Observations on Category Theory

John L. Bell

1. Devised by Eilenberg and Mac Lane in the early 1940s, *category theory* is a presentation of mathematics in terms of the fundamental concepts of *transformation*, and *composition* of transformations. While the importance of these concepts had long been recognized in algebra (for example, by Galois through the idea of a group of permutations) and in geometry (for example, by Klein in his *Erlanger Programm*), the truly universal role they play in mathematics did not really begin to be appreciated until the rise of abstract algebra in the 1930s. In abstract algebra the idea of transformation of structure (homomorphism) was central from the beginning, and it soon became apparent to algebraists that its most important concepts and constructions were in fact formulable in terms of that idea alone. Thus emerged the view that the essence of a mathematical structure is to be sought not in its internal constitution, but rather in the nature of its relationships with other structures of the same kind, as manifested through the network of transformations. This idea has achieved its fullest expression in category theory, an axiomatic framework within which the notions of transformation (as *morphism* or *arrow*) and composition (and also structure, as *object*) are fundamental, that is, *are not defined in terms of anything else.*

From a philosophical standpoint, a category may be viewed as an explicit presentation of a *mathematical form or concept*. The objects of a category $\mathscr{C}$ are the *instances* of the associated form and the morphisms or arrows of $\mathscr{C}$ are the transformations between these instances which in some specified sense "preserve" this form. As examples we have:

Category	***Form***	***Transformations***

Sets	Pure discreteness	Functional correlations
Sets with relations	" " " "	One-many correlations
Groups	Composition/inversion	Homomorphisms
Topological spaces	Continuity	Continuous maps
Differentiable manifolds	Smoothness	Smooth maps

Because the practice of mathematics has, for the past century, been officially founded on set theory, the objects of a category—in particular those of all the above-mentioned categories—are normally constructed as *sets* of a certain kind, synthesized, as it were, from pure discreteness. As sets, these objects manifest set-theoretic relationships—memberships, inclusions, etc. However, these relationships are irrelevant—and in many cases are actually *undetectable*—when the objects are considered as embodiments of a form, i.e., viewed through the lens of category theory. (For example, in the category of groups the additive group of even integers is isomorphic to, i.e. indistinguishable from, the additive group of all integers.) This fact constitutes one of the "philosophical" reasons why certain category theorists have felt set theory to be an unsatisfactory basis on which to build category theory—and mathematics generally. For categorists, set theory provides a kind of ladder leading from pure discreteness to the category-theoretic depiction of the real mathematical landscape. Categorists are no different from artists in finding the landscape (or its depiction, at least) more interesting than the ladder, which should, following Wittgenstein's advice, be jettisoned after ascent.

2. Interpreting a mathematical concept within a category amounts to a kind of refraction or filtering of the concept through the form associated with the category. For example, the interpretation of the concept *group* within the category of topological spaces is *topological group*, within the category of manifolds it is *Lie group*, and within a category of sheaves it is *sheaf of groups*.

This possibility of varying the category of interpretation leads to what I have called in [1] *local mathematics*, in which mathematical concepts are held to possess references, not within a fixed absolute universe of sets, but only *relative* to categories of interpretation of a certain kind—the so-called *elementary toposes*. Absolute truth of mathematical assertions comes then to be replaced by the concept of *invariance*, that is, "local" truth in every category of interpretation, which turns out to be equivalent to constructive provability.

3. In category theory, the concept of *transformation* (morphism or arrow) is an irreducible basic datum. This fact makes it possible to regard arrows in categories as formal embodiments of the idea of *pure variation* or *correlation*, that is, of the idea of *variable quantity* in its original pre-set-theoretic sense. For example, in category theory the variable symbol x with domain of variation X is interpreted as an *identity arrow* (1_X), and this concept is not further analyzable, as, for instance, in set theory, where it is reduced to a set of ordered pairs. Thus the variable x now suggests the idea of pure variation over a domain, just as intended within the usual functional notation $f(x)$. This latter fact is expressed in category theory by the "trivial" axiomatic condition

$$f \circ 1_X = f,$$

in which the symbol x does not appear: this shows that variation is, in a sense, an intrinsic constituent of a category.

4. In certain categories—Lawvere-Tierney's *elementary toposes*—the notion of pure variation is combined with the fundamental principles of construction employed in ordinary mathematics through set theory, viz., forming the *extension of a predicate, Cartesian products*, and *function spaces*. In a topos, as in set theory, every object—and indeed every arrow—can be considered in a certain sense as the extension $\{x: P(x)\}$ of some predicate P. The difference between the two situations is that, while in the set-theoretic case the variable *x* here can be construed *substitutionally*, i.e. as ranging over (names for) individuals, in a general topos this is no longer the case: the *"x"* must be considered as a *true variable.* More precisely, while in set theory one has the rule of inference

$$\frac{P(a) \text{ for every individual } a}{\forall x P(x)}$$

in general this rule fails in the internal logic of a topos. In fact, assuming classical set theory as metatheory, the correctness of this rule in the internal logic of a topos forces it to be a model of classical set theory: this result can be suitably reformulated in a constructive setting.

5. A recent development of great interest in the relationship between category theory and set theory is the invention by Joyal and Moerdijk [3] of the concept of *Zermelo-Fraenkel algebra*. This is essentially a formulation of set theory based on set *operations,* rather than on properties of the membership relation. The two operations are those of *union* and *singleton*, and Zermelo-Fraenkel algebras are the algebras for operations of these types. (One notes, incidentally, the resemblance of Zermelo-Fraenkel algebras to David Lewis' "megethelogical" formulation of set theory in [4].) Joyal and Moerdijk show that the usual axiom system *ZF* of Zermelo-Fraenkel set theory with foundation is essentially a description (in terms of the membership relation) of the *free* or *initial* Zermelo-Fraenkel algebra, just as the Peano axioms for arithmetic describe the free or initial monoid on one generator. This idea can be extended so as to obtain a charcterization of the class of *von Neumann ordinals* as a free Zermelo-Fraenkel algebra of a certain kind. Thus both well-founded set theory and the theory of ordinals can be characterized category-theoretically in a natural way.

6. Category theory does much more than merely reorganize the mathematical materials furnished by set theory: its function far transcends the purely cosmetic. This is strikingly illustrated by the various topos models of *synthetic differential geometry* or *smooth infinitesimal analysis* (see, e.g., [2]). Here we have an explicit presentation of the form of the smoothly continuous incorporating actual infinitesimals which is simply *inconsistent* with classical set theory: a form of the continuous which, in a word, *cannot* be reduced to discreteness. In these models, *all* transformations are smoothly continuous, realizing Leibniz's dictum *natura non facit saltus* and Weyl's suggestions in *The Ghost of Modality* [5], and elsewhere. Nevertheless, extensions of predicates, and other mathematical constructs, can still be formed in the usual way (subject to intuitionistic logic). Two further arresting features of continuity manifest themselves. First, connected continua are *indecomposable:* no proper nonempty part of a connected continuum has a "proper" complement—cf. Anaxagoras' c. 450 B.C. assertion that the (continuous) world has no parts which can be "cut off by an axe". And secondly, any curve can be regarded as being traced out by the motion of an *infinitesimal tangent vector*—an entity embodying the (classically unrealizable) idea of *pure direction*—thus allowing the direct development of the calculus and differential geometry using nilpotent infinitesimal quantities. These near-miraculous, and yet natural ideas, which *cannot* be dealt with coherently by reduction to the discrete or the notion of "set of distinct individuals" (cf. Russell, who in *The Principles of Mathematics* roundly condemned infinitesimals as "unnecessary, erroneous, and self-contradictory"), can be explicitly formulated in category-theoretic terms and developed using a formalism resembling the traditional one.

Establishing the consistency of smooth infinitesimal analysis through the construction of topos models is, it must be admitted, a somewhat laborious business, considerably more complex than the process of constructing models for the more familiar (discrete) theory of infinitesimals known as *nonstandard* *analysis*. I think the situation here can be likened to the use of a complicated film projector to produce a simple image (in the case at hand, an image of ideal smoothness), or to the cerebral activity of a brain whose intricate neurochemical structure contrives somehow to present simple images to consciousness. The point is that, although the fashioning of smooth toposes is by no means a simple process, it is designed to realize simple principles. The path to simplicity must, on occasion, pass through the complex.

References

[1] Bell, John L. *From Absolute to Local Mathematics.* Synthese **69** (1986), 409-426.

[2] Bell, John L. *A Primer of Infinitesimal Analysis.* Cambridge University Press, 1998.

[3] Joyal, A. and Moerdijk, I. *Algebraic Set Theory.* London Mathematical Society Lecture Notes Series 220. Cambridge University Press, 1995.

[4] Lewis, D. *Mathematics is Megethology.* Philosophia Mathematica (3) **1** (1993), 3-23.

[5] Weyl, H. *The Ghost of Modality.* In *Philosophical Essays in Memory of Edmund Husserl.* Cambridge (Mass.), 1940, pp. 278-303.