Avoiding the axiom of choice in general category theory

by M. Makkai (McGill University)

Introduction

In Category Theory, there is an underlying principle according to which the right notion of "equality" for objects in a category is isomorphism. Let me refer to the principle as the *principle of isomorphism*. According to the principle of isomorphism, any object isomorphic to a given one should be able to serve the same categorical purposes as the given one. Of course, the principle of isomorphism may be read as a limitation on what properties of objects are to be considered in category theory; but the principle also carries with it the assertion that by so restricting the properties of objects, we are not losing any essential element of the situation.

Therefore, when singling out an object with a certain property, we should be content with determining the object up to isomorphism only. Indeed, the categorical operations defined by universal properties (products, exponentials, etc) determine the object-parts of their values at given arguments only up to isomorphism. The idea behind the notion of *anafunctor*, the main new concept in this paper (see 2.1.(i) to (v) below; a reference of the form m.n.(...) is to item n.(...) in Section m) is that the same principle should extend to values of functors: their object-values are to be determined up to isomorphism only.

General category theory in its usual form does not quite live up to the principle of isomorphism; the ubiquitous use of the Axiom of Choice in general category theory is a related fact. A simple example is at hand when, for a category $\mathbf{C}$ having binary products of objects, we pass to the consideration of "the" product functor $P=(\ )\times(\ ):\mathbf{C}\times\mathbf{C}\longrightarrow\mathbf{C}$. The definition of P requires the simultaneous *choice* of a specific product $(A\xleftarrow{\pi_1}A\times B\xrightarrow{\pi_2}B)$ corresponding to each pair (A, B) of objects. To be sure, in most *examples* of a category $\mathbf{C}$ such a simultaneous choice can be made without the Axiom of Choice; however, we want to use the product functor in the theory for any category $\mathbf{C}$ with binary products, without knowing anything further about $\mathbf{C}$. Whether or not an *explicit* choice of products is available, something of the *canonicity* of the resulting entity (functor) is lost when we make a *particular* choice of products. Actually, talking about *the* product functor becomes imprecise; there are,

The author's research has been supported by NSERC Canada and FCAR Quebec

in general, many possible product functors.

The general form of the above type of use of the Axiom of Choice is in taking "the" adjoint of a functor on the basis of the representability of a family of `Set`-valued functors derived from the given functor. Every time we use the Adjoint Functor Theorem to get an adjoint, we use the Axiom of Choice in the described manner.

There are similar violations of canonicity and attendant uses of the Axiom of Choice in the definitions of various concrete monoidal categories, and higher dimensional categorical objects.

In this paper, I propose a revision of the notion of functor, that of anafunctor, and consequent revisions of certain higher dimensional concepts, that makes possible a theory based more thoroughly on canonical constructions than ordinary category theory, and specifically, that rectifies the violations described above of the principle of isomorphism. The revisions are non-intrusive in the sense that category theory with anafunctors is of the same general shape as with ordinary functors. It seems that there is no limitation of the applicability of anafunctors in any context where functors are used. The resulting theory avoids Choice to a large extent (although not completely; see below), and still has the same general form as classical general category theory. If one employs the full Axiom of Choice, the new theory reduces to the classical one. Without the Axiom of Choice, we have a product *ana*functor $P=(\ )\times(\ ):\boldsymbol{C}\times\boldsymbol{C}\longrightarrow\boldsymbol{C}$ *defined canonically* on the basis of of $\boldsymbol{C}$ having binary products. The adjoint of a(n) (*ana*)functor, an *ana*functor, is given canonically once the condition mentioned above on representability is fulfilled. Anafunctors have natural transformations, which are the arrows of a category as usual; categories with anafunctors and natural transformations form a bicategory. We have *ana*bicategories, *ana*monoidal categories, with basic theories similar in outline to those of their usual counterparts.

Whereas "anafunctor" is a generalization of "functor", a certain specialization of the notion of "anafunctor", "*saturated anafunctor*" is the one that should be regarded as the finished form of the concept; an ordinary functor is (usually) not a saturated anafunctor. Saturated anafunctors (1.(vi)) satisfy the analog of Leibniz's principle of substitutability of equal for equal: if an object in the codomain category of the saturated anafunctor is isomorphic to a value of the anafunctor, then it is itself a value of the anafunctor at the same argument, "in a uniquely determined way". It turns out that saturated anafunctors are sufficient; there is a canonical way of "saturating" any anafunctor, the result of which, a saturated anafunctor, is isomorphic (via a canonical natural isomorphism) to the given anafunctor.

The most important difference between using anafunctors and using functors is a result of the fact that the category $\mathrm{Ana}(\boldsymbol{X},\boldsymbol{A})$ of (small) anafunctors between two fixed small categories $\boldsymbol{X}$ and $\boldsymbol{A}$ is not small (unless $\boldsymbol{X}$ or $\boldsymbol{A}$ is empty). However, under the assumption of a certain weak consequence, here called the Small Cardinality Selection Axiom (SCSA), of the Axiom of Choice, $\mathrm{Ana}(\boldsymbol{X},\boldsymbol{A})$ is equivalent (in fact, in the strong sense) to a small category. Thus, the SCSA ensures the Cartesian closed character of the bicategory of small categories with anafunctors and natural transformations (with "Cartesian closed" meant in the natural bicategorical sense). The SCSA is closely related to A. Blass' axiom (in [Bl]) of Small Violations of Choice (SVC), another weak choice principle.

There is a well-known and important approach to category theory relative to a largely arbitrary topos. See [Be2], [P], [P/Sch], [Jo]. The theory uses the formalism of indexed categories ([P/Sch], [Jo]), or alternatively and essentially equivalently, that of fibrations ([Be1], [Be2]). Category theory done internally in $\mathcal{E}$ is a part of indexed category theory over $\mathcal{E}$. Indexed category theory over $\mathcal{E}$ may use the axiom of choice externally. For instance, in [P/Sch], a form of the Initial Object Theorem is proved, and from this, an appropriate form of the Adjoint Functor Theorem is inferred, by the same kind of use of the axiom of choice as the one that goes into constructing the product functor mentioned above.

The approach of the present paper is, in a sense, orthogonal to that of indexed category theory; neither approach does what the other does, but they can be combined to work together. When a topos lacks the necessary axiom of choice, the product functor mentioned above for an internal category with products (where the mere existence of products, rather than their specifiability, is assumed internally) does not exist internally, and will not exist for the externalization, an indexed category, of the internal category. However, the present paper's approach will provide an internal anafunctor in place of the product functor without assuming Choice in the topos. In fact, the development of the present paper, can be relativized to any topos. In [M/P], anafunctor theory will be put into the context of indexed category theory over a topos, and a connection will be established with stacks and stack completions. It will be shown that a suitable variant of the SCSA, one that is equivalent to saying that internal categories have internal stack-completions, will ensure that the bicategory of internal categories, internal anafunctors and natural transformations is Cartesian closed.

The present paper is only the beginning of the development of "anafunctor theory". Let me briefly indicate an area of category theory where anafunctors are relevant. This is the general (or universal) algebra of structured categories. The usual kinds of structured categories (lex categories, regular categories, (elementary) toposes (in this case, use only isomorphism 2-cells), and many more) form *locally finitely presentable bicategories*. The latter have a theory formally similar to that of locally finitely presentable categories of [G/U] . This theory has only partly been codified at the present time, but various key elements of it, such as the theory of bicategorical (indexed or weighted) limits (see, *e.g.*, [S]), have been clarified. The sequel [Ma2] will deal with locally finitely presentable bicategories and related matters by employing anafunctors, giving more canonical answers to existence questions than the usual theory, and avoiding the Axiom of Choice. I now give two indications, to be worked out in *loc.cit.*, why anafunctors are useful for a "canonical" version of the general algebra of structured categories.

One may maintain that, when dealing with a category $\boldsymbol{C}$ with finite products, it is not necessary to invoke the product functor $(\ )\times(\ ):\boldsymbol{C}\times\boldsymbol{C}\longrightarrow\boldsymbol{C}$; after all, all that this does is to pick a particular product for each argument-pair, and we have the experience that in most cases this is not necessary. However, if we want to make the theory of categories with finite products (fp categories) a part of the algebra of structured categories along the lines hinted at above, then the product functor is forced on us. In this theory, $\boldsymbol{C}$ induces a functor (a restricted representable functor) $\dot{\boldsymbol{C}}:\mathrm{FP}_{\mathrm{f}}^{\mathrm{op}}\to\mathrm{Cat}$ on the opposite of the bicategory FP_{f} of the finitely presentable fp categories to the bicategory Cat of small categories, and the above product functor is the value of $\dot{\boldsymbol{C}}$ at the 1-cell $[X]\longrightarrow[Y,Z]$ in FP_{f} ; here, $[X]$ is the fp category freely generated by the object X , similarly for $[Y,Z]$, and the arrow is induced by the mapping $X\longmapsto Y\times Z$. (The mapping $\boldsymbol{C}\mapsto\dot{\boldsymbol{C}}$ is the basic identification of the objects of a locally finitely presentable bicategory with a Cat-valued functor. The reader will be familiar with the 1-dimensional analog of the described constructions; replace FP_{f} with $\mathrm{Ring}_{\mathrm{f}}$, the category of finitely presentable commutative rings with 1, replace Cat by Set , take a ring R in place of $\boldsymbol{C}$, take $\times$ to be the multiplication in R , and the above will refer to the multiplication-operation $()\cdot():R\times R\longrightarrow R$.) In brief, the point of view of the bicategorical algebra of structured categories necessitates the consideration of something like the product functor. We have mentioned that anafunctor theory is capable of providing the needed entity in a canonical fashion.

Another example for the use of anafunctors is as follows. Consider the notion of the free structured category $\mathcal{F}(\boldsymbol{G})$ of a given kind generated by the graph $\boldsymbol{G}$. For the sake of a convincing example, let us talk about categories with finite limits and finite colimits (without any further restriction) as the given kind. Suppose $\boldsymbol{G}$ is a finite graph. In this case, $\mathcal{F}(\boldsymbol{G})$ has an explicit description, consisting of iterated formal limits and colimits, starting with the generators; in particular, certainly, there is no need for Choice in the construction of $\mathcal{F}(\boldsymbol{G})$. (Andre Joyal has recently given a beautiful theory of just this free construction, and its enriched generalizations.) However, to verify the universal property of $\mathcal{F}(\boldsymbol{G})$, against *all* maps $\varphi:\boldsymbol{G}\longrightarrow\boldsymbol{C}$ into a category $\boldsymbol{C}$ with finite limits and colimits, in the usual theory we do need some form of the axiom of choice. In fact, we are required to construct a functor $F:\mathcal{F}(\boldsymbol{G})\longrightarrow\boldsymbol{C}$ preserving finite limits and colimits and satisfying the initial conditions given by φ. The construction of F requires a series of choices of limits and colimits in $\boldsymbol{C}$, which cannot be done without Choice. The use of an anafunctor in place of F eliminates the need of the Axiom of Choice, and in fact makes F canonical.

Of course, the last example is a crucial one for the general algebra of structured categories; in this theory, we would not want to do without free objects such as $\mathcal{F}(\boldsymbol{G})$.

Let me turn to remarks on the set-theory used in the paper.

The set-theoretic foundations used in this paper are "minimal", and probably the reader will have no problem following the paper even if he skips these (brief) preliminaries.

We work in a constructive set-theory with sets and classes. For the sake of definiteness, we take as our foundations the Gödel-Bernays (G-B) axioms for sets and classes [G], without the Axiom of Choice, and without the Axiom of Regularity (Foundation), and we employ intuitionistic predicate logic to deduce consequences of the axioms. (We could accommodate ur-elements, but to do so would require some explanations that we do not want to give; thus, all things in our theory are classes, and some classes (precisely those that are elements of some class) are sets; the axiom of extensionality is assumed in an unrestricted form.) We do not use Grothendieck universes.

The use of the adjective "small" will, as usual, signify that the entity it qualifies is a set. Thus, a small class is the same thing as a set.

A *category* $\boldsymbol{A}$ is given by a *class* of objects $\mathrm{Ob}(\boldsymbol{A})$, and a *class* $\mathrm{Arr}(\boldsymbol{A})$ of arrows, with further data as usual. Thus, we do not make the blanket assumption that a category has small hom-sets; if it does, it is said to be *locally small*. A *small* category has both $\mathrm{Ob}(\boldsymbol{A})$ and $\mathrm{Arr}(\boldsymbol{A})$ sets; of course, $\mathrm{Arr}(\boldsymbol{A})$ being a set implies that $\mathrm{Ob}(\boldsymbol{A})$ is one as well. A small category can be regarded as a single set (e.g., as a tuple $(|\boldsymbol{A}|, \mathrm{Arr}(\boldsymbol{A}), ...)$), and we may talk about the *class* (and eventually, the *category*) of all small categories.

Note that a category isomorphic to a small category *is* small (by the Axiom of Replacement).

Within G-B, one cannot talk about the category of all functors $\boldsymbol{X} \to \boldsymbol{A}$ for two fixed, but arbitrary categories $\boldsymbol{X}$, $\boldsymbol{A}$; there are no collections whose members are proper classes. Of course, there is no problem when the categories $\boldsymbol{X}$, $\boldsymbol{A}$ are small, or even when just $\boldsymbol{X}$ is small (since in the latter case functors $\boldsymbol{X} \to \boldsymbol{A}$ are (may be regarded as) sets). However, within the framework of the formal base-theory G-B, we may contemplate *metacategories*; an example is $\mathrm{FUN}(\boldsymbol{X}, \boldsymbol{A})$, the metacategory of all functors $\boldsymbol{X} \to \boldsymbol{A}$ and natural transformations. Formally, a metacategory is given by predicates (formulas) $\mathrm{Ob}(X, \vec{P})$, $\mathrm{Arr}(f, \vec{P})$, $\mathrm{Dom}(f, X, \vec{P})$, $\mathrm{Codom}(f, X, \vec{P})$, $\mathrm{Comp}(f, g, h, \vec{P})$ of the base-theory (in our case, G-B), with the free variables shown, all ranging over *classes*, together with the assumption that, for a fixed value of the parameters $\vec{P}$, the obvious equivalents of the category axioms (which become first order formulas, having only $\vec{P}$ as free variables, built up of the given predicates) hold. The said assumption may be a consequence of an assumption $C(\vec{P})$ on the parameters $\vec{P}$. In the case of $\mathrm{Fun}(\boldsymbol{X}, \boldsymbol{A})$, $\vec{P}$ is $\boldsymbol{X}, \boldsymbol{A}$ [although a category $\boldsymbol{X}$ is given by classes $|\boldsymbol{X}|$, $\mathrm{Arr}(\boldsymbol{X})$, ..., these can be combined, although somewhat artificially, into a single class; if we do not want to do this, $\vec{P}$ will be a longer tuple, listing all the data-classes of both categories $\boldsymbol{X}, \boldsymbol{A}$], and $C(\vec{P})$ is the assumption that $\boldsymbol{X}$, $\boldsymbol{A}$ are indeed categories. Of course, the idea of a metacategory is just one instance of a family of meta-concepts similarly fashioned from a formal concept such as "category". One can e.g. talk about CAT , the meta-2-category of all categories, functors and natural transformations; Cat is the 2-category of *small* categories, functors and natural transformations.

Let me note that I will usually drop the "meta" prefix from constructs such as metafunctor, meta-natural transformation, etc.

Although [F/S] does not mention a formalized base-theory in which the exposition is made, it is rather clear that a class-set theory is meant such as G-B; no universes are employed. On the other hand, the explicit base-theory in [CWM] is Zermelo-Fraenkel (ZF) set theory, a theory of sets without class-variables. One universe (a set U with appropriate properties) is used, and the word "small" is reserved for members of U. [CWM] uses "class" in a somewhat non-standard manner; classes in [CWM] are non-small *sets*.

Our base-theory is like that of [F/S]; in particular our categories and the categories in [F/S] may be large (classes); the word "small" is used here in agreement with [F/S]; however, [F/S] does not mention "metacategories". The "metacategories" of [CWM] are our categories. Our metacategories are introduced on the same principle as those of [CWM], but the difference in the base-theories makes the meanings of the term different.

The use of the prefix "ana" has been suggested by Dusko Pavlovic. He noted the use of "pro-" in category theory (profunctor, proobject), and noted that in biology, the terms "anaphase" and "prophase" are used in the same context.

At a time when the work on this paper had essentially been completed, Robert Paré told me that he had had related ideas in the 1970's, and he had lectured about them at a meeting in New York in 1975, although he had not published his work.

Some time after the first version of this paper was written, I was informed that a special case of the notion of anafunctor, and of the notion of natural transformation of anafunctors, the case when the domain category is $\mathbf{1}$, the terminal category, have been introduced in [J/S], under the name of "clique" and morphism of cliques. In [J/S], cliques are used for certain special purposes; beyond the definition of cliques and their morphisms, there is essentially no overlap between [J/S] and this paper. For more precise references, see toward the end of §1 of this paper.

I thank Dusko Pavlovic, Bill Boshuck, Marek Zawadowski, Gonzalo Reyes, Bill Lawvere, Ervin Fried and Bob Paré for valuable conversations on the subject of this paper. I also thank the Referee of the paper for two very careful readings of the paper, and for several remarks that helped improve the paper substantially.