Contemporary Mathematics
Volume 30, 1984

THE INTERACTION BETWEEN CATEGORY THEORY AND SET THEORY

Andreas Blass

This paper, like the lecture on which it is based, is a survey of a few of the ways that category theory and set theory interact. The topics to be treated were not chosen in any particularly rational way; my own interests and knowledge were the dominant factor in the selection of material. I apologize to all whose work should have been mentioned but was not.

In the first section, we discuss the interaction arising as a special case of set theory's foundational role in mathematics. In the second section, we discuss interactions arising as a special case of category theory's role of clarifying and unifying concepts from a broad range of subfields of mathematics. The third section is devoted to the possibility of building set-theoretic structure into certain categories, particularly topoi, and using the interplay between set-theoretic and category-theoretic ideas to illuminate both. In the final section, we exhibit two situations where it is fairly clear that category theory is trying to tell us something about certain set-theoretical concepts and results, but the message is not yet well understood. The two appendices contain justifications for some of the mathematical assertions made earlier in the paper; non-mathematical assertions are left unjustified.

1. Set theory interacts with everything

It is a remarkable empirical fact[1] that mathematics can be based on set theory. More precisely, all mathematical objects can be coded as sets (in the cumulative hierarchy built by transfinitely iterating the power set operation, starting with the empty set), and all their crucial properties can be proved

0271-4132/84 $1.00 + $.25 per page

from the axioms of set theory. For example, the natural numbers can be coded as von Neumann ordinals, the integers and the rational numbers as equivalence classes of ordered pairs (which in turn can be coded as sets of the form $\{\{x\}, \{x,y\}\}$), real numbers as Dedekind cuts (or equivalence classes of Cauchy sequences), functions as sets of ordered pairs, etc. Then the basic properties of these systems, for example that the real numbers form a complete Archimedean ordered field, are theorems of set theory (by which we mean, for the sake of definiteness, ZFC, Zermelo-Fraenkel set theory with the axiom of choice).

At first sight, category theory seems to be an exception to this general phenomenon. It deals with objects, like the categories of sets, of groups, etc. that are as big as the whole universe of sets and that therefore do not admit any evident coding as sets. Furthermore, category theory involves constructions, like the functor category, that lead from these large categories to even larger ones. Thus, category theory is not just another field whose set-theoretic foundation can be left as an exercise. An interaction between category theory and set theory arises because there is a real question: What is the appropriate set-theoretic foundation for category theory? We shall consider three of the answers that have been given to this question.

Answer 1. None.

The point of this answer is that for its own internal development category theory, like most branches of mathematics, does not need a set-theoretic foundation. Once the basic concepts are clearly understood, their set-theoretic encoding is irrelevant. (For example, an analyst wants the real numbers to form a complete ordered field; he does not care whether they are Dedekind cuts or equivalence classes of Cauchy sequences or anything else.) But this approach is not adequate for answering questions like: Does category theory necessarily involve existential principles that go beyond those of other mathematical disciplines? At first sight, the answer to this question is yes, because of the need for large (and superlarge and . . .) categories; a more careful analysis amounts to an attempt to provide a set-theoretic foundation for category theory.

Answer 2. Grothendieck universes.

The idea here is to assume the existence of an inaccessible cardinal κ (an assumption that goes beyond ZFC) and to think of the "sets" that occur in ordinary mathematics as being only the sets created in the first κ stages of the cumulative hierarchy, the so-called small sets. Then there are further stages available at which large (and superlarge and . . .) categories can be created. This approach has two drawbacks. One is that, as already mentioned, it uses a hypothesis that goes beyond ZFC, so large categories still seem to require stronger existence principles than the rest of mathematics. The other is that, if one wants to use category theory to prove a theorem about ordinary sets, this approach will establish the theorem only for small sets. To get the result for all sets, we need to have not just one inaccessible κ but a proper class of them, so that every set is small for some choice of κ. The need to be able to change from one κ to another (change of universe) results in a lot of technicalities which, from a non-foundational point of view, look irrelevant. On the other hand, there are some positive things to be said for this approach. It made inaccessible cardinals popular in France. It provides a rigorous set-theoretic foundation for free manipulations of large categories. Its set-theoretic assumption that goes beyond ZFC, the existence of a proper class of inaccessible cardinals, is rather mild[2] at least in comparison with the large-cardinal hypotheses that set-theorists are fond of. And it probably helped to motivate the third answer.

Answer 3. Reflection principles.

Here again, one replaces the sets of ordinary mathematics with the small sets, those created in the first κ stages of the cumulative hierarchy. But one does not assume that κ is inaccessible. Instead, one assumes that each first-order statement (in the language of set theory, with small sets as parameters) has the same meaning when the variables are interpreted as ranging over all sets as when they range only over small sets. In other words, the universe of small sets is an elementary substructure of the universe of all sets. This approach, developed in [10], has two advantages. First, the assumptions guarantee that, if we prove a theorem about small sets by using large categories, then the same theorem holds for arbitrary sets; we never need to introduce more κ's to make more sets small. Second, the assumptions do not really go beyond ZFC; any assertion in the first-order language of set theory, not mentioning κ, that can be proved using these assumptions can also be proved without them.[3] Thus, the results obtained for small sets by considering large categories can also be obtained (though with more involved proofs) for arbitrary sets, on the basis of just the ZFC axioms. In this sense, category theory does not really depend on principles beyond those of ZFC, at least as far as theorems about sets are concerned.

Although this approach was first proposed in connection with the problem of foundations for category theory, it is natural to use it whenever objects seem to be too large to be coded as sets. In particular, it seems to me that it should be of some use in clarifying forcing with proper classes [9,20,29] by making the natural (regular open) Boolean algebra available even though it is superlarge.

2. Category theory interacts with nearly[4] everything

It is a remarkable empirical fact that the important structural properties of mathematical objects are often expressible in category-theoretic terms, specifically as universal properties. Among the concepts admitting universal descriptions are, for example, natural numbers, power set, cartesian product, free product, tensor product, universal enveloping algebra, Stone-Čech compactification, universal covering space, and the logical connectives and quantifiers.

Not only do universal descriptions exist, but they are useful in at least two ways. First, they tend to express the more important properties of mathematical structures, so that keeping them in mind helps one to avoid irrelevant complications. For example, there exist long computational proofs that the operation $- \otimes M$ of tensor product with a fixed module preserves the exactness of sequences $A \to B \to C \to 0$ of modules. These proofs are based on an explicit description of the elements of a tensor product. This description, though valuable for many purposes, is not the essential property of the tensor product. The essential (and simpler, once one gets accustomed to thinking categorically) property is a universal one, namely that $- \otimes M$ is left adjoint to $\mathrm{Hom}(M,-)$. From this point of view the right exactness of $- \otimes M$ is an immediate corollary of the general (and quite simple) fact that left adjoints preserve coequalizers (or, more generally, colimits). For another example, consider the statement that, if A is a proper subalgebra of a Boolean algebra B, then there are two distinct ultrafilters in B with the same intersection with A. This is not hard to prove directly, but it becomes trivial if one uses the proper formulation of the Stone representation theorem, the form that (in the spirit of category theory) pays attention to morphisms as well as objects and asserts that the Stone space construction is a contravariant equivalence from the category of Boolean algebras to the category of totally disconnected compact Hausdorff spaces. When such an equivalence is applied to a monomorphism that is not an isomorphism, like the inclusion of A into B, the result is an epimorphism that is not an isomorphism, and such a morphism of totally disconnected compact Hausdorff spaces is not one to one; so two distinct points of the Stone space of B have the same image in the Stone space of A, as desired.

The second use of category-theoretic descriptions is that they can reveal structural similarities between concepts in diverse areas of mathematics. For example, free groups, abelianizations of groups, universal enveloping algebras, and Stone-Čech compactifications are all instances of the same construction, the left adjoint of a foregetful functor. And indeed it is often useful to think of abelianization as "the abelian group freely generated by a group" and Stone-Čech compactification as "the compact Hausdorff space freely generated by a space[5]". Here are some examples involving set theory.

The first involves my thesis, "Orderings of Ultrafilters" [2], which should have been "Categories of Ultrafilters". At the time, I was aware of (and discussed in [2] and [3]) a category, with ultrafilters as objects, closely related to the Rudin-Keisler ordering of ultrafilters. But I only recently realized, though the idea is implicit in Glazer's work [18], that the Rudin-Frolík ordering is also closely related to a category, one whose morphisms are (sequences of) ultrafilters. It turns out that one of the most important properties of the Rudin-Frolík ordering, its treelike character established by Rudin [31], becomes, in a categorical formulation, a simple algebraic condition that is also satisfied (in a stronger form) by the category freely generated by any directed graph. For details, see Appendix A.

A second example involves Boolean-valued models of set theory. A well-known construction due to Scott and Solovay (see [20]) associates to every complete Boolean algebra B a B-valued model of set theory V^B and (as any category theorist would feel compelled to add) functorially associates to every complete homomorphism $B \to B'$ a transformation $V^B \to V^{B'}$. A crucial property of complete homomorphisms is that, if we view them as functors by viewing the Boolean algebras as partially ordered sets and thus as categories, then they have adjoints on both sides. In iterated forcing, the construction known as the inverse limit is really the inverse limit with respect to the left adjoints of the homomorphisms. The construction known as the direct limit is not really a limit of the Boolean algebras, but Šcedrov [32] has shown that the corresponding Boolean valued model is a limit in the sense of fibered topoi [19]; again we find a connection with an entirely different branch of mathematics, algebraic geometry.

Finally, let me mention that the dual Ramsey theorem recently proved by Carlson and Simpson [8] can, as the name implies, be motivated (though unfortunately not proved) by formulating Ramsey's theorem in (partially) category-theoretic terms and then reversing the arrows. As suggested in [8], the same procedure should yield other reasonable concepts and conjectures in combinatorial set theory.

3. Set-theoretic concepts in categories

Many of the elementary concepts of category theory were introduced for the purpose of expressing familiar concepts of set theory and their generalizations in other areas of mathematics. A partial list of such pairs of concepts is given in the following table:

Category theory	Set theory
Object	Set
Morphism	Function
Monomorphism	One-to-one function
Epimorphism	Surjection
Isomorphism	Bijection
Product	Cartesian product
Coproduct	Disjoint union

The full extent of this "categorical set theory" appears in topos theory, and it is there that set theory and category theory interact most strongly. As a prologue to the topos axioms, let us briefly consider the standard ways of describing things in set theory and in category theory.

A set is usually specified by saying what its elements are. The axiom of extensionality guarantees the uniqueness of sets so specified. Of the other axioms, most (null set, pairing, separation, union, power set, replacement) assert the existence of a set specified in this manner, and the axiom of infinity can be rephrased to fit the same pattern. The axiom of choice cannot be so rephrased; indeed the whole point of this axiom is to assert the existence of (a few of the) sets whose members we are unable to specify. It seems that it is precisely this unusual characteristic of the axiom of choice that led to the controversies about it in the early part of this century. The axiom of regularity can, with some effort, be rephrased as the existence of sets with specified members: Every set has an ordinal rank. But this rephrasing is quite distant from the usual formulation of the axiom and, partially as a result, the axiom of regularity has suffered an even worse fate than the axiom of choice; people don't deny the axiom of regularity, they just ignore it.

An object in a category is usually specified by giving a universal property, that is, by saying what the morphisms into or out of it are. Yoneda's lemma says that this information (in a suitably functorial form) suffices to uniquely determine an object, up to isomorphism.

If we view objects in a category as generalized sets (as in the table above), then morphisms into X, from an arbitrary I, are (generalized) I-indexed families of members of X. Thus, one of the two modes of category-theoretic description (using morphisms into, rather than out of, X) generalizes the set-theoretic mode by specifying families of elements rather than single elements. And it turns out to be this mode that expresses the essential existence properties of the category $\mathcal{S}$ of sets. These properties, the topos axioms, assert the existence of objects X with the "incoming morphisms" specified in the following table; we have included in the last line a version (due to Freyd [15]) of the axiom of infinity, which is not among the topos axioms but is often assumed in conjunction with them. (Like the usual set-theoretic form, it does not uniquely determine X because X occurs in the right column.)

Object X	Morphisms $I \to X$
1	Exactly one morphism
Pullback of $\begin{array}{ccc} & & Y \\ & & \downarrow \\ W & \to & Z \end{array}$	Commutative squares $\begin{array}{ccc} I & \to & Y \\ \downarrow & & \downarrow \\ W & \to & Z \end{array}$
Set Ω of truth values	Subobjects of I
Set Z^Y of maps from Y to Z	Morphisms $I \times Y \to Z$
Dedekind-infinite set	Partitions of I as $J_0 \cup J_1$ together with maps $J_1 \to X$.

A remarkable theorem of Juul-Mikkelsen [22] asserts that these axioms (without infinity) imply the existence of objects, like coproducts, whose category-theoretic description involves "outgoing morphisms". Indeed, these axioms suffice for doing set theory in a topos as long as set theory is based on intuitionistic logic [1, 6, 12, 26, 28]. Specifically, one can interpret intuitionistic type theory in any topos, with the objects of the topos as types and with morphisms $1 \to \Omega$ as truth values. If the topos is sufficiently complete, then one can interpret in it an intuitionistic version IZF of ZF; see [13]. These results say that topoi are very like the universe of sets. As Fourman and Scott [12, 33] have emphasized, a topos is essentially the same thing as a theory formulated in a many-sorted intuitionistic logic with higher types.

On the other hand, topoi can be very unlike the universe of sets. The category $\mathcal{S}$ of sets is the most trivial of topoi, in the same sense that a single point is the most trivial of topological spaces. We give a few examples to explain this remark by indicating the rich variety of topoi.

1. The category Sh(X) of sheaves over any topological space X is a topos. This reduces to $\mathcal{S}$ if X is a single point. In general, there is much structure in a topos Sh(X) that is not detected by the internal logic, the intuitionistic type theory discussed above. For instance, if we let X be a circle and let A and B be its two non-isomorphic double coverings (two circles, and one circle winding around twice), then the sentence asserting that A and B <u>are</u> isomorphic is true (i.e. has the largest truth value) in the internal logic, because A and B are locally isomorphic. The difference between A and B implies only that no isomorphism is <u>definable</u> in the internal logic. Quite generally, internal existence corresponds to local existence while internal definability corresponds to global existence. Thus, global phenomena are reflected, not in the internal logic itself, but in its metatheory. In this sense, the internal logic is rather weak for geometric purposes. In another sense, however, it is too strong, for it contains operations (implication and universal quantification) that are not preserved when sheaves are pulled back along a continuous map of spaces. To be fair, one should add that the internal logic lacks some operations (infinite disjunctions) that are preserved. This discrepancy has led to extensive studies [21, 25, 32, 38] of the restricted part of (infinitary) logic that enjoys this preservation property, the so-called geometric logic, and the topoi, called classifying topoi, naturally associated to theories in this logic.

2. As a special case of the preceding example, taking X to be discrete, we have the topos $\mathcal{S}/X$ of X-indexed families of sets. More generally, for any object X of any topos $\mathcal{E}$, we have the topos $\mathcal{E}/X$, whose objects are morphisms into X and whose morphisms are commutative triangles in $\mathcal{E}$.

3. If P is any partially ordered set (poset), there is a topos $\mathcal{S}^P$ whose objects are P-indexed families of sets $(A_p)_{p \in P}$ equipped with functions $\alpha_{qp}: A_p \to A_q$ for $p \leq q$ such that $\alpha_{rq}\alpha_{qp} = \alpha_{rp}$ for $p \leq q \leq r$ and α_{pp} = identity. The morphisms are families of maps $A_p \to B_p$ compatible with the maps within the objects. These topoi are Kripke models of intuitionistic set theory, but we shall see them again in two other roles.

4. If G is any group (or monoid) the category $\mathcal{S}^G$ of sets with a G-action (and G-equivariant maps) is a topos. As another instance of the category-theoretic unity of mathematics, we mention that the notion of geometric morphism of topoi, obtained by abstracting the properties of pulling back sheaves along a continuous function, yields, when applied to topoi $\mathcal{S}^G$ with G a group, restriction of operators along a homomorphism of groups. In particular, one and the same topos-theoretic concept describes the natural morphisms of both topological spaces and groups. Better yet, one can mix the two; a geometric morphism from $Sh(X)$ to $\mathcal{S}^G$ is a one-dimensional cohomology class of X with coefficients in G.

5. Generalizing the last three examples, we have the topos of presheaves, i.e., contravariant set-valued functors, on an arbitrary small category. And generalizing all the preceding examples, we have topoi of sheaves over sites [19], called Grothendieck topoi.

6. Finally, to avoid giving the impression that all topoi are Grothendieck topoi, we mention that the sets and functions of any model (even a nonstandard model) of set theory form a topos.

The dissimilarities between various topoi, indicated by the examples, can be exploited in combination with the similarities in their internal logical structure to produce two sorts of applications.

The first sort of application directly uses the fact that whatever can be defined or (intuitionistically) proved in set theory is available in all topoi. Rousseau [30] has shown that certain theorems in the function theory of one complex variable can be proved intuitionistically and then interpreted in the topos of sheaves over $\mathbb{C}^n$ to yield theorems about $n+1$ variables. In other topoi, one can model the continuum by a ring R (not the usual Dedekind or Cauchy reals) in which the set $D = \{x \in R \mid x^2 = 0\}$ is just large enough so that every function $D \to R$ is $x \mapsto ax + b$ for unique a and b in R. Then derivatives can be defined algebraically, and one can develop a theory of "synthetic differential geometry"; see [23] and the references given there. In the development of this theory, one need not (and usually does not) pay any attention to the topos in which one is working; one argues as in ordinary set theory, except that the arguments must be (formally) intuitionistic, and one tacitly relies on the validity of intuitionistic set theory in the topos.

The second application is to use topoi as models for establishing consistency and independence results in set theory. We cite three examples.

1. Cohen-style independence proofs, using forcing (but not symmetry arguments), can be rewritten in terms of the topoi $\mathcal{S}^P$, where P is a poset (= a notion of forcing), and their double-negation subtopoi (= Gödel negative translation of classical into intuitionistic theories). This was done by Lawvere and Tierney for the independence of the continuum hypothesis [24, 37] and by Bunge [7] for the independence of Souslin's hypothesis; Šcedrov [32] has a general study of forcing from a topos-theoretic point of view.

2. The independence proof for the axiom of choice, using forcing and symmetric submodels can also be cast in topos-theoretic terms [13], but Freyd [16] has given a direct topos-theoretic proof of the independence of the axiom of choice from (classical) ZF, using the double-negation subtopos of $\mathcal{S}^{A^{op}}$, where A^{op} is the dual of the following easily described category A. The objects are the finite von Neumann ordinals, and the morphisms $m \to n$ are the functions from m onto n that leave all members of n fixed. Šcedrov and I [5] and independently Solovay [35] have shown that this model is equivalent

to the following Cohen-style model, which seems not to have been previously considered by set theorists. Adjoin, to some general model M of ZFC, countably many generic sets a_n of natural numbers, let B be the Boolean algebra (of sets of natural numbers) generated by the a_n's, and take the smallest model having $M \cup \{a_n | \text{all } n\} \cup \{B\}$ as a subset.

3. Fourman and Šcedrov [14] used the topos of presheaves on a category that looks like

$$\begin{tikzcd} \bullet \arrow[r, "\alpha"] & \bullet \arrow[loop right, "\beta"] \end{tikzcd}$$

with β^2 = identity to show that one cannot prove in IZF (plus the axiom of dependent choice) the "world's simplest axiom of choice": If a family of two-element sets has at most one member (i.e., every two members are equal), then it has a choice function.

4. Two coincidences

In this section, I shall present two situations where the category-theoretic point of view shows us similarities between things that at first sight seem unrelated. In both cases, these similarities are not yet well understood; they have merely been observed.

The first concerns implications between axioms of choice from finite sets, a subject that begins with Tarski's proof [27] that, if all families of two-element sets have choice functions, then so do all families of four-element sets. (Of course this proof is in ZF, without the axiom of choice.) Implications of this sort were studied by Mostowski [27] and Gauntt [17] (see also [39]) who gave group-theoretical necessary and sufficient conditions for the implications to be provable. For simplicity, we consider first a slightly different sort of implication, for which the necessity of a slightly different group-theoretical condition follows from Gauntt's work while the sufficiency is proved in Appendix B using an extension of Mostowski's technique.

For any set I and any natural number k, let $C_k(I)$ be the statement that every I-indexed family of k-element sets has a choice function. Then the following are equivalent, for any natural numbers $k_1,\dots,k_r$, n.

(1) It is provable in ZF that, whenever a set I satisfies $C_{k_1}(I), C_{k_2}(I), \dots,$ and $C_{k_r}(I)$, then it also satisfies $C_n(I)$.

(2) Any group that can act without fixed points on an n-element set can also act without fixed points on a set of one of the cardinalities $k_1,\dots,k_r$.

Category theory, more specifically topos theory, enters the picture if we introduce, for a natural number k and a topos $\mathcal{E}$, the statement $C_k(\mathcal{E})$: "Every object of $\mathcal{E}$ internally isomorphic to k has a global element". In other words, if X is an object of $\mathcal{E}$ such that the assertion "X has exactly k pairwise distinct elements" is true in the internal logic then there is a morphism $1 \to X$ in $\mathcal{E}$. With this notation, it is not hard to check that a set I has the property $C_k(I)$ in a model $\mathfrak{M}$ of ZF if and only if $C_k(\mathfrak{M}/I)$ holds. Thus, (1) is equivalent to the statement that the implication

(3) If $C_{k_1}(\mathcal{E})$ and . . . and $C_{k_r}(\mathcal{E})$ then $C_n(\mathcal{E})$

holds for all topoi $\mathcal{E}$ of the form $\mathfrak{M}/I$ with $\mathfrak{M}$ a model of ZF and $I \in \mathfrak{M}$. On the other hand, a group G can act without fixed points on a k-element set if and only if it is not the case that $C_k(\mathcal{S}^G)$. (All k-element G-sets are internally isomorphic in $\mathcal{S}^G$ to k, and global elements are fixed points.) Therefore (2) is equivalent to the statement that the implication (3) holds for all topoi $\mathcal{E}$ of the form $\mathcal{S}^G$ with G a group.

The equivalence of (1) and (2) thus amounts to a transfer principle for implications of the form (3) between two classes for topoi, those of the form $\mathfrak{M}/I$ and those of the form $\mathcal{S}^G$. It is reasonable to expect this insight to shed new light on such equivalence results, perhaps leading to a "soft" proof, but it has not yet done so.

For the sake of completeness, we add a few words about the original result of Mostowski and Gauntt. They proved the equivalence of

(1') It is provable in ZF that, if all sets I satisfy $C_{k_1}(I), C_{k_2}(I), \ldots,$ and $C_{k_r}(I)$, then they all satisfy $C_n(I)$ as well. [So the truth of $C_n(I)$ can depend on the truth of $C_{k_i}(J)$ for J different from I.]

and

(2') Any group that can act without fixed points on an n-element set has a subgroup that can act without fixed points on a set of one of the cardinalities $k_1, \ldots, k_r$.

This equivalence can be expressed as the transfer, between the same two classes of topoi as before, of the implication

(3') If, for all objects X of $\mathcal{E}$, $C_{k_1}(\mathcal{E}/X)$ and . . . and $C_{k_r}(\mathcal{E}/X)$, then $C_n(\mathcal{E})$.

The second curiosity concerns Scott's theory of domains for denotational semantics (see [34] and the references there) in the form that involves semilattices P of finite pieces of data. Two of the central notions of Scott's theory, the ideal elements of the domain P and the approximable functions from one domain P to another Q, are precisely equivalent to the notions of geometric morphism $\mathcal{S} \to \mathcal{S}^P$ and $\mathcal{S}^P \to \mathcal{S}^Q$ respectively. Scott's construction of the domain of approximable functions gives us a new domain $[P \to Q]$ with the property that $\mathcal{S}^{[P \to Q]}$ is the exponential $(\mathcal{S}^P)^{\mathcal{S}^Q}$ in the category of Grothendieck topoi. What is striking here is that the same topoi $\mathcal{S}^P$ occur as Kripke models for intuitionistic set theory, as ingredients in the forcing technique, and now as a form of Scott domains. The first two of these have been connected by Fitting [11], but it is not yet clear how the third fits in.

As an application of this curiosity, Scott's construction of a nontrivial domain D isomorphic to $[D \to D]$ immediately yields an example of a nontrivial (i.e., different from $\mathcal{S}$) topos $\mathcal{D}$ equivalent to $\mathcal{D}^{\mathcal{D}}$. I know of no essentially different way to produce such an example.

Incidentally, the information systems used in [34] as an alternative basis for the theory of domains are essentially the same as Horn theories in propositional logic, and the corresponding topoi $\mathcal{S}^P$ are just the classifying topoi of these Horn theories.

Appendix A

This appendix is an explanation of the remark in Section 2 that the Rudin-Frolík ordering is related to a category whose morphisms are sequences of ultrafilters. We begin with the definition of this ordering. An ultrafilter U on a set I is Rudin-Frolík below an ultrafilter V on a set J, written $U \leq_{RF} V$, if and only if there is an I-indexed family of ultrafilters W_i on J such that

(i) there are sets $A_i \in W_i$ such that A_i is disjoint from $A_{i'}$ for all distinct i and i' in I, and

(ii) For all $B \subseteq J$, we have $B \in V$ if and only if $\{i \in I \mid B \in W_i\} \in U$.

(If we view ultrafilters on J as points of the Stone-Čech compactification of J, then (i) says that the family of W_i's is discrete and (ii) says that its limit with respect to U is V.) This ordering is usually studied in the special case where I and J are countable, and in this situation a theorem of Rudin [31] asserts that the Rudin-Frolík ordering is tree-like, that is, the set of predecessors of any ultrafilter is linearly ordered.

The following category $\mathcal{C}$ is intended to formalize the general idea behind the definition of the Rudin-Frolík ordering. The objects of $\mathcal{C}$ are simply sets; a morphism from I to J is an I-indexed family $(W_i)_{i \in I}$ where each W_i is an upward-closed family of subsets of J. (For our present purposes, we could have required the W_i's to be ultrafilters, but the more general situation is also of interest, for example when the sets are finite.) The composite of $(W_i') : I \to J$ and $(W_j'') : J \to K$ is the morphism $(W_i) : I \to K$ where, for $B \subseteq K$ and $i \in I$

$$B \in W_i \text{ if and only if } \{j \in J \mid B \in W_j''\} \in W_i' .$$

The identity morphism of I is the family of principal ultrafilters on I with the obvious I-indexing.

If we restrict attention to the subcategory of $\mathcal{C}$ whose morphisms are discrete (in the sense of (i) above) families of ultrafilters, then $U \leq_{RF} V$ if and only if V factors through U when both are viewed as morphisms with domain 1. Rudin's theorem then asserts that any commutative square in this subcategory, of the form

$$\begin{tikzcd} 1 \arrow[r, "U"] \arrow[d, "V"'] & I \arrow[d] \\ J \arrow[r] \arrow[ur, dash, dashed] & K \end{tikzcd}$$

with I,J,K countable, admits a "diagonal fill-in" either from I to J or from J to I making the upper triangle commute. (The same remains true with arbitrary, not necessarily countable sets in place of 1 and K.)

A similar but stronger fill-in principle, in which both triangles are required to commute, is an important property of the free categories on arbitrary directed graphs. Indeed, such categories are characterized by this principle plus the requirement that the relation "proper factor of" on the set of morphisms be well-founded.

Unfortunately, this stronger fill-in principle does not generally hold in the ultrafilter situation, but it nearly does. If the fill-in morphism goes from I to J, then the two I-indexed families, say (W_i) and $(W'_i) : I \to K$, in the bottom triangle agree modulo U; that is,

$$\{i \in I : W_i = W'_i\} \in U .$$

Returning to the larger category $\mathcal{C}$, we remark that it is dual to the category of power sets (= complete atomic Boolean algebras) and inclusion-preserving functions. The duality sends I to its power set $\mathcal{P}(I)$ and sends a morphism $(W_i)_{i \in I} : I \to J$ to the function

$$\mathcal{P}(J) \to \mathcal{P}(I) : A \mapsto \{i \in I \mid A \in W_i\} .$$

The subcategory whose morphisms are families of ultrafilters is dual, in the same way, to the category of power sets and Boolean homomorphisms.

Appendix B

For each natural number n and each finite set Z of natural numbers, consider the principle

$C^+(n,Z)$: For every set I, if every I-indexed family of k-element sets with $k \in Z$ has a choice function, then so does every I-indexed family of n-element sets.

We shall show that $C^+(n,Z)$ is provable in Zermelo-Fraenkel set theory if and only if n and Z have the following property, first considered by Gauntt [17].

$L(n,Z)$: Every group that can act without fixed points on an n-element set can also act without fixed points on a k-element set for at least one $k \in Z$.

Before we begin the proof, however, a few clarifying remarks are in order.

First observe that, if every I-indexed family of k-element sets has a choice function, then the same is true with I replaced by any of its subsets J, for we can always extend a J-indexed family to an I-indexed family and then restrict the choice function to J. This implies that the ambiguity, in the hypothesis of $C^+(n,Z)$, as to whether k can be different for different sets in the family, is of no consequence; indeed, if the choice function exists when k is constant, then in the case of variable k we can select choice functions for the subfamilies where k is constant and combine these into a choice function for the whole family. (The selecting is permissible because Z is finite so there are only finitely many subfamilies to consider.)

Second, to prevent confusion, it should be emphasized that $C^+(n,Z)$ is not in general equivalent to the similar-looking principle

$C(n,Z)$: If, for each $k \in Z$, every family of k-element sets has a choice function, then so does every family of n-element sets.

Indeed, it is an old theorem of Tarski [27] that $C(4,\{2\})$, but follows from work of Gauntt [17], as well as from the result that we shall prove[6], that there are models of ZF where $C^+(4,\{2\})$ fails, because the alternating group A_4 is a counterexample to $L(4,\{2\})$. Of course the implication from $C^+(n,Z)$ to $C(n,Z)$ is true, since the hypothesis of $C(n,Z)$ allows us to apply $C^+(n,Z)$ to every I, thereby producing the conclusion of $C(n,Z)$.

This last remark, applied to well-ordered I, shows that $C^+(n,Z)$ implies the principle

$C^*(n,Z)$: If, for each $k \in Z$, every well-ordered family of k-element sets has a choice function, then so does every well-ordered family of n-element sets.

The point of introducing C^* is that Gauntt [17] (see [39] for details) has shown that $C^*(n,Z)$ is provable in ZF if and only if $L(n,Z)$ holds. Therefore, the provability of $C^+(n,Z)$ implies the truth of $L(n,Z)$, and we have half of the desired equivalence.

To prove the other half, we consider arbitrary n,Z,I, and $(x_i)_{i \in I}$ such that $L(n,Z)$ is true, all I-indexed families of k-element sets with $k \in Z$ have choice functions, and (x_i) is an I-indexed family of n-element sets. Our goal of proving (in ZF) the existence of a choice function for (x_i) will be reached in a sequence of steps in which the sets x_i are endowed with more and more "structure", culminating in enough structure to determine a specific element. Although the final structure we want for each x_i is quite simple, a single chosen element, the structures along the way to this one will involve sets of sets of . . . subsets of x_i. To simplify the discussion, we shall treat the sets x_i as though they consisted of ur-elements, at least to the extent that, when we form the cumulative hierarchy of finite sets over x_i,

$$V_0(x_i) = x_i$$

$$V_{k+1}(x_i) = x_i \cup \mathcal{P}(V_k(x_i))$$

$$V(x_i) = \bigcup_k V_k(x_i) \ ,$$

the sets that "ought to be" newly added at each stage are in fact new, i.e., $\mathcal{P}(V_k(x_i)) - \mathcal{P}(V_{k-1}(x_i))$ is disjoint from $V_k(x_i)$. This amounts to saying that no member of x_i is a subset of $V_k(x_i)$, and it can be arranged without loss of generality, by suitably "tagging" the members of the x_i's; for a detailed discussion of a similar situation, see [27]. It will be convenient to fix once and for all a particular n-element set, which we call $\underset{\sim}{n}$, and whose members we treat as ur-elements in the same sense as above. We note for future reference that $V(\underset{\sim}{n})$ contains the natural numbers (coded as finite von Neumann ordinals) and is closed under formation of ordered pairs (coded as $\{\{a\},\{a,b\}\}$). We also note that the group $S_{\underset{\sim}{n}}$ of all permutations of $\underset{\sim}{n}$ acts on $V(\underset{\sim}{n})$ in a canonical manner: for $a \in V_{k+1}(\underset{\sim}{n}) - V_k(\underset{\sim}{n})$ and $\pi \in S_{\underset{\sim}{n}}$, $\pi(a)$ is defined (by induction on k) to be $\{\pi(b) \mid b \in a\}$. Thus, each $\pi \in S_{\underset{\sim}{n}}$ preserves the membership relation $\in$ and therefore everything definable from $\in$. In particular, π leaves each natural number fixed and commutes with formation of ordered pairs. For each $a \in V(\underset{\sim}{n})$, we denote its symmetry group $\{\pi \in S_{\underset{\sim}{n}} \mid \pi(a) = a\}$ by Sym (a).

<u>Lemma</u>. Let an action of some subgroup G of $S_{\underset{\sim}{n}}$ on some finite set a be given. Then there exists $a' \in V(\underset{\sim}{n}) - \underset{\sim}{n}$ such that $G \subseteq \mathrm{Sym}\,(a')$ and a', with the G-action obtained by restricting the canonical action of $S_{\underset{\sim}{n}}$ on $V(\underset{\sim}{n})$, is G-isomorphic to a.

<u>Proof</u>. The set a is a disjoint union of subsets (the G-orbits) on which G acts transitively, and it suffices to prove the lemma for these subsets, as we can then take the disjoint union of the resulting sets in $V(\underset{\sim}{n})$ to obtain the desired a'. (To make the sets in $V(\underset{\sim}{n})$ disjoint, we take advantage of the availability in $V(\underset{\sim}{n})$ of natural numbers and ordered pairs.) So we may assume that G acts transitively on a. Then a is (as a G-set)

a quotient of the regular action, i.e., of the set G with G acting by left multiplication. Since we can form quotients in $V(\underset{\sim}{n})$, it suffices to prove the lemma when a is the regular G-set G. But in this case a straightforward calculation shows that we can take a' to be a G-orbit in the set of bijections from the von Neumann ordinal n to the set $\underset{\sim}{n}$. □

In view of the lemma and the finiteness of Z, we can find a natural number k so large that, whenever the set a in the hypothesis of the lemma has cardinality in Z, the set a' in the conclusion can be taken to be in $V_k(\underset{\sim}{n})$. Fix such a k for the rest of the proof.

We are now in a position to describe the structures on the x_i that we shall construct. They are simply choices of specific elements from certain sets in $V_k(x_i)$. Thus, at each stage, we will have a partial choice function on every $V_k(x_i)$. To keep track of these structures, it will be convenient to have canonical examples of them, so we select one representative from each isomorphism class of structures of the form $(V_k(\underset{\sim}{n}),\in,f)$, where f is a partial choice function on $V_k(\underset{\sim}{n})$. This selection involves no use of the axiom of choice since there are only finitely many structures involved. (This is why we introduced k.)

As indicated above, our proof proceeds in stages. At the beginning of each stage, we shall have a family $(f_i)_{i \in I}$ of partial choice functions on the corresponding $V_k(x_i)$. The proof begins with all f_i empty, and it ends when, for each i, the set x_i is in the domain of f_i so the desired choice has been made. We now describe a typical stage in the proof. For each i, the structure $(V_k(x_i), \in, f_i)$ is isomorphic to exactly one of our representative structures $(V_k(\underset{\sim}{n}), \in, f)$. We partition I into finitely many pieces, putting two i's into the same piece if and only if the corresponding $(V_k(x_i), \in, f_i)$'s are isomorphic, and we focus attention on one such piece of I, say J, and the representative structure $(V_k(\underset{\sim}{n}), \in, f)$ associated to it. Let $G = \mathrm{Sym}(f)$; we consider two cases.

Case 1. G fixes some $p \in \underset{\sim}{n}$. Then we can extend all the f_i's, with $i \in J$, so as to have x_i in their domains as follows. For each $i \in J$, consider any isomorphism ϕ from $(V_k(\underset{\sim}{n}), \in, f)$ to $(V_k(x_i), \in, f_i)$. We assert that $\phi(p)$ is independent of the choice of ϕ. Indeed, if ψ is another such isomorphism then $\phi^{-1}\psi$ is an automorphism of $(V_k(\underset{\sim}{n}), \in, f)$, which implies $\phi^{-1}\psi \in G$. By the case hypothesis, $\phi^{-1}\psi(p) = p$ so $\psi(p) = \phi(p)$ as desired. So we can extend f_i by setting $f_i(x_i) = \phi(p)$.

Case 2. G fixes no $p \in \underset{\sim}{n}$. By $L(n,Z)$, G acts without fixed points on some set a with cardinality in Z. By the lemma and the choice of k, we may assume that $a \in V_k(\underset{\sim}{n}) - \underset{\sim}{n}$, $G \subseteq \mathrm{Sym}(\underset{\sim}{a})$, and the fixed-point-free action of G on a is the restriction of the canonical action of $S_{\underset{\sim}{n}}$ on $V_k(\underset{\sim}{n})$. As in Case 1, we find that, for each $i \in J$, all of the isomorphisms from $(V_k(\underset{\sim}{n}), \in, f)$ to $(V_k(x_i), \in, f_i)$ send a to the same element a_i of $V_k(x_i)$.

We assert that a_i is not in the domain of f_i, for any i. In view of the isomorphism, it suffices to verify that a is not in the domain of f. If $f(a)$ were defined, it would have to be fixed by G, since f and a are. But no element of a is fixed by G, and our assertion is therefore proved.

Since all the a_i $(i \in J)$ have cardinality in Z and since $J \subseteq I$, there is a choice function selecting an element p_i from each such a_i. We extend f_i by setting $f_i(a_i) = p_i$ for all $i \in J$.

We have described how to extend f_i for all i in one of the pieces J of I. Do this for all of the pieces. (There are only finitely many, so the axiom of choice is not used here.) This completes the description of one stage of our proof. Since the domain of each f_i gets strictly larger every time Case 2 occurs and contains x_i as soon as Case 1 occurs, we see that after at most as many steps as there are sets in $V_k(\underset{\sim}{n})$ the domain of f_i will contain x_i, and we have achieved our goal. □

Remark. It can be shown that each occurrence of Case 2 makes the group G in the proof strictly smaller, so the number of steps needed in the proof is bounded, independently of k (hence of Z), by the number of prime factors of $n!$, counted with multiplicity.

Footnotes

1. It is not clear to me whether this fact is a mathematical one, a historical one, or a psychological one (or something else). Does set theory have some essential structural property that guarantees its ability to encode other theories? Does set theory serve as a foundation for merely those theories that have been constructed in the past, with no expectation that it will serve for future theories? Or is there something about human brains that prevents them from producing mathematics that cannot be coded in set theory? My guess is that the historical view is closest to the truth, but for psychological reasons; mathematics codable into set theory was produced first (and we have not progressed beyond it) because it is easier for our minds to grasp. I also suspect that we have not yet come close to grasping the full complexity of what can be coded in set theory, so non-codable theories will probably not arise (naturally) for quite some time.

2. Many set theorists (including me) feel that the same intuitive conception of the cumulative hierarchy of sets that justifies the axioms of ZFC also justifies the assumption that there are arbitrarily large inaccessible cardinals.

3. The reason for this remarkable fact, proved in [10], is that any single proof can use only finitely many of the assumptions, i.e., the equivalence of "sets" with "small sets" in finitely many formulas, and, for each such finite subset of our assumptions, the existence of a κ satisfying them is provable in ZFC. Note that it is essential that the assumptions be stated as infinitely many statements, one about each formula. The argument

would not work if we had the single assumption "For any formula ϕ, ϕ is true when the variables range over all sets if and only if ϕ is true when the variables range over small sets". Fortunately, we were in no danger of making such an assumption because, by the well-known theorem of Tarski [36] on undefinability of truth, the concept of a set theoretical formula being "true when the variables range over all sets" is not (uniformly) definable in set theory. The closest one can come to expressing it in set theory is to express it for each ϕ individually (by saying ϕ), and this is what we did in formulating the assumptions.

4. The word "nearly" seems to be necessary here. I know of no interaction between category theory and, say, analytic number theory.

5. I believe that the categorical viewpoint also deserves the credit for the observation that complete regularity of a space X is not needed for the existence of the Stone-Čech compactification βX but only to guarantee that the canonical map $X \to \beta X$ (the unit of the adjunction) is an embedding.

6. An easier construction of a model of ZF with atoms, where every countable family of pairs has a choice function but not every countable family of four-element sets does, is indicated in the last section of [4].

REFERENCES

1. J. Bénabou, unpublished.
2. A. Blass, Orderings of Ultrafilters, thesis, Harvard University, 1970.
3. A. Blass, Two closed categories of filters, Fund. Math. 94 (1977), 129-143.
4. A. Blass, Cohomology detects failures of the axiom of choice, Trans. Amer. Math. Soc., 279 (1983), 257-269.
5. A. Blass and A. Ščedrov, Freyd's models for the independence of the axiom of choice, in preparation. See Abstracts Amer. Math. Soc. 3 (1982), 524.
6. A. Boileau and A. Joyal, La logique des topos, J. Symbolic Logic 46 (1981), 6-16.
7. M. Bunge, Topos theory and Souslin's hypothesis, J. Pure Appl. Alg. 4 (1974), 159-187.
8. T. Carlson and S. Simpson, A dual form of Ramsey's theorem, Adv. Math., to appear.

9. W. Easton, Powers of regular cardinals, Ann. Math. Logic I (1970), 139-178.

10. S. Feferman, Set-theoretical foundations of category theory, in Reports of the Midwest Category Seminar III, ed. S. MacLane, Springer Lecture Notes in Mathematics 106 (1969), 201-247.

11. M. Fitting, Intuitionistic Logic, Model Theory, and Forcing, North-Holland, 1969.

12. M. Fourman, The logic of topoi, in Handbook of Mathematical Logic, ed. J. Barwise, North-Holland, 1977, 1053-1090.

13. M. Fourman, Sheaf models of set theory, J. Pure Appl. Alg. 19 (1980), 91-101.

14. M. Fourman and A. Šc̆edrov, The "world's simplest axiom of choice" fails, Manuscripta Math. 38 (1982), 325 332.

15. P. Freyd, Aspects of topoi, Bull. Austral. Math. Soc. 7 (1972), 1-76 and 467-480.

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

17. R. J. Gauntt, Some restricted versions of the axiom of choice, Notices Amer. Math. Soc. 15 (1968), 351.

18. S. Glazer, Endomorphisms, infrafilters and segments of functions, preprint (1975).

19. A. Grothendieck and J. L. Verdier, Séminaire de Géométrie Algébrique IV, Tome 1, Springer Lecture Notes in Mathematics, 269 (1972), and Tome 2, Springer Lecture Notes in Mathematics, 270 (1972).

20. T. Jech, Set Theory, Academic Press, 1978.

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

22. C. Juul-Mikkelsen, Lattice theoretic and logical aspects of elementary topoi, Aarhus Univ., Mat. Inst. Various Publications Series 25 (1976).

23. A. Kock, Synthetic Differential Geometry, London Math. Soc. Lecture Notes 51 (1981).

24. F. W. Lawvere, Quantifiers and sheaves, Proc. Int. Cong. Math. 1970, 329-334.

25. M. Makkai and G. Reyes, First Order Categorical Logic, Springer Lecture Notes in Mathematics 611 (1977).

26. W. Mitchell, Boolean topoi and the theory of sets, J. Pure Appl. Alg. 2 (1972), 261-274.

27. A. Mostowski, Axiom of choice for finite sets, Fund. Math. 33 (1945), 137-168.

28. G. Osius, Logical and set-theoretical tools in elementary topoi, in Model Theory and Topoi, ed. F. W. Lawvere, C. Maurer, and G. G. Wraith, Springer Lecture Notes in Math. 445 (1975), 297-346.

29. D. Pelletier, Forcing with proper classes: a Boolean-valued approach, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astron. Phys. 22 (1979), 345-352.

30. C. Rousseau, Topos theory and complex analysis, J. Pure Appl. Alg. 10 (1977/78), 299-313.

31. M. E. Rudin, Partial orders on the types of βN, Trans. Amer. Math. Soc. 155 (1971), 353-362.

32. A. Šc̆edrov, Sheaves and forcing and their metamathematical applications, thesis, SUNY at Buffalo, 1981, to appear in Memoirs Amer. Math. Soc.

33. D. Scott, A report on intuitionistic algebra, retiring presidential address to Assoc. Symbolic Logic, Biloxi, January 1979.

34. D. Scott, Domains for denotational semantics, ICALP '82, Aarhus, July 1982.

35. R. Solovay, Boolean valued interpretations of topoi, I, unpublished manuscript.

36. A. Tarski, Der Wahrheitsbegriff in den formalisierten Sprachen, Studia Philoscophica 1 (1936), pp. 261-405.

37. M. Tierney, Sheaf theory and the continuum hypothesis, in Toposes, Algebraic Geometry and Logic, ed. F. W. Fawvere, Springer Lecture Notes in Mathematics 274 (1972), 13-42.

38. M. Tierney, Forcing topologies and classifying topoi, in Algebra, Topology and Category Theory, ed. A. Heller and M. Tierney, Academic Press (1976), 211-219.

39. J. Truss, Finite axioms of choice, Ann. Math. Logic 6 (1973), 147-176.

MATHEMATICS DEPARTMENT
THE UNIVERSITY OF MICHIGAN
ANN ARBOR, MI 48109