

# Why tricategories?

A.J.Power\*

Department of Computer Science  
University of Edinburgh  
King's Buildings  
Edinburgh EH9 3JZ  
Scotland.

April 11, 1994

## Abstract

We outline a logical framework, based on the theory of categories with extra structure, for logics that arise in computer science. We list many representative examples of structures that have arisen, then we classify them in terms of equational, and the more general essentially algebraic, structure. In both cases, we outline the main results and their significance for the logical framework. This study gives rise to coherence questions. We explain the issues, and then outline the category theoretic concepts, such as tricategories, that arise in resolving the coherence problems.

## 1 Introduction

The primary motivation for tricategories that appears in the literature is in algebraic topology: a one-object tricategory is a monoidal bicategory [BW], and a one-object one-arrow tricategory is a braided monoidal category, sometimes called a braided tensor category [JS]; and it is under these guises that tricategories generally appear. However, the computational motivation and use of tricategories that I propose is quite different. So in this article, I wish to explain precisely how tricategories arise from a consideration of computing, and precisely what theorems and insights are gained as a result.

Tricategories arise in attempting to develop a logical framework for the logics that arise in computing. When I say logical framework, I do not mean a type theory which serves as a metatheory for a bunch of logics, along the lines of Martin-Löf Type Theory or  $LF$ , which, owing to their linguistic nature, have immediate computational interpretations. What I mean is an abstract framework in which to study semantically the logics that arise from a consideration of computing. The idea is to move away from the syntax that is inherent in type theories in an attempt to elucidate the relative power of the central constructs of the type theories, rather than concentrating on syntactic translations. More precisely, the aim is for a more semantic approach to the idea of a logical framework.

---

\*This research was supported by ESPRIT Basic Research Action 6453: Types for proofs and programs.

Semantics arises at two levels here. First, we seek a semantic logical framework as outlined above. But also, within that logical framework, we seek to study the semantics, or model theory, of logics. This is a departure from the view taken by many of the extant logical frameworks such as *LF* [HHP], which are based upon syntax. Here, we seek to unify not so much syntax as the way the various logics are used in computer science. Some logics, such as those used in logic programming, are used syntactically. Others, such as higher order intuitionistic logic, are often used semantically, for instance via a topos. In both cases, the use of a logic only makes sense to the extent that it resolves a semantic question. For instance, one may be asked to write a *PROLOG* program to check whether two given nodes of a given finite graph are connected: although the program is a formal construction and it has an underlying set of horn clauses to which it is syntactically related, the program resolves a semantic question. Moreover, constructions on programs are often fundamentally semantic constructions, which are insightful to describe without the syntax. For example, in [PS], in defining a map of logic programs, used for systematic program development of logic programs, although the work is all expressed syntactically, the central definition is inherently a semantic definition (see Example 2.12 herein). So we attempt to unify the role of these logics via a unified study of their semantics.

The study of categories with equational or “essentially algebraic” structure provides such a semantic study. Examples of such structure are detailed in Section 2: they include cartesian closed categories, toposes, and fibrations with certain extra structure. On that basis, I regard this study informally as a logical framework along the lines outlined above. It incorporates logics used in functional and logic programming, as well as a unified and expanding study of semantics of several features of functional programming. The study is far from complete and as yet has not adequately addressed logics used to study the dynamic behaviour of programs. Of course, that is a major deficiency, one that is shared by many logical frameworks. However, I believe there has been sufficient success to warrant further research.

Categories with equational structure are relatively well understood. There has been a succession of papers addressing the major issues that arise, detailed in Section 3. They unify a substantial class of logics, such as minimal logic and higher order intuitionistic logic. More general and much less well understood are categories with “essentially algebraic” structure: these allow operations to be defined not just universally, but on equationally defined categories, as explained in Section 4. They substantially increase the class of logics we can include. Several attempts have been made to study such structure, but the difficulties are substantial. Indeed, the study is so little developed and the problems so profound that one can really only start to resolve one problem at a time and hope that the others are also tractable. There has been enough success to encourage further investigation, but it is by any measure slow. One of the major problems is a coherence problem: in the very definitions that one is led to use, there are so many isomorphisms satisfying so many conditions that even in simple examples it is hard to retain a coherent, rigorous account. The difficulties are like those of a full abstraction problem, or a word problem, or of taking account of the various notions of variable conversion in type theories. So one seeks coherence results. These, in one incarnation, say that one may without loss of generality replace a structure with many isomorphisms by one with fewer isomorphisms. It is in this regard that tricategories arise: as a tool with which to prove a difficult coherence result, specifically a result that shows one may replace “bilimits” by

the simpler “pie-limits”.

This paper is divided into two parts. The first part, which comprises most of the paper, is largely expository, with few precise definitions and results. The aim is to explain the thinking behind the definitions and results rather than give the results themselves. However, they are referenced in detail. This part consists of Sections 2 to 5. In Section 2, categories with structure are suggested as a possible logical framework, with many examples of how they arise. In Section 3, the development of a unified study of categories with equational structure is outlined. Section 4 commences a generalisation from categories with equational structure to categories with essentially algebraic structure. Section 5 explains how a plethora of isomorphisms and equations between them arise, and outlines the coherence problem that must be resolved.

The second part of the paper is shorter, more technical, and written in the more usual style. It consists of Sections 6 and 7, and is meant for readers already familiar with category theory. This part explains the technical category theoretic concepts, such as tricategories, that are required to resolve the above mentioned coherence problem. Section 6 outlines our approach to resolving the problem; and Section 7 explains the role of tricategories in a resolution of the difficulty, and outlines a few of the attendant results.

Specifically, the problem is one of bilimits and pie-limits. Traditionally, the treatment of categories with structure has been addressed in terms of bilimits, as outlined in Section 5. Bilimits require coherent isomorphisms wherever ordinary limits require equality of arrows. This leads to a plethora of isomorphisms and coherence conditions that say which composites of isomorphisms agree. Categorically, the definitions are eminently natural. But the isomorphisms are overwhelming. So one would like to simplify the situation by the replacement of bilimits by pie-limits, since the latter have equality in some places where the former require coherent isomorphisms. There is a tricategory of small 2-categories with finite bilimits, and a tricategory of small 2-categories with finite pie-limits. One may use results about tricategories to prove the two tricategories equivalent.

Of course, the results and definitions quoted here build on the work of many people, but I should especially mention Max Kelly, who has been responsible for many of the category theoretic ideas. Also, I should like to express my gratitude to Stuart Anderson, David Pym and Edmund Robinson for insightful relevant conversations.

## 2 Categories with structure

There are many specific examples of and results about categories with structure that arise, logically although typically not historically, from computing. A selection is as follows.

**Example 2.1** A *cartesian closed category* arises as a model of typed  $\lambda$ -calculus and hence as a model of the proof theory of minimal propositional logic [LS].

In denotational semantics, one also considers cartesian closed categories with extra structure such as the following.

**Example 2.2** A *cartesian closed category with finite sums and a natural numbers object* (see [T]).

For the study of nondeterminism, there arises a consideration of further structure as follows.

**Example 2.3** A cartesian closed category with finite sums, together with a monad  $P$  on it, and a natural transformation  $\cup : P \times P \rightarrow P$  satisfying equational axioms. (See [AP] for a discussion of this.)

On the logical side, there are many examples of categories with structure.

**Example 2.4** A category with finite products, otherwise known as a cartesian category may be used to model many sorted equational logic. (See Makkai and Pare's book [MP] for an excellent general account of this phenomenon, unifying logics, categories with extra structure, and their models in the category of small sets. Their account is in the more restricted setting of a category with finite limits and finite colimits.)

A particularly useful structure is as follows.

**Example 2.5** An elementary topos. This is a model of higher order intuitionistic logic (see [LS]).

**Example 2.6** An elementary topos with natural numbers object allows the construction of classifying toposes for theories and is used by Fourman and Phoa in their proposed semantics for  $ML$  [FP].

In unifying logic with semantics, one also considers the following example.

**Example 2.7** A locally cartesian closed category, such as the category of  $\omega$ -sets.

Such also arise as an extensional version of Martin-Löf Type Theory [S2], but that view is somewhat controversial.

Logic programming is well suited by various fragments of linear logic (see [HP1] and [HP2]), and many of those fragments have models given by categories with extra structure.

**Example 2.8** A symmetric monoidal category or a monoidal category, models  $\otimes$  in linear logic.

**Example 2.9** A cartesian closed category with one symmetric monoidal structure and one symmetric monoidal closed structure, together with finite coproducts, a monad and a comonad, satisfying equational conditions models propositional linear logic. (See Carolyn Brown's thesis [Br]).

There is also structure that corresponds to a first order linear logic (see [A]) but it fits better in a later example.

An important question to ask of the above examples of categories with structure and the way they arise is whether they may all be seen as instances of just one correspondence between logic and categories with structure, and one correspondence between programming languages and such categories. The answer to that is no, but it is a qualified no. Much of the semantics is unified, as one can see in any good text on the subject, such as [T]. However, for the logics, relating minimal propositional logic and cartesian closed categories, propositions of the logic correspond to objects of the category, whereas for higher order intuitionistic logic and elementary toposes, objects of the category amount to types, and propositions are modelled by subobjects of them.

This can be resolved if one adds somewhat more complexity and resorts to a study of fibrations with extra structure, as strongly encouraged by Bénabou in [B2] and as exemplified by the work in [H], [J] and [P]. Bill Lawvere was the first to notice this resolution. A fine recent analysis of this in the case of a first order linear logic appears in Ambler's thesis [A]. Formally, a fibration is essentially a category  $C$  together with a functor from  $C$  into  $Cat$ . That is somewhat misleading, but I have stated it to give some sense to later examples. One also resorts to more complex structure, such as pairs of categories with structure, for semantic reasons. For instance, to model partiality, one may consider a base category of total maps and a larger category of partial maps. Various structures have been used to represent this idea (see Robinson and Rosolini's [RR] for a fine account). Two such structures are as follows.

**Example 2.10** *A category with an admissible class of subobjects* (formerly called a category with a dominion).

**Example 2.11** *A partial cartesian closed category.*

In logic programming, implicit in but central to [PS] is a monoid together with a category with finite products and an action of the monoid on the category. This construction was made on the assumption that terms are untyped, but if one put a typing structure on the terms, with the eminently reasonable assumption in that setting that it be closed under finite products, one would have the following.

**Example 2.12** *A category  $C$  with finite products and an  $obC$ -indexed family of categories with finite products, together with an action of  $C$  on the family, i.e., a fibration with added equational structure.*

Another logical example is given by the Calculus of Constructions and its correspondence, as shown by Hyland and Pitts [HPi] with the following structure.

**Example 2.13** *A fibration with certain "essentially algebraic" structure.* The specific structure of Hyland and Pitts requires substantial definition, so I omit the details here.

This construction is very general, variants giving fibrations with structure corresponding to the  $\lambda\Pi$ -calculus, which is the type theory of the  $LF$  logical framework,  $F_{\omega}$ , and their variants, spawning much recent work. Closely related, and preceding, work is the correspondence established by Seely between first order intuitionistic theories with equality and the following.

**Example 2.14** *A hyperdoctrine* [S1].

Similar structure appears widely in the literature. One such is as follows.

**Example 2.15** *A comprehension category with added structure* as in Jacobs' thesis [J].

We seek a unified theory of all of these structures, with sufficient flexibility and elegance to incorporate the likelihood of many variants.

### 3 Categories with equational structure

In seeking a unified treatment of categories with extra structure, our strategy is simple: we consider the study of sets with extra structure, we attempt to lift the central ideas from sets to categories, and we make whatever modifications are necessitated by the special features of categories in order to deduce the results we seek. One such feature is the existence of the opposite of a category: the opposite of a discrete category, i.e., a set, is just itself. Another important feature is the fact that two objects of a category may be isomorphic without being equal, whereas any two isomorphic objects of a discrete category are necessarily equal. One kind of result we seek is a characterization of the models that arise (see [KP1] Section 5). Another is a characterization of the limiting or colimiting properties of the category of such models ([BKP] Section 2 and Thm 5.8).

In a study of sets with structure, one may profitably start with a study of universal algebra, i.e., a study of sets with equational structure. It has long been known that every category of algebras for a one-sorted signature, subject to equations, is equivalent to the category of algebras for a finitary monad on the category of sets (see [M] Thm 6.8.1; the term "finitary" is a size condition: we will not need a precise definition here). Kelly and others have generalised this analysis from sets to categories, and the theory is now well established.

Let  $Cat_0$  denote the category of small categories and functors. In ordinary universal algebra, an algebra is a set  $X$  together with a family of basic operations  $\sigma : X^n \rightarrow X$ , subject to equations between derived operations. In order to define equational structure on categories, one must of course replace the set  $X$  by a category  $C$ . One also replaces the finite number  $n$  by a finitely presentable category  $c$ . One finally allows not only functions from  $Cat_0(c, C)$  into the set of objects of  $C$ , but also functions from  $Cat_0(c, C)$  into the set of arrows in  $C$ . These remain subject to equations between derived operations. Details appear in [KP1] Section 5, and further discussion and examples appear in [DK] and [Po3]. One may then prove that every category of small categories with specified equational structure is equivalent to the category of algebras for a finitary monad on  $Cat_0$ , and conversely ([KP1] Thm 5.1). Indeed, it follows that for each of examples 2.1 to 2.10, the category of small such categories with structure and functors that strictly preserve the structure, i.e., preserve the structure up to equality in the usual naive algebraic way, is equivalent to the category of algebras,  $T_0 - Alg$ , for a finitary monad  $T_0$  on  $Cat_0$  (see [DK] for many of the examples).

This is encouraging, but is not delicate enough for our purposes. An example of a central problem is as follows. Consider cartesian categories, Example 2.4. We seek a study not only of cartesian categories, but also of the maps between them. For instance, a model in *Sets* of an equational theory is a finite product preserving functor from a cartesian category  $Th$  into *Sets*, but that functor need not strictly preserve finite products. For a specific example, let  $M$  be the Lawvere (equational) theory of a monoid. Then, the terminal object  $1$  of  $M$  has the property that  $X \times 1 = X$  for every object  $X$  of  $M$ . However, that is in general false in the category of sets, i.e., for a set  $X$ , it is in general not true that  $X \times 1 = X$  unless  $X$  is empty: they are merely isomorphic. It follows that the category  $FP_s(M, Sets)$  of functors from  $M$  to *Sets* that strictly preserve finite products is  $\emptyset$ , the empty category. The solution to this problem is to weaken the notion of preservation of structure. One says that a functor  $H : C \rightarrow D$  preserves binary products

if, for every pair of objects  $(X, Y)$  in  $C$ , the uniquely determined comparison map from  $H(X \times Y)$  into  $H(X) \times H(Y)$  is an isomorphism, not necessarily an equality. One can routinely generalise this to a notion of preservation of finite products. Then, the category  $FP(M, Sets)$  of finite product preserving functors from  $M$  to  $Sets$  is equivalent to the category of monoids, as one wishes.

For another example, if we wish to speak in this setting of an encoding of a logic  $L$  in some syntactic logical framework with signature  $\Sigma_L$ , then we need to consider an equivalence of categories with structure between  $C(L)$  and  $C(\Sigma_L)$ , the categories freely formed from  $L$  and  $\Sigma_L$  respectively, as in Philippa Gardner's thesis [G]. We want the structure of the category  $C(L)$  to transport along equivalence to corresponding structure on  $C(\Sigma_L)$ , but if we restrict ourselves to functors that preserve structure strictly, then not even the simplest of structure is transported. For instance, suppose  $F \dashv U : C \rightarrow D$  is an equivalence of categories, and suppose  $C$  has terminal object  $1$ . One would like to deduce that  $D$  also has a terminal object and that the equivalence respects the terminal objects. Indeed,  $U(1)$  is a terminal object in  $D$ . However,  $F$  does not strictly preserve it since  $UF(1)$  is in general not equal to  $1$ : in general, the unique map from  $UF(1)$  to  $1$  is an isomorphism, not an equality. Again, in order to capture this phenomenon, we need an account of those functors that preserve structure up to coherent isomorphism rather than those that preserve it strictly.

The above are just two examples, but they are representative. Many of the deeper results of category theory such as the syntax/semantics equivalence between finite limit theories and locally presentable categories (see [GU], [MP], and [K1] Thm 9.8) break down if we restrict attention to functors strictly preserving structure. So we consider functors that preserve structure up to coherent isomorphism. By "coherent", I mean that one typically does not need the mere existence of an isomorphism, but one needs that isomorphism to satisfy equations. For instance, in our monoid example, when looking at preservation of binary products, we needed our isomorphism to be a specific map, the comparison map from  $H(X \times Y)$  to  $H(X) \times H(Y)$ . In more delicate examples such as that of monoidal categories, i.e., Example 2.8, we need several and more delicate equations (see [BKP] Section 1.2 for the general situation).

In fact, there is a well developed analysis of categories with equational structure and functors that preserve that structure. Blackwell, Kelly and Power's [BKP] is the central paper and is devoted specifically to the connection between preservation and strict preservation of equational structure. The theory developed there incorporates all our examples from 2.1 to 2.9. It rests upon the fact that each of the monads  $T_0$  corresponding to each example lifts uniquely to a 2-monad  $T$  on the 2-category  $Cat_g$  of small categories, functors, and natural isomorphisms [KP3]. Some of the central results of the theory are as follows.

For any monad on the category of sets, its category of algebras is complete, with limits given pointwise. So, in particular, it has equalizers. However, if we denote the 2-category of  $T_0$ -algebras and functors that preserve the structure in the correct sense by  $T-Alg$ , then in the case of cartesian categories,  $T-Alg$  does not have equalizers: the unit category  $1$  is cartesian, and so is the unique two-object category  $Iso$  that is equivalent to  $1$ . There are precisely two functors  $F$  and  $G$  from  $1$  to  $Iso$ , and both preserve finite products. So they should have an equalizer. However, the only functor  $H : E \rightarrow 1$  such that  $FH = GH$  is the unique functor from the empty category, but the empty category is not cartesian since

it does not have a terminal object. This situation is typical, and indeed this argument is general:  $1$  is in  $T - Alg$  in general; if structure transports along equivalence, then  $Iso$  is also in  $T - Alg$ ; and the two functors  $F$  and  $G$  preserve all structure of  $1$ . So  $T - Alg$  is rarely complete and we seek an account of what limits such 2-categories do have in general.

A precise study of the limits available in  $T - Alg$  is fundamental to carrying out semantical studies of logical frameworks in this setting. For instance, in her thesis [Br], Carolyn Brown gives an appropriate definition of a map of Petri nets, then proves that the category of Petri nets is a model of propositional linear logic. The way she does this is first to prove that any category with the equational structure of Example 2.9 is a model of propositional linear logic. Then she proves that the category of Petri nets has such structure: in order to prove this, she uses a limiting property of the category of  $T$ -algebras to give a particularly elegant proof ([Br] Section 5.1). So it is fundamental to have a general, precise analysis of the limiting properties of  $T - Alg$ : and in the absence of completeness in general, we must state precisely what limits do exist.

A useful class of such limits is the class of *pie-limits* [PR]. Pie-limits may be described as follows. *Products* in 2-categories are as one would expect, subject to a mild extra property called the “two-dimensional property” (see [K2]). An *serter* between parallel maps  $f, g : A \rightarrow B$  is the universal pair  $(h, \alpha)$  consisting of a map  $h : I \rightarrow A$  and a 2-cell  $\alpha : fh \Rightarrow gh$ . An *equifier* of parallel 2-cells  $\alpha, \beta : f \Rightarrow g : A \rightarrow B$  is the universal map  $k : E \rightarrow A$  such that  $\alpha k = \beta k$ . A pie-limit is any limit that can be formed from products, inserters, and equifiers. A pie-limit is finite if only finite products need be used. A precise account and characterization appear in [PR] Thm 2.2. Observe that using inserters and equifiers, one may approximate the definition of an equalizer, replacing the equality of parallel pairs by an isomorphism between them. Specifically, an *iso-serter* between parallel maps  $f, g : A \rightarrow B$  is the universal pair  $(h, \alpha)$  consisting of a map  $h : I \rightarrow A$  and an invertible 2-cell  $\alpha : fh \Rightarrow gh$ . This iso-serter plays the role of an equalizer up to isomorphism. See [K2] for all the definitions and for examples.

The situation for colimits in  $T - Alg$  is more complex. Observe that  $T - Alg$  need not even have an initial object: in the case of cartesian categories, every cartesian category must have at least one object, the terminal object, so there are at least two finite product preserving functors from any cartesian category into  $Iso$ , namely the constant functors at each of the two objects of  $Iso$ . This leads us to *bicolimits* (see [St1] and [BKP] Thm 5.8), which are discussed in Section 5.

There are several other substantial results about categories with equational structure, [BKP] being the primary paper. In particular, there is an account of the problem of contravariance for categories with equational structure, and an account of that structure that transports along equivalence (see [KP2] and [Po7]). Those are delicate issues that took several years to resolve, and we shall see more of them in the next two sections.

## 4 Categories with essentially algebraic structure

Having unified examples 2.1 to 2.9, we seek to broaden this unity to incorporate the remaining examples. We look again to the study of sets with structure for inspiration. One step in generality above the study of sets with equational structure is the study of

sets with essentially algebraic structure, i.e., the study of models of a finite limit theory. Here, a theory is defined to be a small category  $Th$  with finite limits, and a model is a finite limit preserving functor from  $Th$  to  $Sets$ . The primary text in this area is [GU], and a version in Australian English is [K1], but I heartily recommend [MP].

The passage from equational theories, equivalently finitary monads, to finite limit theories allows us to define operations and equations not only universally but on equationally defined subobjects. For instance, in defining a category, one has a set  $O$  of objects, a set  $A$  of arrows, and functions  $dom, cod : A \rightarrow O$ ; but then one defines composition of arrows not on the set  $A \times A$  of all pairs of arrows but on the set of pairs  $(f, g)$  of arrows such that  $cod(f) = dom(g)$ . So the theory of categories is not equational with respect to the category of sets, but it is a finite limit theory, in fact a primary example of a finite limit theory. All the examples of categories with structure in Section 2 are models of finite limit theories but with the wrong maps: as in Section 3, we will introduce 2-cells to account for the right maps.

The category of models  $Mod(Th)$  of a finite limit theory has several useful properties. For instance, the inclusion of  $Mod(Th)$  into the functor category  $[Th, Sets]$  has a left adjoint. It follows that  $Mod(Th)$  is complete and cocomplete, with limits given pointwise. Although it is easy to prove directly that  $Mod(Th)$  is complete, cocompleteness is not so obvious: in particular, colimits are not given pointwise in general. Further, for any finite limit preserving functor from  $Th$  to  $Th'$ , the induced functor from  $Mod(Th')$  to  $Mod(Th)$  has a left adjoint. Indeed, such induced functors can be characterized as those that have left adjoints and preserve a certain class of colimits. Also useful is the fact that any colimit preserving functor from  $Mod(Th)$  into a cocomplete category necessarily has a right adjoint. Categories of the form  $Mod(Th)$  can be characterized as being complete categories subject to a delicate size condition (see [K1] Thm 9.8). There is not a good logical characterization of these categories although there are characterizations of variants such as those corresponding to Horn clauses.

An important notion for the practical application of these results is that of *sketch*. A sketch  $S$  consists of a small category  $C$  together with a finite collection of cones in  $C$ . One may define a *model* of a sketch in a category with finite limits  $L$  to be a functor from  $C$  to  $L$  that sends each of the cones to a limiting cone in  $L$ ; a *map of models* is a natural transformation between two such functors. Let  $Mod(S)$  denote the category of models of the sketch  $S$  in the category of sets. Then, there is a finite limit theory  $Th_S$  and a model  $\eta$  of  $S$  in  $Th_S$  such that composition with  $\eta$  yields an equivalence from  $Mod(Th_S)$  to  $Mod(S)$ . This is the usual way that one obtains a finite limit theory. For instance, in the study of categories, one generally just studies the sketch for a category and rarely requires an explicit description of the corresponding finite limit theory.

An attractive feature of finite limit theories is that one can speak of models not merely in  $Sets$ , but more generally in any category with finite limits. For instance, a monoid in  $Cat$  is a finite limit preserving functor from the theory of a monoid into  $Cat$ , and this is precisely a strict monoidal category. Similarly, a group in the category of topological spaces is precisely a topological group. An interesting example for us is that for any small category  $B$ , a category in the functor category  $[B^{op}, Sets]$  is, up to equivalence, a fibration with base  $B$ : as usual, the maps in  $Cat([B^{op}, Sets])$  are wrong.

A first attempt to generalise this to a study of categories with structure may be to define a *strict 2-theory* as a small 2-category  $Th$  with finite limits (see [K2] for definitions)

and a *model* as a finite limit preserving 2-functor from  $Th$  to  $Cat$ , the 2-category of small categories, functors, and natural transformations. A *strict map of models* would be a 2-natural transformation.

However, this does not incorporate all our examples. In particular, it does not account for contravariance or nonfunctoriality. Take for example the case of cartesian closed categories. In defining a cartesian closed category  $C$ , one needs to define the exponential of two objects. This does not form a functor from  $C \times C$  to  $C$ , but rather a functor from  $C^{op} \times C$  to  $C$ . But  $C^{op} \times C$  is not an allowable domain of definition for us. A further problem arises for elementary toposes, because the function giving the characteristic map of a subobject is not functorial.

Nevertheless, this framework can account for fibrations with certain extra structure, such as all the structures of example 2.12, and we can account for the covariant part of the structure of all the other examples. This includes all those examples from 2.1 to 2.9 for which the monad on  $Cat_0$  extends not merely to a 2-monad on  $Cat_g$  (see Section 3), but to a 2-monad on  $Cat$ , the 2-category of small categories, functors, and natural transformations. (Such an extension, if it exists, is necessarily unique (see [KP3]).) For example, cartesian categories with finite sums, categories with a natural numbers object, symmetric monoidal categories, and categories with finite limits and colimits are the algebras for 2-monads on  $Cat$ . The last are central to Makkai and Pare's work [MP].

A concrete example of how this works is the case of binary products, as in Example 2.4. This is an equational theory in which all the structure is covariant, so fits into this framework. Let  $F$  denote the category depicted by  $Y \leftarrow X \rightarrow Z$ . Let  $F'$  denote  $F$  together with a freely adjoined initial object. Then to give a category with binary products is to give a category  $C$  together with functors  $lim : C^2 \rightarrow C^{F'}$  and  $comparison : C^{F'} \rightarrow C^{F'}$  subject to the commutativity of some derived diagrams in  $Cat$  to force  $lim$  to yield a cone over each pair of objects and  $comparison$  to yield the unique comparison map. Unicity of the comparison map follows by functoriality. Details appear in [KP3], and [DK] has many similar calculations. From this description, it is routine to derive a strict 2-theory with the desired models.

For an example that seems not to be equational, so is not covered by the analysis of Section 3, consider a fibration. A fibration consists of a functor  $p : E \rightarrow B$  subject to the condition that for each object  $X$  of  $E$  and each arrow of the form  $f : A \rightarrow pX$  in  $B$ , there is an arrow with codomain  $X$  lying over  $f$  that satisfies a universal property. This seems not to be equational because  $f$  has codomain  $pX$ , so the operation we have described is not universally defined on objects or arrows or diagrams of a particular shape in a base category. However, it is routine to complete this outline to derive a strict 2-theory for fibrations.

The absence of an account of contravariance and nonfunctoriality is an important omission, but the issues have not been fully investigated yet. As mentioned in Section 3, these issues arose for categories with equational structure, but they were resolved satisfactorily there by use of  $Cat_g$ : in order to get the right maps there, we enriched over  $Cat_g$ , and although much structure of interest to us did not enrich over  $Cat$ , all enriched over  $Cat_g$ , so we had a satisfactory account of the right maps for all our structure. Here, we have enrichment over  $Cat$  embedded into our very definitions, so the analysis of contravariance and nonfunctoriality is inherently more delicate, and these issues remain the subject of current research in denotational semantics (see [F]). We leave them aside

for now in the hope and expectation that they will be readily resolved for essentially algebraic structure: later difficulties are more worrying.

A second difficulty with this approach, as in Section 3, is that the maps are the wrong maps. This is resolvable, essentially as in Section 3, by replacing 2-natural transformations as maps of models by pseudo-natural transformations (see [K2] Section 5): these generalise 2-natural transformations in that one replaces the commuting square in the definition of 2-naturality by a coherent isomorphism, i.e., an isomorphism satisfying two equations corresponding to those in the definition of a functor.

This works quite well. We can prove several results along the lines of those for ordinary finite limit theories as follows. A *biadjoint*, defined in [K2] Section 6, is the most common two-dimensional generalisation of adjoint, generalising all equalities in the definition of adjunction to coherent isomorphisms. Given a strict 2-theory  $Th$ , if we denote the 2-category of finite limit preserving 2-functors into  $Cat$  and pseudo-natural transformations by  $Mod(Th)$ , we have an inclusion of  $Mod(Th)$  into  $Ps(Th, Cat)$ , the 2-category of all 2-functors and pseudo-natural transformations. We can prove the following.

- The inclusion of  $Mod(Th)$  into  $Ps(Th, Cat)$  has a left biadjoint.
- $Mod(Th)$  has all pie-limits, given pointwise.
- $Mod(Th)$  has all bicolimits.
- Every finite limit preserving 2-functor from  $Th$  to  $Th'$  induces a left biadjoint to the induced 2-functor from  $Mod(Th')$  to  $Mod(Th)$ .

These results generalise several of the results concerning equational structure in [BKP]: we outlined the situation for limits in Section 3. In fact, perhaps surprisingly, one can use the results we have developed for categories with equational structure to deduce corresponding results here [KP4]. The reason is that for any small 2-category  $Th$ , there is a 2-monad  $T$  on  $[ob(Th), Cat]$  for which  $T - Alg$  is  $Ps(Th, Cat)$ . The strict maps of algebras are the 2-natural transformations (see [BKP] Section 6.6). It is routine to verify that  $Mod(Th)$  is closed under pie-limits in  $Ps(Th, Cat)$ , and these two facts give us enough to deduce the above results from those for equational structure.

A third problem arises when one seeks to generalise the notion of sketch to 2-sketch and generalise the result that every sketch induces a finite limit theory with the same models. There is an evident definition of 2-sketch, which is a small 2-category  $C$  with a finite collection of generalised cones (see [K2] and [PR] and the example of categories with binary products above). There are also evident definitions of a model of a 2-sketch and a map of models. The latter is a pseudo-natural transformation between two 2-functors from  $C$  to  $Cat$ . Let us consider a 2-sketch with no generalised cones, i.e., a small 2-category  $C$ . Then,  $Mod(C)$  is just  $Ps(C, Cat)$ . We should like the inclusion  $\eta$  of  $C$  into the free 2-category with finite limits,  $Th_C$ , on  $C$  to induce an equivalence (or at least a biequivalence, to be defined in Section 6) from  $Mod(Th_C)$  to  $Mod(C)$ : but it does not.

For a counterexample, let  $T$  be the 2-category with two objects  $X$  and  $Y$ , two maps  $f, g': X \rightarrow Y$ , and no nontrivial 2-cells. Define  $H : T \rightarrow Cat$  to be the constant 2-functor at 1, and define  $K : T \rightarrow Cat$  by the two functors from 1 to  $Iso$ . There are two pseudo-natural transformations from  $H$  to  $K$ . Neither extends to  $Th_C$  because the equalizer in

$Th_C$  of  $f$  and  $g$  is sent by the extension of  $H$  to 1 and by the extension of  $K$  to  $\phi$ , the empty category.

This problem, though not the other two, may be resolved by replacing “all” finite limits in the definitions of strict 2-theory and model by a restricted class of them. Two classes that have been proposed are *flexible* limits [BKPS] and pie-limits [PR]. They are almost identical, and since we have already seen pie-limits in Section 3, we remain with them. We make the following definitions.

- A 2-theory is a small 2-category  $Th$  with finite pie-limits.
- A model is a 2-functor from  $Th$  to  $Cat$  that preserves finite pie-limits.
- A map of models is a pseudo-natural transformation.

These definitions admit the results for strict 2-theories outlined above and more, detailed in [KP4], namely

- The inclusion of  $Mod(Th)$  into  $Ps(Th, Cat)$  has a left biadjoint.
- $Mod(Th)$  has all pie-limits, given pointwise.
- $Mod(Th)$  has all bicolimits.
- Every finite pie-limit preserving 2-functor from  $Th$  to  $Th'$  induces a left biadjoint to the induced 2-functor from  $Mod(Th')$  to  $Mod(Th)$ .
- The inclusion  $\eta$  of a small 2-category  $C$  into its free finite pie-limit completion  $Th_C$  induces an equivalence from  $Mod(Th_C)$  to  $Mod(C)$ .

A proof of the final result rests upon the fact that to give two 2-functors from  $C$  into  $Cat$  together with a pseudo-natural transformation between them is to give a 2-functor from  $C$  into  $Ps(T, Cat)$ , where  $T$  is defined as above; and to give two finite pie-limit preserving 2-functors from a 2-category with finite pie-limits  $D$  into  $Cat$  together with a pseudo-natural transformation between them is to give a finite pie-limit preserving 2-functor from  $D$  to  $Ps(T, Cat)$ . Now use the freeness of  $Th_C$ . This does not generalise to all finite limits because  $Ps(T, Cat)$  does not have pointwise equalizers, as our counter-example showed.

Our pie-limit definitions still incorporate all finitary 2-monads on  $Cat$ , and include all the above mentioned examples. I should mention that this uses the definition of one-sided fibration of [St1]. This is less common than that used in most texts, but agrees with the more common definition up to equivalence: it has coherent isomorphisms where the more common definitions have equalities. We also use a definition of map of fibrations in the same spirit: this is a mild generalisation of the more common definition, but each of our maps is isomorphic to one in the more usual sense.

Moreover, since many 2-categories of small categories with structure, such as all those of the form  $\mathcal{T} - Alg$ , have all finite pie-limits but not all finite limits, we may now speak of models of a 2-theory in any such 2-category. In particular, the 2-category  $Fib/B$  of fibrations over a base category  $B$  has pie-limits, so we can speak of fibrations with structure determined by pie-limits. Recall that a fibration over  $B$  is, up to equivalence, a category in  $[B^{op}, Sets]$ . Bénabou stresses the importance of regarding fibrations as

generalised categories [B2], and that view is implicit in much work on fibrations or indexed categories, such as [H], [J], [P] and [St2]. At least some of that work may be expressed in terms of models of a 2-theory in  $Fib/B$ .

There is an alternative to the use of 2-categories with finite pie-limits. It involves bicategories with finite bilimits. We outline that approach and discuss its relative merits in the latter part of the next section. Categorically, it is generally regarded as more natural. Moreover, a priori, it seems more general. However, its complexity overwhelms any attempt at calculation with it, and the main results we outline in the rest of the paper show that its extra generality is illusory, that it is in fact equipotent with pie-limits. That is why we have stressed the simpler theory of pie-limits here.

## 5 Transport of structure along equivalence

As explained in Section 3, one reason for requiring an analysis of structure preserving functors rather than functors that strictly preserve structure is to study the transport of structure along equivalence. As remarked there, given a logic  $L$  with encoding  $\Sigma_L$ , we should like to study structure transporting along an equivalence from  $C(L)$  to  $C(\Sigma_L)$ . Another reason for studying such structure is that it often enables us to move to a more simply presented example. This allows us to reduce the level of detail we need to consider in computations through the use of a coherence theorem. For instance, every fibration is equivalent to a split fibration, which is a simpler structure. So it is convenient to be able, without loss of generality, to restrict attention from fibrations with structure to split fibrations with structure. This is only possible if the structure under consideration transports along equivalence. For a third example, recall from Section 3 that the category of monoids is equivalent to the category  $FP(M, Sets)$ . So if we are to deduce results about the category of monoids from a general theory about categories of the form  $FP(Th, Sets)$ , we need an account of that structure that transports along equivalence. Not all structure is of this kind, even with the maps we now have: an example follows.

The central difficulty in the transport of such structure is the distinction between equality and isomorphism. For instance, consider the following category with equational structure: a category  $C$  together with an endofunctor  $H : C \rightarrow C$  such that  $H^2 = 1$ . Suppose  $C$  is equivalent to  $D$  with the equivalence given by  $F \dashv U : C \rightarrow D$ . Then, in general, it is not true that  $(UHF)^2 = 1 : D \rightarrow D$ . It is true that there is an isomorphism between  $(UHF)^2$  and 1. So it seems, and in fact there exist precise theorems (see [KP2] and [Po7]) to the effect, that structure transports if and almost only if it involves no equations between objects. The idea of no equations between objects is true of all the structure of all the examples of Section 2, albeit sometimes subtly.

Indeed, this is regarded as such a fundamental idea in the development of categories with structure that a preponderance of work in the area, such as unpublished work of Michael Makkai, Street's work [St1], [St2], and implicitly Johnstone's work [Joh] (see also Bénabou's [B2] for the centrality in which this question is placed), has been directed solely towards such structure, and this has led to the preponderance of work being based not on our definitions of 2-theory and model thereof but the following: a *bicategorical theory* is a bicategory with finite bilimits; a *model* is a pseudo-functor into  $Cat$  that preserves finite bilimits; a *map of models* is a pseudo-natural transformation: n.b., any 2-category with

finite pie-limits can be regarded as a bicategory with finite bilimits. Precise definitions appear in [B1] and [St1], but the idea is simple: replace all equalities in the definitions of “2-category with finite limits” and “2-functor that preserves finite limits” by coherent isomorphisms.

As the following examples show, the details are ghastly.

**Example 5.1** Consider a bipullback in a bicategory: it is a pullback with all equalities, even for composition of arrows, replaced by coherent isomorphisms. Given  $B \rightarrow A \leftarrow C$ , one requires an object  $lim$  together with maps from  $lim$  to each of  $A$ ,  $B$  and  $C$ , plus two isomorphisms. To express the one-dimensional part of the limiting property requires at least nine isomorphisms: given an object  $X$ , maps from  $X$  to each of  $A$ ,  $B$  and  $C$ , and two isomorphisms, one requires a map from  $X$  to  $lim$ , together with three isomorphisms, one each to account for the maps from  $X$  into  $A$ ,  $B$  and  $C$ ; and because it is in a bicategory, one needs two isomorphisms to account for associativity of the maps that lead from  $X$  via  $lim$  and  $A$  to each of  $B$  and  $C$ . This data is subject to two equations to the effect that the isomorphisms compose well with respect to the maps out of  $A$ . Moreover, to account for the two-dimensional property requires at least eighteen isomorphisms and four other 2-cells, all subject to numerous equations.

**Example 5.2** Suppose  $A$  and  $B$  are 2-categories with inserters. Consider what it means for a pseudo-functor  $H : A \rightarrow B$  to preserve inserters regarded as finite bilimits. A pseudo-functor generalises a 2-functor in that it preserves composition and identities only up to isomorphism, subject to three coherence equations corresponding to the equations in the definition of a category. Given  $f, g : X \rightarrow Y$ , let  $(h : I \rightarrow X, \alpha : fh \Rightarrow gh)$  be an inserter of  $(f, g)$  and let  $(k : J \rightarrow Y, \beta : (Hf)k \Rightarrow (Hg)k)$  be an inserter of  $(Hf, Hg)$ . Now  $Hf.Hh$  is isomorphic to  $H(fh)$ , and  $H\alpha$  is a 2-cell from  $H(fh)$  into  $H(gh)$ , which is in turn isomorphic to  $Hg.Hh$ . By the universal property of an inserter, this yields a unique map  $c : HI \rightarrow J$ , satisfying two equations, the one giving coherence with respect to  $\alpha$  and  $\beta$  involving four 2-cells. For our pseudo-functor  $H$  to preserve the inserter as a finite bilimit, we require that  $c$  be an equivalence. Now suppose  $c^*$  is an equivalence inverse of  $c$ . In general,  $c^*$  need not commute with  $Hh$  and  $k$ , but need only commute up to coherent isomorphism. So for instance, if  $B$  was  $Cat$ , then  $HI$  need not be isomorphic to the inserter in  $Cat$ , but only equivalent to it, and the equivalence inverse need not satisfy the equations that one so freely uses.

A general theory based upon such definitions as those of bilimit and bilimit preserving pseudo-functor, although natural, is unpleasantly complex. We should like to take these natural but complex definitions and prove that we may subsume them into our simpler setting of Section 4. That has two attractions. First, it yields results in this setting as immediate corollaries of easy results in the setting of pie-limits (see [KP4]). Second, it is a central completeness result for our pie-limit setting, as it ensures we incorporate any natural example that may arise in these terms. In fact, all our examples of the pie-limit setting may equally be seen as examples of the bilimit setting, which of course makes a result reducing the latter to the former particularly pressing.

In the second part of the paper, which constitutes the remaining sections, we show that it is possible to reduce from bilimits to the simpler setting of pie-limits. In doing so, we exhibit the technical category theoretic concepts, such as tricategories, required for the proofs.

## 6 Coherence for bicategories with finite bilimits

We now revert to a more usual style of presentation in order to show how to prove the results we sought at the end of the last section, and to show specifically how concepts such as tricategories arise in the proof.

**Notation 6.1** For any small bicategory  $B$  with finite bilimits, we write  $FBil(B, Cat)$  for the 2-category of finite bilimit preserving pseudo-functors into  $Cat$ .

**Definition 6.2** A pseudo-functor  $H : B \rightarrow C$  is a *biequivalence* if every object  $Y$  of  $C$  is equivalent to an object of the form  $HX$ , if every map of the form  $g : HX \rightarrow HX'$  is isomorphic to one of the form  $Hf$  for some  $f : X \rightarrow X'$ , and if every 2-cell of the form  $\beta : Hf \Rightarrow Hf' : HX \rightarrow HX'$  is uniquely of the form  $H\alpha$  for  $\alpha : f \Rightarrow f' : X \rightarrow X'$ .

Our central result is

**Theorem 6.3** For every small bicategory  $B$  with finite bilimits, there exists a small 2-category  $Th$  with finite pie-limits such that  $FBil(B, Cat)$  is biequivalent to  $FPie(Th, Cat)$ , the 2-category of models of the 2-theory  $Th$ .

The result can be proved in several parts. First, we prove the following.

**Theorem 6.4** Every small bicategory with finite bilimits is biequivalent to a small 2-category with finite pie-limits.

**Proof:** A detailed proof appears in [Po2] Thm 4.1. Given  $B$ , consider the Yoneda embedding of  $B$  into  $Hom(B^{op}, Cat)$ , the 2-category of pseudo-functors and pseudo-natural transformations. On homs, it is an equivalence. Moreover, the Yoneda embedding preserves finite bilimits, and  $Hom(B^{op}, Cat)$  has all pie-limits. So the closure of  $B$  in  $Hom(B^{op}, Cat)$  provides the 2-category we seek.

**Corollary 6.5** For any small bicategory  $B$  with finite bilimits, there is a small 2-category  $Th$  with finite pie-limits such that  $FBil(B, Cat)$  is biequivalent to  $FBil(Th, Cat)$ .

It is well-known (for instance, see [Po1] Thm 3.4 and Example 4.2) that for any small 2-category  $Th$ , every pseudo-functor from  $Th$  to  $Cat$  is equivalent to a 2-functor from  $Th$  to  $Cat$ . One may hope even more: that every 2-functor from  $Th$  to  $Cat$  that preserves finite bilimits may be equivalent to one that preserves finite pie-limits. If that were true, we could immediately deduce Theorem 6.3. However, it is not so easy.

**Example 6.6** Suppose a 2-category  $Th$  has a terminal object. Then  $Th$  is trivially a bicategory with a biterminal object. A 2-functor  $H : Th \rightarrow Cat$  preserves that biterminal object if and only if the unique map from  $H1$  to the unit category  $1$  is an equivalence. However, a terminal object is also an example of a pie-limit, and a 2-functor  $K : Th \rightarrow Cat$  preserves the terminal object as a pie-limit if and only if the unique map from  $K1$  to  $1$  is an isomorphism. There exists a 2-category  $Th$  with a terminal object together with a 2-functor  $H : Th \rightarrow Cat$  that preserves it as a bilimit but for which there is no 2-functor  $K$  equivalent to  $H$  such that  $K$  preserves the terminal object as a pie-limit (see [Po5]).

Although this is not strictly a counterexample to our hope, it is clear that our hope is doomed. This is a setback, but it is not fatal.

**Theorem 6.7** For any small 2-category  $Th$  with finite pie-limits, there is another small 2-category  $Th'$  with finite pie-limits, necessarily biequivalent to  $Th$ , such that to give a finite bilimit preserving 2-functor from  $Th$  to  $Cat$  is to give a finite pie-limit preserving 2-functor from  $Th'$  to  $Cat$ .

This theorem, together with Corollary 6.5 and the remark after it, suffices to prove our main result, Theorem 6.3. To build  $Th'$  is delicate, as replacing an equivalence by an isomorphism at one point in  $Th$  may automatically replace an isomorphism by an equivalence somewhere else. So the only proof I have of its existence involves the proof of a more general result, then this deduction as one of several corollaries. For the proof, we need tricategories.

## 7 Why tricategories?

In Section 3, we considered finitary 2-monads on  $Cat_0$ . There exists a 2-monad  $T$  on  $Cat_0$  for which the algebras are small categories with finite limits; arrows in  $T_0 - Alg$  are functors that strictly preserve finite limits; and the maps of primary interest are functors that preserve finite limits (see Example 2.4). It is a consequence of Theorem 3.13 of [BKP] that for every small category  $Th$  with finite limits, there is a small category  $Th'$  with finite limits such that  $FL(Th, Cat)$  is equivalent to  $FL_*(Th', Cat)$ . We seek to emulate that result here, replacing  $Cat$  by the 3-category  $2 - Cat$  of small 2-categories, and replacing all finite limits by all finite pie-limits.

In short, this strategy works, but only after modification. Specifically, the 2-cells in  $2 - Cat$  are inadequate: they are 2-natural transformations, whereas we need pseudo-natural transformations, as illustrated by the following example.

**Example 7.1** Let  $A$  and  $B$  be 2-categories with binary products. Let  $H : A \rightarrow B$  be a 2-functor. Then the limiting property of binary products in  $B$  ensures that for every pair of objects  $(X, Y)$  in  $A$ , there is a unique comparison map  $\gamma_{X,Y} : H(X \times Y) \rightarrow H(X) \times H(Y)$  satisfying two coherence equations. It follows that  $\gamma$  is a 2-natural transformation between two 2-functors from  $A \times A$  to  $B$ , one given by taking the binary product in  $A$  then applying  $H$ , the other given by applying  $H \times H$  then taking the binary product in  $B$ . Now suppose that  $H$  preserves binary biproducts. That is equivalent to saying that for each pair  $(X, Y)$ , the comparison map  $\gamma_{X,Y}$  is an equivalence. Suppose an equivalence inverse is given by  $\gamma_{X,Y}^*$ . Alas,  $\gamma^*$  need not be 2-natural: it need only be pseudo-natural. In fact, it is routine to verify that  $H$  preserves binary biproducts if and only if the 2-natural transformation  $\gamma$  is an equivalence in the 2-category  $Ps(A \times A, B)$  of 2-functors from  $A \times A$  to  $B$  and pseudo-natural transformations.

The observation in the example that an equivalence inverse to  $\gamma$  is necessarily pseudo-natural, hence an equivalence in  $Ps(A \times A, B)$ , applies to all pie-limits; it is central to expressing finite bilimit preserving 2-functors elegantly as generalised maps of algebras.

In passing to the extra generality of pseudo-natural transformations, we no longer have a 3-category. In fact, small 2-categories, 2-functors and pseudo-natural transformations do

not even form a 2-category: the interchange law, asserting the commutativity of horizontal and vertical composition, fails. An example of the failure is the following.

**Example 7.2** Let  $\alpha : H \Rightarrow K : A \rightarrow B$  be a pseudo-natural transformation, and let  $f : X \rightarrow Y$  be any map in  $A$ . Then let  $\bar{f} : \bar{X} \Rightarrow \bar{Y} : 1 \rightarrow A$  be the pseudo-natural transformation determined by  $f$ , i.e.,  $\bar{f} = f$ . By the interchange law, it would follow that  $\alpha$  is 2-natural with respect to  $f$ . This holds for all  $f$ . So  $\alpha$  is necessarily 2-natural, a contradiction.

Small 2-categories, 2-functors and pseudo-natural transformations naturally form a tricategory. Fortunately, they form a special sort of tricategory called a *Gray*-category, which has better properties than an arbitrary tricategory. A *Gray*-category is a category enriched over the symmetric monoidal closed category *Gray* which is defined as follows.

**Definition 7.3** (see [GPS] Def 4.8) *Gray* is the symmetric monoidal closed category whose underlying category is the category  $2 - Cat_0$  of small 2-categories and 2-functors, and with exponential given by  $Ps(A, B)$ .

It is a routine to verify that *Gray* is indeed symmetric monoidal closed as claimed. Details appear in [GPS] Section 4.

In this setting, we can deduce Theorem 6.7 and hence, as a corollary, Theorem 6.3 (see [Po5] and [Po6]).

**Proof:** (of Theorem 6.7) The proof is essentially a routine generalisation of that of Theorem 3.13 of [BKP]. First observe that the category of small 2-categories with finite pie-limits and 2-functors that strictly preserve finite pie-limits is monadic over  $2 - Cat_0$ . Then define the notion of a 2-functor preserving monadic structure not necessarily strictly, in terms of the monad and its enrichment over *Gray*; and verify that this agrees, in our specific case, with the definition of finite bilimit preserving 2-functor. Denote the category of algebras for a *Gray*-monad  $T$ , with 2-functors that preserve the structure, by  $T - Alg$ , and the subcategory determined by those 2-functors that strictly preserve the structure by  $T - Alg_s$ . Prove, as in [BKP] Thm 3.13, that the inclusion has a left adjoint  $(\ )'$  and that for each object  $A$  of  $T - Alg$ ,  $A$  is biequivalent to  $A'$ . Make an easy adjustment to account for the fact that the 2-category *Cat* is not small, and we are done.

Just as the proof of Theorem 3.13 of [BKP] required two-dimensional limits, our proof requires three-dimensional limits; we need further three-dimensional limits in proving that our specific example is in fact an example of that setting. Also, of course, we need to compose 2-cells and 3-cells in the course of the proof, leading to several natural questions.

First, in the definition of a 3-category, there are three sorts of composition, and for simplicity of exposition, we should prefer a single unifying definition. That led to a pasting result [Po4] Thm 6.16, better suited to our purposes than the known one [Jo].

Also, we asked whether every tricategory is triequivalent to a 3-category. If so, we could for some purposes replace our tricategory by a 3-category, giving us a simpler exposition. In fact, that is not the case: a one-object one-arrow tricategory is a braided monoidal category; a one-object one-arrow 3-category is a strict monoidal category with a strict symmetry on it; and it is well known that there exists even a symmetric monoidal category that is not equivalent to any monoidal category with a strict symmetry. An example is the following.

**Example 7.4** Consider the category *Sets* with finite product. It is cartesian closed, hence symmetric monoidal closed. The symmetry is given by the switch map from  $X \times Y$  to  $Y \times X$ . In the special case that  $X = Y$ , the switch map is not the identity if  $X$  has more than one object. Any equivalence of categories must preserve identity maps and be fully faithful. Hence, *Sets* is not equivalent to a monoidal category with a strict symmetry.

The best positive result we have is as follows.

**Theorem 7.5** Every tricategory is triequivalent to a *Gray*-category.

**Proof:** A proof appears in [GPS] Thm 8.1. Essentially, given a tricategory  $T$ , use the Yoneda embedding to embed  $T$  into  $Tricat(T^{op}, Bicat)$ , the tricategory of trihomomorphisms from  $T^{op}$  into  $Bicat$ , where each of the definitions is the natural replacement of all equalities by isomorphisms in the definitions of 3-category, 3-functor, etcetera. One then proves that  $Bicat$  is triequivalent to a *Gray*-category, and deduces the result. The proof of [GPS] Thm 8.1 is a modified version of this, as a few tricks can eliminate some detail.

This result, although not directly used for our theorem, helped to give the definitive notion of map between 2-categories with equational structure [Po6]: a one object tricategory is a monoidal 2-category, an example of a 2-category with equational structure; a trihomomorphism between such gives precisely the right data and axioms for a general map of algebras. We require that general definition of map of algebras for our example.

An unexpected but pleasant consequence of the analysis at this level of generality is that it seems to shed light on several issues of Section 4. For instance, regarding the contravariance question, it seems that the notion of a 2-category with an involution, a model being a 2-functor that sends the involution to  $( )^{op}$ , may be insightful for an account of contravariance. A 2-category with an involution is another example of a 2-category with equational structure. Also, regarding our problem with a sketch lifting to a strict 2-theory, our counterexample may be expressed essentially as saying that the left adjoint to the evident functor from the category of small 2-categories with finite limits into  $2 - Cat_0$  does not enrich to a *Gray*-adjunction between the *Gray*-categories determined by adding pseudo-natural transformations. Our claim that this problem does not occur for pie-limits is essentially the assertion that the corresponding left adjoint for 2-categories with finite pie-limits does enrich.

## 8 Bibliography

- [A] S.J.Amblar, First order linear logic in symmetric monoidal closed categories, Ph.D. thesis, Edinburgh (1992) CST-87-92
- [AP] S.O.Anderson and A.J.Power, A representable approach to nondeterminism (submitted)
- [B1] J.Bénabou, Introduction to bicategories, in: Lecture Notes in Math. 47 (1967) 1-77
- [B2] J.Bénabou, Fibred categories and the foundations of naive category theory, J. Symb. Logic. 50 (1985) 10-37
- [BW] R.Betti and R.F.C.Walters, The calculus of ends over a base topos, J. Pure Appl. Algebra 56 (1989) 211-220

- [BKPS] G.J.Bird, G.M.Kelly, A.J.Power and R.H.Street, Flexible limits for 2-categories, *J. Pure Appl. Algebra* 61 (1989) 1-27
- [BKP] R.Blackwell, G.M.Kelly and A.J.Power, Two-dimensional monad theory, *J.Pure Appl. Algebra* 59 (1989) 1-41
- [Br] C.T.Brown, Linear logic and petri nets: categories, algebra and proof, Ph.D. thesis, Edinburgh (1991) CST-71-91
- [DK] E.J.Dubuc and G.M.Kelly, A presentation of topoi as algebraic relative to categories or graphs, *J. Algebra* 81 (1983) 420-433
- [FP] M.P.Fourman and W.K.Phoa, A proposed categorical semantics for pure *ML*, in: ICALP '92, *Lecture Notes in Computer Science* 623 (1992) 533-544
- [F] P.J.Freyd, Algebraically complete categories, in: *Category Theory Proceedings, Como 1990, Lecture Notes in Math.* 1488 (1991) 95-104
- [GU] P.Gabriel and F.Ulmer, Lokal präsentierbare Kategorien, in: *Lecture Notes in Math.* 221 (1971)
- [G] P.A.Gardner, Representing logics in type theory, Ph.D. thesis, Edinburgh (1992)
- [GPS] R.Gordon, A.J.Power and R.H.Street, Coherence for tricategories, *Memoir of the American Math Soc* (to appear)
- [HP1] James Harland and David Pym, The uniform proof-theoretic foundation of linear logic programming, in: *Proc. International Symp. on Logic Programming*, MIT Press (1991) 304-318
- [HP2] J.A.Harland and D.J.Pym, On resolution in fragments of classical linear logic, in: *Proc. LPAR '92, Lecture Notes in Artificial Intelligence* (to appear)
- [HHP] Robert Harper, Furio Honsell and Gordon Plotkin, A framework for defining logics, in: *Proc. Symp. on Logic in Computer Science* (1987) 194-204
- [H] Claudio Hermida, Fibrations, logical predicates, and indeterminates, Ph. D. thesis, Edinburgh (1993)
- [HP] J.M.E.Hyland and A.M.Pitts, The Theory of Constructions: categorical semantics and topos-theoretic Models, in: *Categories in Computer Science and Logic, Contemporary Math.* 92, AMS (1989) 137-199
- [J] Bart Jacobs, Categorical type theory, Ph.D. thesis, Nijmegen (1991)
- [Jo] Michael Johnson, The combinatorics of n-categorical pasting, *J. Pure Appl. Algebra* 62 (1989) 211-225
- [Joh] P.T.Johnstone, Fibrations and partial products in a 2-category, preprint (1992)
- [JS] A.Joyal and R.H.Street, Braided tensor categories, *Advances in Math.* (to appear)
- [K1] G.M.Kelly, Structures defined by finite limits in the enriched context I, *Cahiers de Top. et Geom. Diff.* 23 (1982) 3-41
- [K2] G.M.Kelly, Elementary observations on 2-categorical limits, *Bull. Austral. Math. Soc.* 39 (1989) 301-317
- [KP1] G.M.Kelly and A.J.Power, Adjunctions whose counits are coequalizers, and presentations of finitary enriched monads, *J. Pure Appl. Algebra* 89 (1993) 163-179
- [KP2] G.M.Kelly and A.J.Power, Flexibility for a monad, preprint (1988)
- [KP3] G.M.Kelly and A.J.Power, Enrichability of algebraic structures on categories (in preparation)
- [KP4] G.M.Kelly and A.J.Power, Essentially algebraic structure on categories (in preparation)
- [LS] J.Lambek and P.J.Scott, *Introduction to higher-order categorical logic*, Cambridge

Studies in Advanced Mathematics 7 (1986)

[M] S. Mac Lane, *Categories for the working mathematician*, Graduate Texts in Mathematics 5, Springer (1971)

[MP] Michael Makkai and Robert Pare, *Accessible categories: the foundations of categorical model theory*, in: *Contemporary Mathematics* 104 (1989)

[P] D. Pavlović, *Predicates and fibrations*, Ph.D. thesis, Utrecht (1990)

[Po1] A.J. Power, *A general coherence result*, *J. Pure Appl. Algebra* 57 (1989) 165-173

[Po2] A.J. Power, *Coherence for bicategories with finite bilimits I*, in: *Categories in Computer Science and Logic*, *Contemporary Mathematics* 92 (1989) 165-173

[Po3] A.J. Power, *An algebraic formulation for data refinement*, in: *Proc. Conf. Mathematical Foundations of Programming Semantics*, *Lecture Notes in Comp. Sci.* 442 (1990) 390-401

[Po4] A.J. Power, *An n-categorical pasting theorem*, in: *Category Theory Proceedings, Como 1990*, *Lecture Notes in Math.* 1488 (1991) 326-358

[Po5] A.J. Power, *Coherence for bicategories with finite bilimits II* (in preparation)

[Po6] A.J. Power, *Coherence for bicategories with equational structure* (in preparation)

[Po7] A.J. Power, *A simple criterion for the flexibility of a monad* (in preparation)

[PR] John Power and Edmund Robinson, *A characterization of pie-limits*, *Math. Proc. Camb. Phil. Soc.* 110 (1991) 33-47

[PS] A.J. Power and Leon Sterling, *A notion of map between logic programs*, in: *Logic Programming, Proc. 7th International Conf.*, MIT Press (1990) 390-404

[RR] E. Robinson and G. Rosolini, *Categories of partial maps*, *Information and Computation* 79 (1988) 95-130

[S1] R.A.G. Seely, *Hyperdoctrines, natural deduction and the Beck condition*, *Zeitschr. f. math. Logik und Grundlagen d. Math. Bd.* 29 (1983) 505-542

[S2] R.A.G. Seely, *Locally cartesian categories and type theories*, *Math. Proc. Camb. Phil. Soc.* 95 (1984) 33-48

[St1] R.H. Street, *Fibrations in bicategories*, *Cahiers de Top. et Geom. Diff.* 21 (1980) 111-160

[St2] R.H. Street, *Conspectus of variable categories*, *J. Pure Appl. Algebra* 21 (1981) 307-338

[T] R.D. Tennent, *Semantics of Programming Languages*, Prentice-Hall (1991)