $\bigcup$ and $\sum$ behave reasonably in general, that $a\eta CL$ and $a \subseteq CL \to (\bigcup a)\eta CL$ and that $a\eta CL$ and $h: a \to CL$ implies $\left(\sum_{z\eta a} (hz)\right)\eta CL$. For $\bigcap$ we get the appropriate defining condition when $a\eta CL$, and we have $(\bigcap a)\eta CL$ when also $\mathbf{a} \subseteq CL$. For $\prod_{z\eta a} (hz)$ we get the appropriate defining condition when $a\eta CL$ and this product is total when $h: a \to CL$.

It is possible to define power classifications by

(5) $\mathcal{P}(a) = \hat{x}[x\eta CL \wedge \forall u(u\bar{\eta} x \vee u\eta a)]$.

Thus for $x\eta CL$, $x\eta\mathcal{P}(a) \leftrightarrow x \subseteq a$. But $\mathcal{P}(a)$ is not in general total even if a is total, e.g. not for $\mathcal{P}(V)$.

If e is an equivalence relation, its equivalence classes may be defined by $[z] = \hat{x}[(x, z)\eta]$. However, since we lack extensionality we cannot conclude that $(x, z)\eta e \leftrightarrow [x] = [z]$, only that $(x, z)\eta e \leftrightarrow [x] \equiv [z]$ where $u \equiv v \leftrightarrow \forall w(w\eta u \leftrightarrow w\eta v)$. But this shows at any rate that all equivalence relations can be reduced to the single one of $\equiv$. Though extensionality is generally thought to be essential for mathematics it is dispensable if we return to an older manner of speaking. When dealing with structures $\mathfrak{A} = (a, \ldots)$ we usually want to consider also some 'equality' relation defined on $\mathbf{a}$, i.e. a congruence relation for the structure $\mathfrak{A}$. This is standard in constructive mathematics (cf. Bishop [1]) and is only slightly complicating; cf. also Section 4 below.

3.5. We have satisfied much of (R3) in $\hat{\mathrm{T}}$. However, there remains the question of introducing the natural numbers as a total classification. In other words, we want the existence of an object N satisfying the scheme

(NN) (i) $N\eta CL$.
(ii) $0\eta N \wedge \forall x(x\eta N \rightarrow x'\eta N)$.
(iii) $\varphi(0)\wedge\forall x[\varphi(x) \rightarrow \varphi(x')] \rightarrow \forall x(x\eta N \rightarrow \varphi(x))$
for all formulas φ of $\mathscr{L}(=, \eta, \bar{\eta})$.

It is not difficult to modify the proof of Theorem 2 to obtain the following:

THEOREM 4. $\hat{\mathrm{T}}$+NN *is conservative over* $\mathrm{T}_{=}$, *hence over* T.

An alternative is to apply Theorem 2 to T regarded as expressed in a logic stronger than 1st order predicate calculus, by adjoining a 'quantifier' which determines inductive generation in general. For this see [6] Section 4c.

4. *Structure of structures and unrestricted category theory in* $\hat{\mathrm{T}}$

It is here convenient to use letters of any style as variables ranging over the objects of the theory $\hat{\mathrm{T}}$, including Latin, Greek and German letters, as well as the letters $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$, … . We proceed informally but in such a way that all the work can be formalized in $\hat{\mathrm{T}}$. For each $k\geq 1$ define a^k by $a^1=a$, $a^{k+1}=a\times a^k$. Then for each $\tau=((k_1, \ldots, k_n), (l_1, \ldots, l_m), p)$ we define $\mathfrak{A}$ to be a *structure of type* τ if it is of the form:

(1) $\mathfrak{A}=(a, r_1, \ldots, r_n, f_1, \ldots, f_m, 0_1, \ldots, 0_p)$
where $r_i\subseteq a^{k_i}$ $(1\leq i\leq n)$, $f_i\colon a^{l_i}\rightarrow a$ and $0_i\eta a$.

Here **a** and each r_i are in general only partial classifications. We write $\mathrm{Str}_\tau(\mathfrak{A})$ if (1) holds and $!\mathrm{Str}_\tau(\mathfrak{A})$ if $\mathfrak{A}$ is a *total structure*, i.e. if

(2) $a\eta CL$ and $r_i\eta CL$ $(1\leq i\leq n)$.

We can form

(3) $!S_\tau=\hat{\mathfrak{A}}\ !\mathrm{Str}_\tau(\mathfrak{A})$,

so that $!S_\tau$ *is the partial classification of all total structures of type* τ.

As pointed out in 3.4, since we lack extensionality, in practice each structure $\mathfrak{A}$ must carry along a relation e which acts as an *equality relation for* $\mathfrak{A}$. Thus, for example, a semi-group (associative binary system) is a structure $\mathfrak{A}=(a, e, f)$ where $e\subseteq a^2$, $f\colon a^2\rightarrow a$, $(f(x, f(y, z)), f(f(x, y)z))\eta e$ and e is an equivalence relation which preserves f. We have a formula Sem Grp $(\mathfrak{A})$ which expresses that $\mathfrak{A}$ is a semi-group and !Sem Grp $(\mathfrak{A})$ which expresses that $\mathfrak{A}$ is a total semi-group. There is then a classification $!\mathrm{SG}=\hat{\mathfrak{A}}$!Sem Grp $(\mathfrak{A})$ *of all total semi-groups*. Write $x\equiv_{\mathfrak{A}}y$ for $(x, y)\eta e$. Given two semi-groups $\mathfrak{A}=(a, \equiv_{\mathfrak{A}}, f)$ and $\mathfrak{B}=(b, \equiv_{\mathfrak{B}}, g)$, a *homomorphism* $\alpha\colon\mathfrak{A}\rightarrow\mathfrak{B}$ is a triple $(h, \mathfrak{A}, \mathfrak{B})$ where h is an operation $h\colon a\rightarrow b$ such that $x\equiv_{\mathfrak{A}}y\rightarrow hx\equiv_B hy$. We shall write αx for hx. An isomorphism is a pair of homomorphisms $\alpha\colon\mathfrak{A}\rightarrow\mathfrak{B}$, $\beta\colon\mathfrak{B}\rightarrow\mathfrak{A}$ such that $x\equiv_{\mathfrak{A}}\beta\alpha x$ for $x\eta a$ and $z\equiv\alpha\beta z$ for $z\eta b$. We write $\mathfrak{A}\cong\mathfrak{B}$ if there exists such an isomorphism. The relation $E=(\hat{\mathfrak{A}}, \mathfrak{B})\cdot[\mathfrak{A}\cong\mathfrak{B}]$ is an equivalence relation on SG.

Next, given semi-groups $\mathfrak{A}$, $\mathfrak{B}$, define $\mathfrak{A}\times\mathfrak{B}=(a\times b, \equiv_{\mathfrak{A}\times\mathfrak{B}}, f\times g)$ as usual. We thus have an operation q such that for any $\mathfrak{A}$, $\mathfrak{B}$, $q\mathfrak{A}\mathfrak{B}=\mathfrak{A}\times\mathfrak{B}$. By the natural isomorphism $\mathfrak{A}\times(\mathfrak{B}\times\mathfrak{C})\cong(\mathfrak{A}\times\mathfrak{B})\times\mathfrak{C}$ we conclude that q is an associative operation up to $\cong$ on semi-groups. Hence we can prove,

(4) Sem Grp $(\mathfrak{S})$ where $\mathfrak{S}=(!\mathrm{SG}, \cong, q)$.

In other words the structure of all total semi-groups with the operation of products forms 'the semi-group of all total semi-groups'.

This is a paradigm for the treatment of other algebraic notions in $\hat{\mathrm{T}}$ together with an illustration of the possibilities of self-application. Proceeding similarly we can define the notions: Grp $(\mathfrak{A})$ ($\mathfrak{A}$ is a group), !Grp $(\mathfrak{A})$ ($\mathfrak{A}$ is a total group), $!G=\hat{\mathfrak{A}}$!Grp $(\mathfrak{A})$ (the partial classification of all total groups), $\alpha\colon\mathfrak{A}\rightarrow\mathfrak{B}$ (α is a homomorphism from the group $\mathfrak{A}$ to the group $\mathfrak{B}$), Hom $(\mathfrak{A}, \mathfrak{B})=\hat{\alpha}(\alpha\colon\mathfrak{A}\rightarrow\mathfrak{B})$ for $\mathfrak{A}$, $\mathfrak{B}\eta !G$, $\mathrm{Hom}_{!G}=\hat{\alpha}[\exists\mathfrak{A}, \mathfrak{B}(\mathfrak{A}\eta !G\wedge\mathfrak{B}\eta !G\wedge\alpha\colon\mathfrak{A}\rightarrow\mathfrak{B})]$ and $\alpha\circ\beta\equiv\gamma$ (γ is a composition of α, β). Similarly we can define: Cat $(\mathfrak{A})$ ($\mathfrak{A}$ is a category), !Cat $(\mathfrak{A})$ ($\mathfrak{A}$ is a total category, $!C = \hat{\mathfrak{A}}$!Cat $(\mathfrak{A})$ (the partial classification of all total categories), $\varphi: \mathfrak{A} \to \mathfrak{B}$ (φ is a functor from $\mathfrak{A}$ to $\mathfrak{B}$), $\mathfrak{B}^{\mathfrak{A}} =$ Funct $(\mathfrak{A}, \mathfrak{B}) = \hat{\varphi}(\varphi: \mathfrak{A} \to \mathfrak{B})$ for $\mathfrak{A}$, $\mathfrak{B}$ total categories, $\text{Funct}_{!C} = \hat{\varphi}[\exists \mathfrak{A}, \mathfrak{B}(\mathfrak{A}\eta!C \wedge \mathfrak{B}\eta!C \wedge \varphi: \mathfrak{A} \to \mathfrak{B})]$ and $\varphi \circ \psi \equiv \vartheta$ (ϑ is a composition of φ, ψ). From $!G$, $\text{Hom}_{!G}$ we can assemble a category $\mathbb{G}$ which is the *category of all total groups.* $\mathbb{G}$ itself is not a total structure, but we have

(5) Cat $(\mathbb{G})$.

Similarly, we can assemble from $!C$, $\text{Funct}_{!C}$ a category $\mathbb{C}$ which is the *category of all total categories*, so that

(6) Cat $(\mathbb{C})$.

In this way we achieve a form of the requirement (R1) in $\hat{\text{T}}$.

Now for (R2), given any *total categories* $\mathbb{A}$, $\mathbb{B}$, we can form as above Funct $(\mathbb{A}, \mathbb{B})$ – which are the objects of $\mathbb{B}^{\mathbb{A}}$ considered as a category; its morphisms are the natural transformations between functors. We obtain

(7) $!\text{Cat}(\mathbb{A}) \wedge !\text{Cat}(\mathbb{B}) \to !\text{Cat}(\mathbb{B}^{\mathbb{A}})$

as a consequence of the closure of CL under exponentiation. In this way requirement (R2) is met. More generally we can form $\mathbb{B}^{\mathbb{A}}$ with reasonable properties when it is merely assumed that $\mathbb{A}$ is total:

(8) $!\text{Cat}(\mathbb{A}) \wedge \text{Cat}(\mathbb{B}) \to \text{Cat}(\mathbb{B}^{\mathbb{A}})$.

A basic structure to consider is $\mathbb{CL}$, the *category of all total classifications.* This has the property that for any objects A, B of $\mathbb{CL}$, Hom $(A, B) = B^A$ is also total. More generally, call a category $\mathbb{A}$ *locally total* if for any objects $\mathfrak{A}$, $\mathfrak{B}$ of $\mathbb{A}$, $\text{Hom}_{\mathbb{A}}(\mathfrak{A}, \mathfrak{B})$ is total. Then $\mathbb{CL}$, $\mathbb{G}$, $\mathbb{C}$, etc. are all locally total. With any locally total $\mathbb{A}$ and object $\mathfrak{A}$ of $\mathbb{A}$ is associated the Yoneda functor

(9) $h^{\mathfrak{A}}: \mathbb{A} \to \mathbb{CL}$,

which is defined on objects by $h^{\mathfrak{A}}(\mathfrak{B}) = \text{Hom}_{\mathbb{A}}(\mathfrak{A}, \mathfrak{B})$ and is defined on morphisms in an obvious way. Now *Yoneda's Lemma* (YL) may be expressed in $\hat{\text{T}}$ as follows: *If $\mathbb{A}$ is locally total and φ is any functor from $\mathbb{A}$ to $\mathbb{CL}$ then for each object $\mathfrak{A}$ of $\mathbb{A}$, φ is in $1-1$ correspondence with the class of natural transformations from $h^{\mathfrak{A}}$ to φ.*

Call a category *complete* if it is closed under equalizers and products $\prod_{i\eta c} a_i$ where $c\eta CL$. It may be shown that for any total category $\mathbb{A}$, the category $\mathbb{CL}^{\mathbb{A}}$ is complete. Further the Yoneda map is an embedding of $\mathbb{A}$ into $\mathbb{CL}^{\mathbb{A}}$. Hence every total category can be embedded in a complete (non-total) category.

Suppose we are working over set theory, i.e. T = ZF or an extension of ZF. In this case we work in $\text{T}^{\#}$ instead of $\hat{\text{T}}$. The objects of principal interest are structures on sets a; these have associated classifications $\bar{a} = \hat{x}(x \in a)$. Denote by Set the classification consisting of all $\bar{a}$; we also call a classification *small* if it belongs to Set. Note that Set is total: (Set)ηCL. We can form a category from Set which we denote $\mathbb{CL}_s$, *the category of all small classifications;* then further we can form $\mathbb{G}_s$, *the category of all small groups*, $\mathbb{C}_s$, *the category of all small categories*, etc. Each of these is a total category and is locally small in the usual sense. In this respect one considers *small-completeness* of a category, i.e. closure under $\prod_{i\eta c} a_i$ for c small. A question of interest would be whether we can associate with any locally small category a small-completion.

These examples relate to several mentioned by MacLane [15] Section 7 as being problematic for the present set-theoretic accounts of category theory. It is evident that some of the hoped-for freedom is gained by passing to theories like $\hat{\text{T}}$ or $\text{T}^{\#}$. It is true that statements of results must now make distinctions between being *partial* and *total* which previously were made between being *large* and *small.* However, this is no disadvantage if one takes the objects of primary interest to be the categories of small objects such as $\mathbb{CL}_s$, $\mathbb{G}_s$, $\mathbb{C}_s$, etc., all of which are total.

To conclude I wish to emphasize that the kind of expansion of accepted language pursued here is theoretically useful *as a matter of convenience.* That is, the theorems of Section 3 are conservative extension results, which permit us to extend an already accepted theory T to $\hat{T}$ (or $T^{\#}$) without getting new theorems in the language of T. This parallels the classical cases such as introduction of complex numbers. However, there is not yet evidence that expansions such as $\hat{T}$ have any significant mathematical advantage (comparable, say, to the use of complex numbers to obtain results about the real numbers). A further pursuit of the mathematics involved is needed in addition to the logical and aesthetic considerations to guide us to a fully satisfactory foundation of unrestricted structure theory.

Stanford University

NOTES

* Text of a talk for a Symposium on Category Theory in the 5th International Congress of Logic, Methodology and Philosophy of Science (London, Ontario, Aug. 27–Sept. 2, 1975). Research supported by NSF grant MPS74-07505 A01.

[1] This is detailed in [6] and in a succeeding paper which is in preparation.

[2] There are some problems about Lawvere's theory which have been raised by Isbell [11] and others. However, these do not affect its guiding point of view nor the viability of some such theory.

[3] One should mention particularly those raised by Kreisel in his appendix to [3] and in his review [12].

[4] In addition to its appearance in a certain way in the discussions of ftn. 4, I have been told by P. Martin-Löf that he has also raised similar objections.

[5] This is disputed by Myhill [18], which attempts to give an extensional constructive set theory.

[6] Some refinements of this idea should also be mentioned. In [14] MacLane showed that one universe is enough. In [3] I made use of the reflection principle for ZF to show that weaker closure conditions on U suffice for most work.

[7] It is indicated in [12] how to modify Bernays' relation η for the theory of sets and classes (read: properties) so that, e.g., $\mathbb{G}$ can also be construed as the category of all large groups. However, this modification does not satisfy (R2). It is also suggested by Kreisel in [12] that one look at 'suitably indexed collections of functors (which can be defined in BG)' rather than 'the' functor category. One development of this idea of getting around (R2) was carried out by the author in some unpublished seminar notes 'Set-theoretical formulation of some notions and theorems of category theory' (Stanford, October 1968). While it is viable and in one way more fitting to a thorough-going algebraic-axiomatic approach to mathematics, in another way it goes against the grain *never* to be able to talk about 'the' category of all functors between any two large categories.

[8] Robinson's non-standard analysis may be viewed as an attempt to justify direct use of infinitesimals in the calculus.

BIBLIOGRAPHY

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

[2] A. Chauvin: 1974, *Theorie des objets et théorie des ensembles*, Thèse, Université de Clermont-Ferrand.

[3] S. Feferman: 1967, 'Set-Theoretical Foundations of Category Theory' (with an Appendix by G. Kreisel), in *Reports of the Midwest Category Seminar III*, Lecture Notes in Mathematics 106, Springer, 201–247.

[4] S. Feferman: 1974, 'Some Formal Systems for the Unlimited Theory of Structures and Categories' (abstract), *J. Symbolic Logic* **39,** 374–375.

[5] S. Feferman: 1975, 'A Language and Axioms for Explicit Mathematics', in *Algebra and Logic*, Lecture Notes in Mathematics 450, Springer, 87–139.

[6] S. Feferman: 1975, 'Non-Extensional Type-Free Theories of Partial Operations and Classifications', I, in *Proof Theory Symposion, Kiel* 1974, Lecture Notes in Mathematics 500, Springer, 73–118.

[7] F. B. Fitch: 1963, 'The System CΔ of Combinatory Logic', *J. Symbolic Logic* **28,** 87–97.

[8] F. B. Fitch: 1966, 'A Consistent Modal Set Theory' (abstract), *J. Symbolic Logic* **31,** 701.

[9] P. C. Gilmore: 1974, 'The Consistency of Partial Set Theory Without Extensionality', *Axiomatic set theory* (UCLA 1967), *Proc. Symposia in Pure Math. XIII*, Part II, A.M.S., Providence (1974) 147–153.

[10] G. E. Hughes and M. J. Cresswell: 1968, *An Introduction to Modal Logic*, Methuen, London.

[11] J. Isbell: 1967, Review of Lawvere [13], *Mathematical Reviews* **34,** 7332.

[12] G. Kreisel: 1972, Review of MacLane [15], *Mathematical Reviews* **44,** 25.

[13] F. W. Lawvere: 1966, 'The Category of Categories as a Foundation of Mathematics', *Proc. Conference Categorical Algebra* (LaJolla 1965), Springer, N.Y. 1–20.

[14] S. MacLane: 1969, 'One Universe as a Foundation for Category Theory', in *Reports of the Midwest Category Seminar* III, Lecture Notes in Mathematics 106, Springer, 192–200.

[15] S. MacLane: 1971, 'Categorical Algebra and Set-Theoretical Foundations', in *Axiomatic Set Theory* (UCLA 1967) *Proc. Symposia in Pure Math. XIII*, Part I, A.M.S., Providence, 231–240.

[16] P. Martin-Löf: 1974, 'An Intuitionist Theory of Types: Predicative Part', preprint No. 3, University of Stockholm.

[17] Y. Moschovakis: 1969, 'Abstract First-Order Computability II', *Trans. Am. Math. Soc.* **138,** 465–504.

[18] J. Myhill: 1975, 'Constructive Set Theory', *J. Symbolic Logic* **40,** 347–382.

[19] A. Oberschelp: 1973, 'Set Theory Over Classes', *Dissertationes Mathematicae*, **106,** 1–66.

[20] D. Scott: 1970, 'Constructive Validity', in *Symposium on Automatic Demonstration*, Lecture Notes in Mathematics 125, Springer, 237–275.

[21] S. MacLane: 1975, 'Sets, Topoi and Internal Logic in Categories', in *Logic Colloquium '73*, North-Holland Publ. Co., Amsterdam, 119–133.

[22] G. Osius: 1974, 'Categorical Set Theory: A Characterization of the Category of Sets', *J. Pure Appl. Algebra* **4,** 79–119.