Abstract Diagrams

R. Banach
Computer Science Dept., Manchester University, Manchester, M13 9PL, U.K.
banach@cs.man.ac.uk

Abstract: The awkwardnes of 'up to isomorphism' diagrammatic constructions is recalled, and one repost, via skeleton categories and standard isomorphisms, is reviewed. An alternative approach is introduced, which defines abstract diagrams as natural isomorphism classes of concrete diagrams, and is related to the previous one. Maximal abstract diagrams yield canonical diagrammatic constructions, where only 'up to isomorphism' constructions were available previously.

1 Introduction

The fact that in conventional diagrammatic reasoning in category theory, most constructions only yield an answer 'up to isomorphism', is something that people have learnt to live with, rather than something that is held to be intrinsically good. The wealth of coherence results that are generated as a consequence of needing to reconcile the outputs of different instances of essentially the same construction, which differ only in the order in which some component operations are performed, is the tangible mathematical response to the phenomenon. Nevertheless despite these, the feeling that a neater handling of these matters would be nice, is hard to stifle. Aside from aesthetics, ambiguity up to isomorphism is more troublesome when the outputs of diagrammatic constructions are used as a semantic vehicle for some purpose. Cases in point arise when the entities in play are fundamentally graph theoretical (eg. [Rozenberg (1997), Ehrig et al. (1999)]), whereupon the ambiguity can make desired semantic manipulations problematic. This paper offers a new way to tackle these issues.

'Up to isomorphism' signals the necessity to exercise choice. We avoid this need by trading choice for closure, which is an old trick. By this means we come up with a notion of 'abstract diagram' in contrast to the conventional notion of 'concrete diagram'. Previous approaches have defined 'abstract diagrams' out of equivalence classes of objects and arrows of the category of interest, but this does not fully succeed due to the need to refer implicitly or explicitly to a skeleton of the underlying category and which therefore comes down to a choice again (of the skeleton). In our approach, an abstract diagram is a functor category of some or all of the relevant concrete diagrams. It turns out that these functor categories have properties that bear comparison to what is done with conventional diagrams. The price to be paid for this is that (depending on foundations) one has to deal with large categories almost immediately. One further consequence is that the framework of abstract diagrams does not obviate the need to reason at the concrete level, basically because the diagrams that figure in a conventional derivation do not fully encode the derivation, but only some aspects of it. Abstract diagrams provide a way of displaying the results in a neater way; so there is no case of something for nothing, which is reassuring.

The strategy just outlined differs from most of the existing approaches for dealing with choice in categorical constructions. These do not so much seek to avoid it as look for ways of making it more canonically. Largely, one feels, this is prompted by the desire to avoid troubling foundational issues. Our alternative approach is based on purely functorial constructions, and of all constructions that might be considered foundationally suspect, purely functorial ones are the least suspect of all. Accordingly, our foundational basis is to tacitly employ universes, i.e. formally 'everything is a "set"'. (Despite this, we are unable to resist using the phraseology of 'equivalence classes' and 'isomorphism classes', since 'equivalence set' and 'isomorphism set' just sound wrong.)

The rest of this paper is as follows. Section 2 revisits the familiar problems with factoring a category through isomorphisms in the context of the category of graphs, and reviews the familiar solution via standard isomorphisms. Section 3 recalls the definition of the category of diagrams, and extrapolates this to a construction of abstract diagrams using functor categories, yielding a category of abstract diagrams $\mathit{ADiag}(\mathsf{C})$, and showing that maximal abstract diagrams are unique. Section 4 notes that $\mathit{ADiag}(\mathsf{C})$ has too many morphisms, and constructs a fresh category of abstract diagrams that has fewer of them, $\boldsymbol{\mathit{MADiag}}(\mathsf{C})$. These two sections are the key sections of the paper. Section 5 relates the functor category approach to abstract diagrams, to the work using standard isomorphisms, and may be mostly skipped on a first reading. Section 6 shows how diagrammatic constructions can be recast in a canonical way using abstract diagrams: they become inclusion morphisms in $\boldsymbol{\mathit{MADiag}}(\mathsf{C})$. The treatment of products is described in detail, and some other examples are noted briefly. Section 7 briefly outlines some prospective uses of abstract diagrams in semantic theory. Section 8 discusses the relationship of our theory to earlier work, particularly to Makkai's anafunctors [Makkai (1996)], and revisits some foundational issues. Section 9 concludes, after which there is an appendix which reviews feeble functors which are used in Section 5.

2 The Abstraction Problem

In this section we briefly cover some necessary technical preliminaries on graphs, and introduce the abstraction problem for diagrams which motivates the constructions in the main part of the paper.

Definition 2.1 A graph G is a tuple (E, V, s, t) where E and V are sets of edges and vertices, and $s, t : E \to V$ are two set functions that send each edge to its source and target vertex respectively. A graph morphism $f : G \to G'$ is a pair of maps $f_E : E \to E'$, $f_V : V \to V'$ such that $f_V \circ s = s' \circ f_E$ and $f_V \circ t = t' \circ f_E$. This gives us the category Gr of (concrete) graphs and morphisms with obvious identities and composition of morphisms. We will also refer to Gr by the name S; using the name Gr when we are interested in isomorphism classes of various kinds, and using S when we need individual shape graphs in the discussion of diagrams. The convention of having both S and Gr in principle also yields the opportunity to make the two categories of different sizes in a foundationally different treatment.

In Cat, the functor U that forgets arrow composition, yields from a category C, the underlying graph $U\mathsf{C}$ of C.

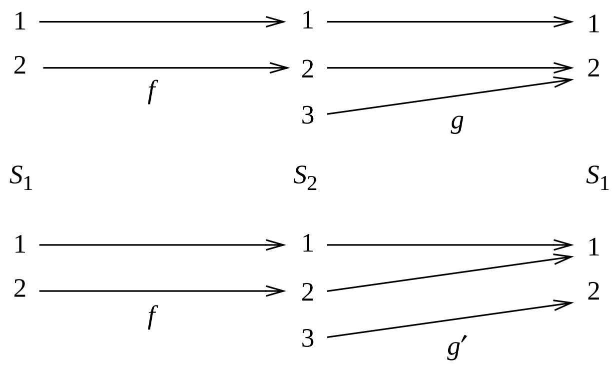


Fig. 1 A counterexample to simple equivalence classes up to isomorphism.

One thing that has been contemplated for *Gr* (and for similar categories whose objects generally admit nontrivial automorphisms), is to raise the level of abstraction from individual graphs and graph morphisms to more abstract notions, particularly since such categories seldom admit canonical choices for representatives of the isomorphism classes of their objects. This is the abstraction problem, and an obvious strategy which suggests itself is to form isomorphism classes of graphs and of morphisms and to proceed from there. Unfortunately this is easier said than done. A familiar example in *Set* (which we can regard as a category of discrete graphs and thus a subcategory of *Gr*), illustrates the problem.

Example 2.2 Let $S_1 = \{1, 2\}$ and $S_2 = \{1, 2, 3\}$. Consider the maps $f : S_1 \to S_2$ and g, $g' : S_2 \to S_1$ of Fig. 1. In a naive construction of abstract sets and abstract maps between them, the abstract set containg a set S would be all sets equipolent to S, and the abstract map containing a map $s : S_1 \to S_2$ would be the collection of all maps $t : T_1 \to T_2$ such that there are isomorphisms $j_1 : S_1 \to T_1$ and $j_2 : S_2 \to T_2$ such that $s = j_2^{-1} \circ t \circ j_1$. In Fig. 1 we claim that g and g' would be in the same isomorphism class because if we take j_1 as the map $\{1 \mapsto 3, 2 \mapsto 1, 3 \mapsto 2\}$ and take j_2 as the map $\{1 \mapsto 2, 2 \mapsto 1\}$ then indeed $g = j_2^{-1} \circ g' \circ j_1$. Now the composition of two abstract maps would be the abstract map containing at least all composites of respective individual maps which are directly composable. So in the example, $g \circ f$ and $g' \circ f$ would be in the same abstract map. However, $g \circ f$ is monic while $g' \circ f$ is not, so this is impossible because monicity is invariant under isomorphism.

The reason why we get this unpleasant phenomenon is clear. When we form the composite, we have ‘forgotten’ that we have to relate g and g' by j_1 and j_2 in this particular instance, because the formation of equivalence classes does not remember this information. The technique of standard graphs and isomorphisms addresses this problem.

Definition 2.3 A choice of standard isomorphisms in *Gr* assigns to each pair of isomorphic graphs G_1 and G_2, a standard isomorphism $\sigma(G_1, G_2)$ such that:

(1) $\sigma(G,G) = \mathrm{id}_G$

(2) $\sigma(G_2,G_3) \circ \sigma(G_1,G_2) = \sigma(G_1,G_3)$

(3) $\sigma(G_2,G_1) = \sigma(G_1,G_2)^{-1}$

If we disallow all isomorphisms other than standard ones, the problems of Example 2.2 disappear because j_1 and j_2 are not standard by (1) above; hence g and g' fall into different equivalence classes.

Definition 2.4 We can construct a choice of standard isomorphisms in Gr as follows:

(1) We choose one graph $\sigma(G)$ from each isomorphism class $[G]$ of graphs isomorphic to G to be standard.

(2) For each G' in $[G]$, we choose one isomorphism $\sigma(\sigma(G),G')$ to be standard (with $\sigma(\sigma(G),G')$ chosen to be $\mathrm{id}_{\sigma(G)}$ if $G' = \sigma(G)$).

(3) For all G_1, G_2 in $[G]$, we set $\sigma(G_1,G_2) = \sigma(\sigma(G),G_2) \circ \sigma(\sigma(G),G_1)^{-1}$.

For the sequel we assume fixed some choice of standard isomorphisms in Gr. The collection of standard graphs and all morphisms between them forms a skeleton category Gr^{K} of Gr. It is not too hard to see that Gr^{K} is isomorphic to the category $\langle \mathsf{Gr} \rangle$, whose objects are isomorphism classes of concrete graphs up to standard isomorphism called (in this approach) abstract graphs and written $\langle G \rangle$, and whose arrows are equivalence classes of concrete morphisms under the relation that relates $g : G \to H$ and $g' : G' \to H'$ iff $g = \sigma(G',H')^{-1} \circ g' \circ \sigma(G,H)$, called abstract morphisms and written $\langle g : G \to H \rangle$. The use of only standard isomorphisms in this relation means that there is a bijection between concrete arrows $g : G \to H$ in $\langle g : G \to H \rangle$, and ordered pairs G, H taken from $\langle G \rangle$ and $\langle H \rangle$. Identities are the equivalence classes of concrete identities, and composition of arrows $\langle g : G \to H \rangle$ and $\langle h : H \to K \rangle$ is given by composing the concrete arrows in the two respective classes in the only possible way using the standard isomorphisms, which forms another equivalence class.

3 Concrete and Abstract Diagrams in an Arbitrary Category

In this section we abandon the approach of the last few paragraphs and embark on a fresh tack.

Definition 3.1 Let $\underline{\mu}$ be a graph, i.e. an object of S, let C be any category, and let $\underline{\gamma} : \underline{\mu} \to U\mathsf{C}$ be a graph morphism from $\underline{\mu}$ to the underlying graph of C. Then $\underline{\gamma}$ is a concrete prediagram of shape $\underline{\mu}$ in C. Let $\mathsf{Pth} : \mathsf{S} \to \mathsf{Cat}$ be the functor that sends graphs to their path categories, which is left adjoint to U. Then the standard free construction extends $\underline{\gamma} : \underline{\mu} \to U\mathsf{C}$ to a functor $\gamma : \mu \to \mathsf{C}$ from the path category μ of $\underline{\mu}$ to C. We call γ a plain concrete diagram of shape μ in C.

Thus far plain concrete diagrams do not have to commute. Let $\gamma : \mu \to \mathsf{C}$ be a plain concrete diagram, and suppose that for two paths $(e^1{}_1, \ldots, e^1{}_k)$ and $(e^2{}_1, \ldots, e^2{}_l)$ in μ, the internal compositions $(\gamma(e^1{}_k) \circ \ldots \circ \gamma(e^1{}_1))$ and $(\gamma(e^2{}_l) \circ \ldots \circ \gamma(e^2{}_1))$ yield the same arrow $f : \gamma(m_0) \to \gamma(m_1)$ in C. Then we say that the two paths commute in γ. Usually $(e^1{}_1, \ldots, e^1{}_k)$ and $(e^2{}_1, \ldots, e^2{}_l)$ have a common starting point m_0, and a common endpoint m_1, in μ, but this is not strictly necessary.

Note that the extent to which a diagram needs to commute depends on the use to which it is being put. Take as an example, an equaliser diagram: the two parallel arrows contemplated at the start don't need to be equal; however when when they are prepended with the equaliser arrow, the compositions must be equal.

Definition 3.2 Let $\gamma : \mu \to C$ be a plain concrete diagram of shape μ in C, and let $\theta_{\mu,\gamma}$ be a set of pairs of paths in μ. If for all pairs in $\theta_{\mu,\gamma}$, the internal compositions under γ yield equal arrows of C, then we say γ is $\theta_{\mu,\gamma}$-commuting. We write $\theta_{\mu,\gamma}$-commuting diagrams using the notation $(\gamma : \mu \to C,\ \theta_{\mu,\gamma})$, and refer to them as a concrete diagrams.

Definition 3.3 We write $[\mu, C]$ for the functor category whose objects are plain concrete diagrams of shape μ, and whose arrows are plain concrete diagram morphisms, i.e. natural transformations $n : \gamma \to \delta$. We write $[\mu, C]_\theta$ for the enriched category whose objects are concrete diagrams of shape μ, and whose arrows are commutativity nondecreasing natural transformations, i.e. natural transformations $n : \gamma \to \delta$ such that $\theta_{\mu,\gamma} \subseteq \theta_{\mu,\delta}$.

From now on all concrete diagrams not explicitly stated to be plain are assumed to have a commutativity specification $\theta_{\mu,\gamma}$ and to be $\theta_{\mu,\gamma}$-commuting. Since we do not want to be restricted to just one shape, we introduce $Diag(C)$ the category of concrete diagrams in C over arbitrary shapes.

Definition 3.4 In $Diag(C)$ the objects are concrete diagrams in C over arbitrary shapes $(\gamma : \mu \to C,\ \theta_{\mu,\gamma})$, and the arrows are pairs $(\xi, \alpha) : (\gamma : \mu \to C,\ \theta_{\mu,\gamma}) \to (\delta : \nu \to C,\ \theta_{\nu,\delta})$, such that $\alpha : \mu \to \nu$ is a change of shape, i.e. an arrow of $Pth(S)$, ξ is a natural transformation from $\gamma : \mu \to C$ to $\delta \circ \alpha : \mu \to C$, and $\theta_{\alpha(\mu),\gamma} \subseteq \theta_{\nu,\delta}$ holds, where $\theta_{\alpha(\mu),\gamma}$ is the image of the pairs of paths in $\theta_{\mu,\gamma}$ under α. Viewed another way, ξ is a collection of C arrows such that $\delta \circ \alpha = \xi \circ \gamma$ holds in the expected way. We say that the morphisms of $Diag(C)$ are mediated by the collections of C arrows ξ. The composition of (ξ, α) : $(\gamma : \mu \to C,\ \theta_{\mu,\gamma}) \to (\delta : \nu \to C,\ \theta_{\nu,\delta})$ and $(\zeta, \beta) : (\delta : \nu \to C,\ \theta_{\nu,\delta}) \to (\varepsilon : \lambda \to C,\ \theta_{\lambda,\varepsilon})$ is $(\alpha(\zeta) \circ \xi, \beta \circ \alpha) : (\gamma : \mu \to C,\ \theta_{\mu,\gamma}) \to (\varepsilon : \lambda \to C,\ \theta_{\lambda,\varepsilon})$, where $\alpha(\zeta)$ is the action of α on ζ, and $\theta_{\beta \circ \alpha(\mu),\gamma} \subseteq \theta_{\lambda,\varepsilon}$ holds.

This lays the foundation for the ensuing definitions.

Definition 3.5 A plain abstract diagram D (of shape μ in C) is a connected subcategory of $[\mu, C]$ all of whose arrows are natural isomorphisms. We write $\iota_D : D \to [\mu, C]$ for the inclusion functor. An abstract diagram D (of shape μ in C) is a plain abstract diagram D together with a commutativity specification $\theta_{\mu,D}$ shared by all the concrete diagrams in D, i.e. it is a subcategory of $[\mu, C]_\theta$. For abstract diagrams we use the notation $(\iota_D : D \to [\mu, C],\ \theta_{\mu,D})$ to reveal the various components.

Definition 3.6 An abstract diagram D of shape μ is maximal iff it is nonempty and for every object $\gamma : \mu \to C$ of D and natural isomorphism $n : \gamma \to \delta$ in $[\mu, C]$, n is an arrow of D.

An abstract diagram D_0 is a subdiagram of an abstract diagram D_1 (both of shape μ) iff D_0 is a subcategory of the category D_1, (and $\theta_{\mu,D_0} = \theta_{\mu,D_1}$). Trivial to prove, but of key importance is the following.

Proposition 3.7 Every abstract diagram is a subdiagram of a unique maximal abstract diagram (MAD).

As a consequence, the more concepts we can reformulate in terms of abstract diagrams, the more results are liable to come out uniquely, rather than 'up to isomorphism'.

Definition 3.8 A morphism $c : \mathit{D}_0 \to \mathit{D}_1$ of abstract diagrams (of shape μ in C) is a functor from the underlying plain abstract diagram D_0 to the underlying plain abstract diagram D_1 (where both D_0 and D_1 are considered simply as categories in their own right), such that $\theta_{\mu,\mathit{D}_0} \subseteq \theta_{\mu,\mathit{D}_1}$ holds. This gives rise to the category of abstract diagrams of shape μ in C, denoted $\mathit{Abs}(\mu,\mathit{C})$. If $c : \mathit{D}_0 \to \mathit{D}_1$ arises as a natural transformation between the inclusion functors $(\iota_{\mathit{D}_0} : \mathit{D}_0 \to [\mu,\mathit{C}], \theta_{\mu,\mathit{D}_0})$ and $(\iota_{\mathit{D}_1} : \mathit{D}_1 \to [\mu,\mathit{C}], \theta_{\mu,\mathit{D}_1})$ we say that c is mediated by the family of arrows of the natural transformation, which we denote by Ξ_c.

Given a mediated morphism $c : \mathit{D}_0 \to \mathit{D}_1$ as above, let $\Xi_c(\gamma)$ denote the element of Ξ_c at concrete diagram γ, where γ is an object of D_0. Since $\gamma : \mu \to \mathit{C}$ is itself a functor, $\Xi_c(\gamma)$ is itself a natural transformation, i.e. a family of arrows $\Xi_c(\gamma)(m)$ in C, one for each object m of μ (or vertex m of $\underline{\mu}$). Thus while an arbitrary morphism of abstract diagrams merely associates concrete diagrams and morphisms between them in a natural manner, a mediated morphism of abstract diagrams must be sensitive to any internal structure of objects captured by the structure of C.

For an example let C be Gr. Then an arbitrary morphism $c : \mathit{D}_0 \to \mathit{D}_1$ of abstract diagrams associates concrete diagrams $\gamma : \mu \to \mathit{Gr}$ and natural isomorphisms $n_0 : \gamma \to \gamma'$ in D_0 with concrete diagrams $\delta : \mu \to \mathit{Gr}$ and natural isomorphisms $n_1 : \delta \to \delta'$ in D_1, such that $c(\mathrm{id}_\gamma) = \mathrm{id}_{c(\gamma)}$, and $c(n_0 : \gamma \to \gamma') = c(n_0) : c(\gamma) \to c(\gamma')$, and compositions of them behave well, (and $\theta_{\mu,\mathit{D}_0} \subseteq \theta_{\mu,\mathit{D}_1}$). So each concrete graph $G_0 = \gamma(m)$ occurring at a vertex m of shape μ in diagram γ in D_0, is mapped to $G_1 = c(\gamma)(m)$ in D_1, and each concrete graph morphism $E_0 : G_0 \to G_0{}' = \gamma(e) : \gamma(m) \to \gamma(m')$ above edge $e : m \to m'$ of μ in D_0, is mapped to a corresponding concrete graph morphism $E_1 : G_1 \to G_1{}' = c(E_0) : c(G_0) \to c(G_0{}') = c(\gamma)(e) : c(\gamma)(m) \to c(\gamma)(m')$ above the same edge $e : m \to m'$ of $c(\mu)$ in D_1. And the mapping of the natural isomorphisms of D_0 under c respects this additional structure.

However if c is mediated, not only does all this have to hold, but each association of G_0 at m of $\gamma(\mu)$ in D_0 with G_1 at m of $c(\gamma)(\mu)$ in D_1 arises via an actual concrete graph morphism $\Xi_c(\gamma)(m) = f_{\gamma,m} : G_0 \to G_1$, such that these $f_{\gamma,m}$ preserve all the other structures.

Just as we wanted to change shape with concrete diagrams, we also want to do so with abstract ones.

Definition 3.9 In $\mathit{ADiag}(\mathit{C})$ the objects are abstract diagrams in C over arbitrary shapes eg. $(\iota_{\mathit{D}_0} : \mathit{D}_0 \to [\mu,\mathit{C}], \theta_{\mu,\mathit{D}_0})$, each with its own $\theta_{\mu,\mathit{D}_0}$, and the arrows are pairs $(c,\alpha) : (\iota_{\mathit{D}_0} : \mathit{D}_0 \to [\mu,\mathit{C}], \theta_{\mu,\mathit{D}_0}) \to (\iota_{\mathit{D}_1} : \mathit{D}_1 \to [\nu,\mathit{C}], \theta_{\nu,\mathit{D}_1})$, such that $\alpha : \mu \to \nu$ is a change of shape, $c : \mathit{D}_0 \to \alpha(\mathit{D}_1)$ is a functor (though not necessarily a natural transformation) between the inclusion functors $\iota_{\mathit{D}_0} : \mathit{D}_0 \to [\mu,\mathit{C}]$ and $\iota_{\alpha(\mathit{D}_1)} : \alpha(\mathit{D}_1) \to [\nu \circ \alpha,\mathit{C}]$, and $\theta_{\alpha(\mu),\mathit{D}_0} \subseteq \theta_{\nu,\mathit{D}_1}$. Here $\alpha(\mathit{D}_1)$ is the plain abstract diagram obtained by precomposing each plain concrete diagram $\delta : \nu \to \mathit{C}$ in D_1 with α. Moreover if c arises as a natural transformation from ι_{D_0} to $\iota_{\alpha(\mathit{D}_1)}$, then it is mediated by a family of arrows, denoted by Ξ_c, each arrow being itself a family of individual C arrows. The composition of the two arrows $(c,\alpha) : (\iota_{\mathit{D}_0} : \mathit{D}_0 \to [\mu,\mathit{C}], \theta_{\mu,\mathit{D}_0}) \to (\iota_{\mathit{D}_1} : \mathit{D}_1 \to [\nu,\mathit{C}], \theta_{\nu,\mathit{D}_1})$ and

$(d,\beta) : (\iota_{D_1} : \mathcal{D}_1 \to [\nu,\mathcal{C}],\ \theta_{\nu,D_1}) \to (\iota_{D_2} : \mathcal{D}_2 \to [\lambda,\mathcal{C}],\ \theta_{\lambda,D_2})$ is the arrow given by the data $(\alpha(d) \circ c, \beta \circ \alpha) : (\iota_{D_0} : \mathcal{D}_0 \to [\mu,\mathcal{C}],\ \theta_{\mu,D_0}) \to (\iota_{D_2} : \mathcal{D}_2 \to [\lambda,\mathcal{C}],\ \theta_{\lambda,D_2})$, where $\alpha(d)$ is the action of d on $\alpha(\mathcal{D}_1)$, and $\theta_{\beta\circ\alpha(\mu),D_0} \subseteq \theta_{\lambda,D_1}$ holds. If both (c,α) and (d,β) are mediated, then the composition is mediated by $\Xi_{\alpha(d)} \circ \Xi_c$, in a notation whose interpretation should be obvious.

For the remainder of this paper we will focus exclusively on mediated morphisms since these cover all the cases that arise in practice.

4 Morphism Overload and the Category *MADiag*($\mathcal{C}$)

The category *ADiag*($\mathcal{C}$) has the kind of objects we want, but its arrows are far too fine-grained. If $(c,\alpha) : \mathcal{D}_0 \to \mathcal{D}_1$ is an arrow, it is sufficient to change the value of the functor c at just one object γ of $\mathcal{D}_0$ to get a different arrow. The next section explores this in detail when relating the main thread of the paper to the standard isomorphism approach of Section 2. We want a less sensitive notion of arrow between abstract diagrams, which we manufacture via the following route.

Definition 4.1 Let $\underline{\alpha} : \underline{\mu} \to \underline{\nu}$ be a shape graph morphism. Let $\underline{\eta}_{\underline{\alpha}}$ be the shape graph constructed as follows:

Vertices: $V_{0,\underline{\mu}} \cup V_{1,\underline{\nu}}$ where:
$V_{0,\underline{\mu}} = \{(m,0) \mid m \in V_{\underline{\mu}}\}$
$V_{1,\underline{\nu}} = \{(m,1) \mid m \in V_{\underline{\nu}}\}$

Edges: $E_{0,\underline{\mu}} \cup E_{1,\underline{\nu}} \cup E_{01,\underline{\mu}}$ where:
$E_{0,\underline{\mu}} = \{(e,0) : (m,0) \to (m',0) \mid e : m \to m' \in E_{\underline{\mu}}\}$
$E_{1,\underline{\nu}} = \{(e,1) : (m,1) \to (m',1) \mid e : m \to m' \in E_{\underline{\nu}}\}$
$E_{01,\underline{\mu}} = \{(m,m',01) : (m,0) \to (m',1) \mid m \in V_{\underline{\mu}},\ m' = \underline{\alpha}_{V_{\underline{\mu}}}(m)\}$

There are obvous injections $\underline{\sigma} : \underline{\mu} \to \underline{\eta}$ and $\underline{\tau} : \underline{\nu} \to \underline{\eta}$. We say that $\underline{\eta}_{\underline{\alpha}}$ is an arrow-shape from $\underline{\mu}$ to $\underline{\nu}$ that represents $\underline{\alpha}$. (N.B. This representation is imperfect since the edge map of $\underline{\alpha}$ is not representated yet.)

Definition 4.2 Let $\mathcal{S}_E$ be the category whose objects are those of $\mathcal{S}$ and whose arrows are pairs $(\underline{\eta}_{\underline{\alpha}}, \underline{\alpha}_E) : \underline{\mu} \to \underline{\nu}$ where $\underline{\eta}_{\underline{\alpha}}$ is the arrow-shape of a morphism $\underline{\alpha} : \underline{\mu} \to \underline{\nu}$ and $\underline{\alpha}_E$ is the edge map component of $\underline{\alpha}$. The identities are $(\underline{\eta}_{\text{id}}, \text{id}_E) : \underline{\mu} \to \underline{\mu}$, where $\underline{\eta}_{\text{id}}$ consists of two copies of $\underline{\mu}$ joined by the obvious family of edges, and the composition of $(\underline{\eta}_{\underline{\alpha}}, \underline{\alpha}_E) : \underline{\mu} \to \underline{\nu}$ and $(\underline{\eta}_{\underline{\beta}}, \underline{\beta}_E) : \underline{\nu} \to \underline{\lambda}$ is given by $(\underline{\eta}_{\underline{\beta}\circ\underline{\alpha}}, \underline{\beta}_E \circ \underline{\alpha}_E) : \underline{\mu} \to \underline{\lambda}$.

Since the composition of $(\underline{\eta}_{\underline{\alpha}}, \underline{\alpha}_E) : \underline{\mu} \to \underline{\nu}$ and $(\underline{\eta}_{\underline{\beta}}, \underline{\beta}_E) : \underline{\nu} \to \underline{\lambda}$ in $\mathcal{S}_E$ is defined in terms of the composition of the underlying shape graph morphisms α and β, it clearly associates on the nose, because the composition of α and β does so. This definition is equivalent to a more convoluted one, that decomposes $\underline{\eta}_{\underline{\alpha}}$ and $\underline{\eta}_{\underline{\beta}}$ into their constituent parts, and assembles $\underline{\eta}_{\underline{\beta}\circ\underline{\alpha}}$ out of them directly (see the proof of Proposition 4.2). Similar observations hold for the constructions in the subsequent definitions.

Further, we note that the only nontrivial information supplied by $\underline{\alpha}_E$ in $(\underline{\eta}_{\underline{\alpha}}, \underline{\alpha}_E)$ is the disambiguation of how edges in $\underline{\sigma}(E_{\underline{\mu}})$ are to be mapped if there are parallel edges in the range of $\underline{\alpha}_E$ in $\underline{\tau}(E_{\underline{\nu}})$, in order to faithfully represent $\underline{\alpha}$. In many simple cases of course, there are no parallel edges in the range of $\underline{\alpha}$ in $\underline{\tau}(E_{\underline{\nu}})$ and $\underline{\alpha}_E$ is redundant.

Proposition 4.3 There is an isomorphism between the categories S and S_E.

Proof. The objects are identical so the only issue is to consider the arrows. From S to S_E the construction of $(\underline{\eta}_\alpha, \underline{\alpha}_E)$ from $\underline{\alpha} : \underline{\mu} \to \underline{\nu}$ gives a unique result. Conversely, given $(\underline{\eta}_\alpha, \underline{\alpha}_E) : \underline{\mu} \to \underline{\nu}$, the vertices of $\underline{\eta}_\alpha$ are tagged 0 or 1, its edges are tagged 0 or 1 or are triples $(m, m', 01)$ for 0-tagged m and 1-tagged m'; the $(m, m', 01)$ triples yield the graph of a total function on the 0-tagged vertices, and $\underline{\alpha}_E$ respects this function. From this information, the shape graphs $\underline{\mu}$ and $\underline{\nu}$ are easy to recover, as is the unique shape graph morphism $\underline{\alpha} : \underline{\mu} \to \underline{\nu}$. It is clear that these maps extend to functors in the right way. ☺

The preceding constructions lift to Pth-arrow-shapes η_α and corresponding path categories. Moreover Proposition 4.3 now allows us to rework some of the previous section in a different way.

Definition 4.4 The category $\boldsymbol{\mathsf{Diag}}(\mathsf{C})$ is given by the following data. Its objects are concrete diagrams in C over arbitrary shapes $(\gamma : \mu \to \mathsf{C}, \theta_{\mu,\gamma})$. Its arrows arise as follows. Let $(\xi, \alpha) : (\gamma : \mu \to \mathsf{C}, \theta_{\mu,\gamma}) \to (\delta : \nu \to \mathsf{C}, \theta_{\nu,\delta})$ be an arrow of $\mathsf{Diag}(\mathsf{C})$. Let $\underline{\eta}_\alpha$ be the arrow-shape that represents $\underline{\alpha}$, with injections $\underline{\sigma} : \underline{\mu} \to \underline{\eta}$ and $\underline{\tau} : \underline{\nu} \to \underline{\eta}$ and edge map $\underline{\alpha}_E$. Define the arrow-diagram $(\xi_\alpha : \eta_\alpha \to \mathsf{C}, \theta_{\eta_\alpha,\xi_\alpha})$ as the concrete diagram in C which agrees with γ on $\sigma(\mu)$, with δ on $\sigma(\nu)$, and with ξ on $E_{01,\underline{\mu}}$. In more detail:

$$\xi_{\alpha V}((m,0)) = \gamma_V(m) \quad \text{for } m \in V_\mu$$
$$\xi_{\alpha E}((e,0)) = \gamma_E(e) \quad \text{for } e \in E_\mu$$

$$\xi_{\alpha V}((m,1)) = \delta_V(m) \quad \text{for } m \in V_\nu$$
$$\xi_{\alpha E}((e,1)) = \delta_E(e) \quad \text{for } e \in E_\nu$$

$$\xi_{\alpha E}((m,\alpha(m),01)) = \xi(m) : \gamma_V(m) \to \delta_V(\alpha(m)) \quad \text{for } m \in V_\mu$$

The commutativity specification $\theta_{\eta_\alpha,\xi_\alpha}$ is equal to $\theta_{\sigma(\mu),\xi_\alpha} \cup \theta_{\tau(\nu),\xi_\alpha} \cup \theta_{\alpha_E}$, which is the image of $\theta_{\mu,\gamma}$ under σ, together with the image of $\theta_{\nu,\delta}$ under τ, together with θ_{α_E} which represents the pairs of commuting paths that arise in η_α from the edge map α_E and the requirement for ξ to be a natural transformation:

$$((m',\alpha(m'), 01) : (m',0) \to (\alpha(m'),1)) \circ ((e,0) : (m,0) \to (m',0)) =$$
$$((\alpha_E(e),1) : (\alpha(m),1) \to (\alpha(m'),1)) \circ ((m,\alpha(m), 01) : (m,0) \to (\alpha(m),1))$$

Each such arrow-diagram $(\xi_\alpha : \eta_\alpha \to \mathsf{C}, \theta_{\eta_\alpha,\xi_\alpha})$ is an arrow of $\boldsymbol{\mathsf{Diag}}(\mathsf{C})$ with domain and codomain $(\gamma : \mu \to \mathsf{C}, \theta_{\mu,\gamma})$ and $(\delta : \nu \to \mathsf{C}, \theta_{\nu,\delta})$. Identities are $(\xi_{\mathrm{id}} : \eta_{\mathrm{id}} \to \mathsf{C}, \theta_{\eta_{\mathrm{id}},\xi_{\mathrm{id}}})$ where $\theta_{\eta_{\mathrm{id}},\xi_{\mathrm{id}}} = \theta_{\sigma(\mu),\xi_{\mathrm{id}}} \cup \theta_{\tau(\mu),\xi_{\mathrm{id}}} \cup \theta_{\mathrm{id}_E}$. And the composition of $(\xi_\alpha : \eta_\alpha \to \mathsf{C}, \theta_{\eta_\alpha,\xi_\alpha})$ where $\theta_{\eta_\alpha,\xi_\alpha} = \theta_{\sigma(\mu),\xi_\alpha} \cup \theta_{\tau(\nu),\xi_\alpha} \cup \theta_{\alpha_E}$ and $(\zeta_\beta : \eta_\beta \to \mathsf{C}, \theta_{\eta_\beta,\zeta_\beta})$ where $\theta_{\eta_\beta,\zeta_\beta} = \theta_{\sigma(\nu),\zeta_\beta} \cup \theta_{\tau(\lambda),\zeta_\beta} \cup \theta_{\beta_E}$ is given by $((\alpha(\zeta) \circ \xi)_{\beta\circ\alpha} : \eta_{\beta\circ\alpha} \to \mathsf{C}, \theta_{\eta_{\beta\circ\alpha},(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}})$ where $\theta_{\eta_{\beta\circ\alpha},(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}} = \theta_{\sigma(\nu),(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}} \cup \theta_{\tau(\lambda),(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}} \cup \theta_{\beta_E\circ\alpha_E}$ and where $\theta_{\beta_E\circ\alpha_E}$ depicts the commuting paths:

$$((m',\beta\circ\alpha(m'),01) : (m',0) \to (\beta\circ\alpha(m'),1)) \circ ((e,0) : (m,0) \to (m',0)) =$$
$$((\beta_E \circ \alpha_E(e),1) : (\beta\circ\alpha(m),1) \to (\beta\circ\alpha(m'),1)) \circ$$
$$((m,\beta\circ\alpha(m),01) : (m,0) \to (\beta\circ\alpha(m),1))$$

Proposition 4.5 There is an isomorphism between the categories $\mathsf{Diag}(\mathsf{C})$ and $\boldsymbol{\mathsf{Diag}}(\mathsf{C})$.

Proof. The objects are identical so we turn to the arrows. From $\mathsf{Diag}(\mathsf{C})$ to $\boldsymbol{\mathsf{Diag}}(\mathsf{C})$ the construction described gives a unique result. Conversely, given an arrow $(\xi_\alpha : \eta_\alpha \to \mathsf{C},$

$\theta_{\eta_\alpha,\xi_\alpha}$) of ***Diag***($\mathcal{C}$), an arrow of $\mathit{Diag}(\mathcal{C})$ is easy to recover as follows. Firstly we extract the domain, codomain and vertex map of the shape graph morphism $\alpha : \mu \to \nu$ from the objects in η_α, and the arrows of length one in η_α, as in Proposition 4.3. Then $\theta_{\eta_\alpha,\xi_\alpha}$ is decomposed into the part contained in $\sigma(\mu)$, the part contained in $\tau(\nu)$, and the remainder θ_{α_E}, which connects $\sigma(\mu)$ and $\tau(\nu)$. The edge map α_E is now easy to recover from θ_{α_E} and α is obtained. The mediating arrows of the natural transformation ξ are now read off as the arrows of ξ_α above the $(m,m',01)$ edges of η_α. It is clear that these maps extend to functors in the right way. ☺

So far we have done little other than to acquire complexity, trading one picture of the category of concrete diagrams and their morphisms, for another perhaps slightly clumsier one. The payoff comes at the abstract level, when the various individual arrow-diagrams ($\xi_\alpha : \eta_\alpha \to \mathcal{C}$, $\theta_{\eta_\alpha,\xi_\alpha}$) of ***Diag***($\mathcal{C}$) which differ only by natural isomorphisms, are absorbed into a single maximally abstract diagram representing them all, overcoming thereby the proliferation of finegrained arrows of $\mathit{ADiag}(\mathcal{C})$.

Definition 4.6 The category ***MADiag***($\mathcal{C}$) is given by the following data. Its objects are maximal abstract diagrams in $\mathcal{C}$ over arbitrary shapes ($\iota_{D_0} : D_0 \to [\mu,\mathcal{C}]$, θ_{μ,D_0}). The arrows arise as follows. Let $\alpha : \mu \to \nu$ be a change of shape, let ($\iota_{D_0} : D_0 \to [\mu,\mathcal{C}]$, θ_{μ,D_0}) and ($\iota_{D_1} : D_1 \to [\nu,\mathcal{C}]$, θ_{ν,D_1}) be two objects of ***MADiag***($\mathcal{C}$), and let $\gamma : \mu \to \mathcal{C}$ and $\delta : \nu \to \mathcal{C}$ be plain concrete diagrams in D_0 and D_1. Let ($\xi_\alpha : \eta_\alpha \to \mathcal{C}$, $\theta_{\eta_\alpha,\xi_\alpha}$) be an arrow-diagram in ***Diag***($\mathcal{C}$) from ($\gamma : \mu \to \mathcal{C}$, $\theta_{\mu,\gamma}$) to ($\delta : \nu \to \mathcal{C}$, $\theta_{\nu,\delta}$), where $\theta_{\mu,\gamma} = \theta_{\mu,D_0}$ and $\theta_{\nu,\delta} = \theta_{\nu,D_1}$. Let ($\iota_{E_{\xi_\alpha}} : E_{\xi_\alpha} \to [\eta_\alpha,\mathcal{C}]$, $\theta_{\eta_\alpha,E_{\xi_\alpha}}$) be the maximal abstract diagtram that contains ($\xi_\alpha : \eta_\alpha \to \mathcal{C}$, $\theta_{\eta_\alpha,\xi_\alpha}$). Then ($\iota_{E_{\xi_\alpha}} : E_{\xi_\alpha} \to [\eta_\alpha,\mathcal{C}]$, $\theta_{\eta_\alpha,E_{\xi_\alpha}}$) is an abstract arrow-diagram and an arrow of ***MADiag***($\mathcal{C}$); moreover all ***MADiag***($\mathcal{C}$) arrows arise this way. We note that any two arrow-diagrams in ***Diag***($\mathcal{C}$) that are related by a natural isomorphism lead to the same abstract arrow-diagram of ***MADiag***($\mathcal{C}$). The identity arrows are ($\iota_{E_{\xi_{\mathrm{id}}}} : E_{\xi_{\mathrm{id}}} \to [\eta_{\mathrm{id}},\mathcal{C}]$, $\theta_{\eta_{\mathrm{id}},E_{\xi_{\mathrm{id}}}}$), the MADs containing the identity arrows of ***Diag***($\mathcal{C}$).

For the composition of ***MADiag***($\mathcal{C}$) arrows, let ($\iota_{E_{\xi_\alpha}} : E_{\xi_\alpha} \to [\eta_\alpha,\mathcal{C}]$, $\theta_{\eta_\alpha,\xi_\alpha}$) from ($\iota_{D_0} : D_0 \to [\mu,\mathcal{C}]$, θ_{μ,D_0}) to ($\iota_{D_1} : D_1 \to [\nu,\mathcal{C}]$, θ_{ν,D_1}), be one ***MADiag***($\mathcal{C}$) arrow, and let it contain the concrete ***Diag***($\mathcal{C}$) arrow-diagram ($\xi_\alpha : \eta_\alpha \to \mathcal{C}$, $\theta_{\eta_\alpha,\xi_\alpha}$) from ($\gamma : \mu \to \mathcal{C}$, $\theta_{\mu,\gamma}$) to ($\delta : \nu \to \mathcal{C}$, $\theta_{\nu,\delta}$), where $\theta_{\mu,\gamma} = \theta_{\mu,D_0}$ and $\theta_{\nu,\delta} = \theta_{\nu,D_1}$. Let ($\iota_{E_{\zeta_\beta}} : E_{\zeta_\beta} \to [\eta_\beta,\mathcal{C}]$, $\theta_{\eta_\beta,\zeta_\beta}$) from ($\iota_{D_1} : D_1 \to [\nu,\mathcal{C}]$, θ_{ν,D_1}) to ($\iota_{D_2} : D_2 \to [\lambda,\mathcal{C}]$, θ_{λ,D_2}) be a second ***MADiag***($\mathcal{C}$) arrow, containing the concrete ***Diag***($\mathcal{C}$) arrow-diagram ($\zeta_\beta : \eta_\beta \to \mathcal{C}$, $\theta_{\eta_\beta,\zeta_\beta}$) from ($\delta : \nu \to \mathcal{C}$, $\theta_{\nu,\delta}$) to ($\varepsilon : \lambda \to \mathcal{C}$, $\theta_{\lambda,\varepsilon}$), where $\theta_{\nu,\delta} = \theta_{\nu,D_1}$ and $\theta_{\lambda,\varepsilon} = \theta_{\lambda,D_2}$. Then the composition of the two ***MADiag***($\mathcal{C}$) arrows mentioned, is the maximal abstract diagram containing the composition of the two concrete ***Diag***($\mathcal{C}$) arrow-diagrams mentioned. Thus it is the abstract arrow-diagram containing ($(\alpha(\zeta) \circ \xi)_{\beta \circ \alpha} : \eta_{\beta \circ \alpha} \to \mathcal{C}$, $\theta_{\eta_{\beta\circ\alpha},(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}}$), and is denoted ($\iota_{E_{(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}}} : E_{(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}} \to [\eta_{\beta\circ\alpha},\mathcal{C}]$, $\theta_{\eta_{\beta\circ\alpha},E_{(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}}}$), where $\theta_{\eta_{\beta\circ\alpha},(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}} = \theta_{\eta_{\beta\circ\alpha},E_{(\alpha(\zeta)\circ\xi)_{\beta\circ\alpha}}}$.

Note that maximality guarantees that there is an abundant choice of concrete arrow-diagrams ξ_α and ζ_β that compose on the nose inside $\iota_{E_{\xi_\alpha}}$ and $\iota_{E_{\zeta_\beta}}$. We can always pick two arbitrary arrow-diagrams in $\iota_{E_{\xi_\alpha}}$ and $\iota_{E_{\zeta_\beta}}$, and can then apply a natural isomorphism to one or the other so that they compose properly, while remaining within $\iota_{E_{\xi_\alpha}}$ and $\iota_{E_{\zeta_\beta}}$ and producing the same composite arrow of ***MADiag***($\mathcal{C}$). The easy lemmas which show that this is so will be omitted.

Proposition 4.7 There is a functor from $\mathbf{Diag}(\mathcal{C})$ to $\mathbf{MADiag}(\mathcal{C})$.

Proof. It is sufficient to observe that the construction of Definition 4.6 is functorial. ☺

We now have our goal, which can be summarised in the following theorem.

Theorem 4.8 There is a functor $\mathbf{MAbs} : \mathit{Diag}(\mathcal{C}) \to \mathbf{MADiag}(\mathcal{C})$.

Proof. $\mathbf{MAbs}$ is given by composing the functor $\mathit{Diag}(\mathcal{C}) \to \mathbf{Diag}(\mathcal{C})$ from Proposition 4.4 with the functor $\mathbf{Diag}(\mathcal{C}) \to \mathbf{MADiag}(\mathcal{C})$ in Proposition 4.7. ☺

The truth of Theorem 4.8 should be contrasted with the fact that there is no naturally arising functorial route from $\mathit{Diag}(\mathcal{C})$ to $\mathbf{MADiag}(\mathcal{C})$ via $\mathit{ADiag}(\mathcal{C})$ due to the enormous nondeterminism in mapping concrete $\mathit{Diag}(\mathcal{C})$ arrows to $\mathit{ADiag}(\mathcal{C})$ arrows. This is despite the existence of a functor from $\mathit{ADiag}(\mathcal{C})$ to $\mathbf{MADiag}(\mathcal{C})$ that maps objects to their maximal closure, and maps arrows to the maximal abstract arrow-diagrams constructed by picking the value of an $\mathit{ADiag}(\mathcal{C})$ arrow at a concrete diagram of its domain and then proceeding as in Definition 4.4 and Definition 4.6, (since for a fixed $\mathit{ADiag}(\mathcal{C})$ arrow, all such choices will lead to the same $\mathbf{MADiag}(\mathcal{C})$ arrow).

Nevertheless it should not be thought that $\mathit{ADiag}(\mathcal{C})$ is a pointless diversion. Its construction via routine functorial reasoning and the existence of a functor from $\mathit{ADiag}(\mathcal{C})$ to $\mathbf{MADiag}(\mathcal{C})$ bolsters the defence against (especially) foundational assaults.

5 Automorphisms and Kinded Abstract Diagrams

In this section we examine the consequences of objects in $\mathcal{C}$ having nontrivial automorphisms, in order to relate the standard isomorphism approach to the concepts of Section 3. This entails examining $\mathit{ADiag}(\mathcal{C})$ in more detail. The part after Notation 5.3 may be skipped on a first reading.

We assume chosen a skeleton subcategory $\mathcal{C}^{K}$ of $\mathcal{C}$, leading to a choice of standard isomorphisms $\sigma(–,–)$ between objects. Also $\langle \mathcal{C} \rangle$ will be the category of abstract $\mathcal{C}$ objects and arrows, consisting of equivalence classes up to standard isomorphisms, of $\mathcal{C}$ objects and arrows.

Let *Kind* be $\{\mathsf{id} \le \mathsf{std} \le \mathsf{iso}\}$, partially ordered as shown. We will use *Kind* as a label set for shape vertices, thus for an abstract diagram of shape μ there will be a map, *kind*, from its vertices to *Kind*, and we will speak of shapes and vertices of kind such and such.

Definition 5.1 Let $(\iota_{D} : D \to [\mu,\mathcal{C}], \theta_{\mu,D})$ be a abstract diagram of kinded shape μ, then D conforms to its kind (i.e. is a kinded abstract diagrams (or KAD)) iff for each vertex m in μ:

(1) $\mathit{kind}(m) = \mathsf{id}$ iff for each natural isomorphism $n : \gamma \to \delta$ in D, the component of n at m is an identity in $\mathcal{C}$, i.e. $n(m) : \gamma(m) \to \delta(m) = \mathrm{id}_{\gamma(m)}$,

(2) $\mathit{kind}(m) = \mathsf{std}$ iff for each natural isomorphism $n : \gamma \to \delta$ in D, the component of n at m is a standard isomorphism in $\mathcal{C}$, i.e. $n(m) : \gamma(m) \to \delta(m) = \sigma(\gamma(m),\delta(m))$,

(3) $\mathit{kind}(m) = \mathsf{iso}$ iff for each natural isomorphism $n : \gamma \to \delta$ in D, the component of n at m is an arbitrary isomorphism in $\mathcal{C}$, i.e. $n(m) : \gamma(m) \to \delta(m)$ is an arbitrary iso.

Obviously every shape gives rise to a family of kinded shapes, kinded in all possible ways, and each giving rise to KADs of the appropriate kinds. The abstract diagrams of Sections 3 and 4 can be viewed as KADs with shapes entirely of kind iso.

Definition 5.2 A KAD $(\iota_D : \mathcal{D} \to [\mu, \mathcal{C}], \theta_{\mu,D})$ of shape μ is maximal iff for every object $\gamma : \mu \to \mathcal{C}$ of $\mathcal{D}$ and natural isomorphism $n : \gamma \to \delta$ in $[\mu, \mathcal{C}]$ of the appropriate kind, n is an arrow of $\mathcal{D}$.

Notation 5.3 In figures below, we will encode the kinds of the vertices of an abstract diagram by the following convention: unadorned vertices imply that the kind is id; vertices in angle brackets imply that the kind is std; and vertices in square brackets imply that the kind is iso. Thus $A \leftarrow \langle B \rangle \to [C]$ is a MAD (with various details suppressed) in which A occurs up to identity, B occurs up to standard isomorphisms, and C occurs up to arbitrary isomorphisms. To forestall possible confusion, concrete diagrams in the category $\langle \mathcal{C} \rangle$ will be indicated thus: $A^{\lozenge} \leftarrow B^{\lozenge} \to C^{\lozenge}$.

When we contemplate incorporating change of shape into the theory, we can both change the geometry, i.e. the underlying elements of $\mathcal{S}$, and the kinds of related vertices.

Definition 5.4 A change of shape morphism $\alpha : \mu \to \nu$ between kinded shapes μ and ν is strict at m in μ iff $kind_\mu(m) = kind_\nu(\alpha(m))$; it is lenient at m in μ iff $kind_\mu(m) \leq kind_\nu(\alpha(m))$; it is coercive at m in μ iff $kind_\mu(m) \geq kind_\nu(\alpha(m))$. It is strict, lenient, coercive, iff it is strict, lenient, coercive, at all m in μ. It is said to be general if it is not any of strict, lenient, coercive.

For each object $(\iota_D : \mathcal{D} \to [\mu, \mathcal{C}], \theta_{\mu,D})$ of $\mathcal{ADiag}(\mathcal{C})$ there is a least attribution of kinds to the vertices m of μ arising from an examination of the isomorphisms of the $\mathcal{C}$ objects above m in the concrete diagrams of $\mathcal{D}$. Any greater kinding of μ gives rise also to a valid KAD.

Definition 5.5 The category $\mathcal{ADiag}^k(\mathcal{C})$ has as objects all valid KADs and as arrows all $\mathcal{ADiag}(\mathcal{C})$ morphisms between them. (So a $\mathcal{ADiag}(\mathcal{C})$ morphism is a morphism between two KADs iff it is a morphism between the same objects with kinds forgotten).

Now if $kind_\mu : \mu \to Kind$ and $kind_\nu : \nu \to Kind$ are two kinded shapes, and $\alpha : \mu \to \nu$ is a change of shape, we can factor α through $\mu*$ where $\mu*$ is μ but with kind map given by $kind_{\mu*}(m) = kind_\mu(\alpha(m))$. So any kinded change of shape $\alpha : \mu \to \nu$ factors as $\alpha = \alpha* \circ \alpha_\mu$, where $\alpha_\mu : \mu \to \mu*$ is an identity on the shape but can change the kinds, and $\alpha* : \mu* \to \nu$ is a strict change of shape morphism. This enables us to separate study of change of geometry from change of the kind map.

Given this canonical factorisation, we now study the effects of change of the kind map with a fixed geometry μ. For good measure, we include the concrete diagrams with values in both $\mathcal{C}$ and for those with values in $\langle \mathcal{C} \rangle$, (the category of abstract $\mathcal{C}$ objects $\langle C \rangle$ and abstract $\mathcal{C}$ arrows $\langle f : C \to D \rangle$ up to standard isomorphism).

Fig. 2 provides a route map between various possibilities of interest. The left column depicts concrete diagrams, the middle column depicts abstract diagrams conveniently related to concrete ones, and the right column depicts unrestricted abstract diagrams. The top two rows show the situation for $\mathcal{C}$, and the bottom row shows the situation for the related category $\langle \mathcal{C} \rangle$ (given the latter, we do not bother separately with the situation

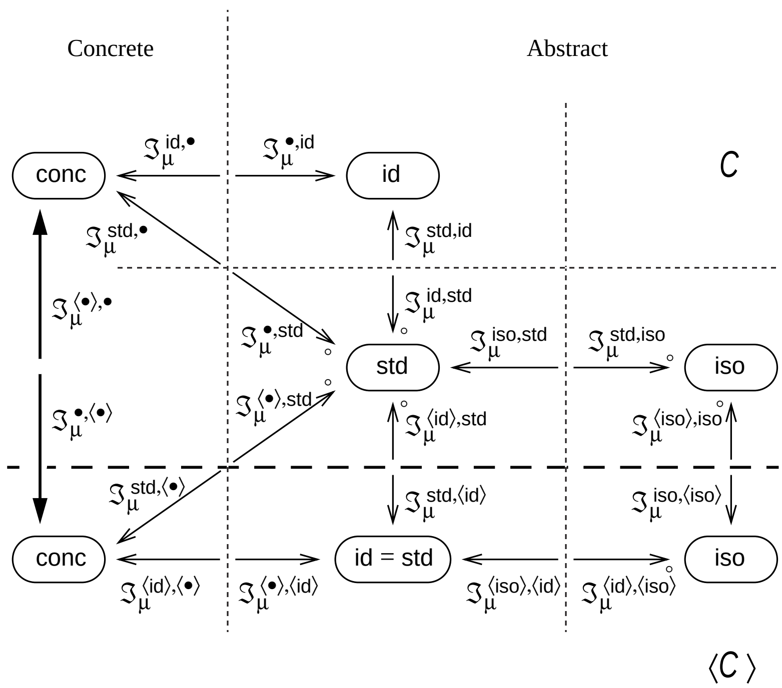


Fig. 2 A roadmap of concrete and kinded abstract diagrams.

for C^K). The rest of this section surveys fairly tersely the connections between the various possibilities that can be constructed canonically. These are described by (sometimes feeble) functors $\mathfrak{I}_{\mu}^{-,-}$ relating the categories of kinded diagrams and their morphisms (most of which are subcategories of $ADiag^{k}(C)$). (Many more such functors arise by composition of the ones illustrated, and by applying nonidentity natural isomorphisms, where appropriate, to the domain or codomain objects of the $\mathfrak{I}_{\mu}^{-,-}$ functors.[1]) For the rest of this discussion we will suppress mention of μ as much as possible, we will assume all abstract diagrams are maximal regarding their kinds, and we will ignore the commutativity data in $ADiag(C)$ morphisms, which thus become $c : D_0 \to D_1$.

Note first that as $\langle C \rangle$ is isomorphic to C^K, standard isomorphisms in $\langle C \rangle$ are just identities; so there is no distiction between abstract diagrams in $\langle C \rangle$ entirely of kind std and those entirely of kind id. This explains the middle element of the bottom row of Fig. 2.

We recall that an object of $\langle C \rangle$ is an equivalence class of objects of C containing in particular a unique skeleton object from C^K, and that an arrow of $\langle C \rangle$ is an equivalence

1. The latter often proves to be equivalent to a change of choice of skeleton.

class of arrows of C in bijective correspondence with ordered pairs of representatives from its domain and codomain objects. We start with the relationship between concrete diagrams in C, and concrete diagrams in $\langle \mathsf{C} \rangle$. Thus $\Im_\mu^{\bullet,\langle\bullet\rangle}$ takes a concrete diagram γ in C to the concrete diagram $\gamma^\Diamond$ in $\langle \mathsf{C} \rangle$, such that the objects and arrows of γ are members of the equivalence classes which constitute the objects and arrows of $\gamma^\Diamond$. Conversely $\Im_\mu^{\langle\bullet\rangle,\bullet}$ sends a $\langle \mathsf{C} \rangle$ diagram $\gamma^\Diamond$ to the concrete diagram γ in C, for which the objects are the skeleton objects drawn from the equivalence class objects of $\gamma^\Diamond$, and the arrows are the unique arrows between the skeleton objects drawn from the arrow equivalence classes of $\gamma^\Diamond$. $\Im_\mu^{\bullet,\langle\bullet\rangle}$ and $\Im_\mu^{\langle\bullet\rangle,\bullet}$ constitute an equivalence of categories.

Proceeding to the top row of Fig. 2, we have the isomorphism between concrete diagrams in C and abstract diagrams entirely of kind id in C, given by functors $\Im_\mu^{\bullet,\mathsf{id}}$ and $\Im_\mu^{\mathsf{id},\bullet}$. This is essentially the correspondence between an item and the singleton containing it. A similar situation prevails on the bottom row between concrete diagrams in $\langle \mathsf{C} \rangle$ and abstract diagrams entirely of kind id or std in $\langle \mathsf{C} \rangle$, given by functors $\Im_\mu^{\langle\bullet\rangle,\langle\mathsf{id}\rangle}$ and $\Im_\mu^{\langle\mathsf{id}\rangle,\langle\bullet\rangle}$. That these are isomorphisms, follows readily from the only possible action on mediated morphisms of abstract diagrams of kind id.

We next discuss the middle column of Fig. 2. The object map of the functor $\Im_\mu^{\mathsf{std},\mathsf{id}}$ takes a maximal abstract diagram D of kind std, to the singleton containing the unique concrete diagram in D consisting of skeleton objects and arrows between them. The arrow map of the functor $\Im_\mu^{\mathsf{std},\mathsf{id}}$ takes a morphism $c : D_0 \to D_1$, to the morphism $\chi \circ \Xi_c(\gamma) = \{f\} : \{\gamma\} \to \{\delta\}$ where: γ is the unique concrete diagram in D_0 consisting of skeleton objects and arrows between them; δ is the corresponding one in D_1; and f is the natural transformation $\chi \circ \Xi_c(\gamma)$ given by the natural transformation $\Xi_c(\gamma)$ (taking γ to its target in D_1) postcomposed with the unique family of standard isomorphisms χ that takes the target of $\Xi_c(\gamma)$ to δ.

Conversely the object map of the feeble functor $\Im_\mu^{\mathsf{id},\mathsf{std}}$ takes a singleton containing an individual concrete diagram γ, to the abstract diagram D_0 consisting of the set of concrete diagrams related to γ by families of standard isomorphisms. The arrow map of $\Im_\mu^{\mathsf{id},\mathsf{std}}$ takes a morphism $\{f\} : \{\gamma\} \to \{\delta\}$ between singletons, mediated by a single natural transformation f, and sends it to the set of mediated morphisms determined as follows. Let φ be a function that maps each χ_γ, a natural transformation of γ consisting of standard isomorphisms, to a natural transformation χ_δ of δ consisting of standard isomorphisms. Such a function determines a morphism $c_\varphi : D_0 \to D_1$ of abstract diagrams of kind std, by mapping each concrete diagram in D_0 via $\chi_\delta \circ f \circ \chi_\gamma^{-1}$. The collection of all such morphisms for all possible choices of φ, determines the arrow map of $\Im_\mu^{\mathsf{id},\mathsf{std}}$.

The above makes $\Im_\mu^{\mathsf{id},\mathsf{std}}$ and $\Im_\mu^{\mathsf{std},\mathsf{id}}$ into a weak equivalence of categories, weakness being in the sense that $\Im_\mu^{\mathsf{id},\mathsf{std}}$ is a weak left adjoint to $\Im_\mu^{\mathsf{std},\mathsf{id}}$. The above also fixes the properties of the pair $\Im_\mu^{\bullet,\mathsf{std}}$ and $\Im_\mu^{\mathsf{std},\bullet}$ by requiring that the upper triangle in Fig. 2 commutes in the expected way. This means that $\Im_\mu^{\bullet,\mathsf{std}}$ is feeble and a weak left adjoint to $\Im_\mu^{\mathsf{std},\bullet}$.

Moving down, the object map of the functor $\Im_\mu^{\mathsf{std},\langle\mathsf{id}\rangle}$ takes an abstract diagram D of kind std in C to the singleton containing the concrete diagram γ in $\langle \mathsf{C} \rangle$ formed as follows. We select for each vertex m in the shape μ, the isomorphism class of concrete objects of C up to standard isomorphisms, occurring above m in the concrete diagrams of D; we select for each edge $e : m_0 \to m_1$ in the shape μ, the isomorphism class of concrete arrows of C up to standard isomorphisms, occurring above e in the concrete diagrams of D. The arrow map of the functor $\mathfrak{I}_{\mu}^{\mathsf{std},\langle\mathsf{id}\rangle}$ takes a morphism $c : \mathsf{D}_0 \to \mathsf{D}_1$ of abstract diagrams of kind std, and sends it to the morphism $\{\langle\Xi_c(\Gamma)(m)\rangle_m\} : \{\gamma\} \to \{\delta\}$ between singletons containing γ and δ, the images of D_0 and D_1, as follows. Let $\Xi_c(\Gamma)$ be the natural transformation Ξ_c at the object Γ of D_0 where Γ consists exclusively of skeleton objects and arrows (i.e. C^{K} objects and arrows only). Let $\langle\Xi_c(\Gamma)(m)\rangle_m$ be the collection of isomorphism classes up to standard isomorphisms of the arrows of $\Xi_c(\Gamma)$ as m ranges over the vertices of μ. These are arrows in $\langle\mathsf{C}\rangle$ forming a natural transformation of γ. We write $\{\langle\Xi_c(\Gamma)(m)\rangle_m\} : \{\gamma\} \to \{\delta\}$ for the natural extension of $\langle\Xi_c(\Gamma)(m)\rangle_m$ to an action on the singleton $\{\gamma\}$ containing γ.

Conversely the object map of the feeble functor $\mathfrak{I}_{\mu}^{\langle\mathsf{id}\rangle,\mathsf{std}}$ takes each singleton containing a concrete diagram γ in $\langle\mathsf{C}\rangle$ whose objects and arrows are isomorphism classes of C objects and arrows up to standard isomorphisms, and maps it to the abstract diagram D_0 consisting of the set of concrete diagrams in C which can be constructed as follows. We select for each vertex m in the shape μ, an element of the equivalence class which is the object of γ above it; and for each edge $e : m_0 \to m_1$ in the shape μ, we select from the equivalence class above e in γ, the unique element with the just chosen domain and codomain objects. The arrow map of $\mathfrak{I}_{\mu}^{\langle\mathsf{id}\rangle,\mathsf{std}}$ takes a morphism $\{\langle f\rangle_m\} : \{\gamma\} \to \{\delta\}$ between singletons, mediated by a single natural transformation $\langle f\rangle_m$ consisting of isomorphism classes of C arrows up to standard isomorphisms, containing in particular the collection $\{f_m\}$ all of whose domains and codomains are skeleton objects, and maps it as follows. Let Γ and Δ be the unique concrete diagrams in D_0 and D_1 all of whose objects and arrows are skeleton objects and arrows. Let φ be a function that maps each χ_Γ, a natural transformation of Γ formed by standard isomorphisms, to χ_Δ a natural transformation of Δ formed by standard isomorphisms. Such a function determines a morphism $c_\varphi : \mathsf{D}_0 \to \mathsf{D}_1$ of abstract diagrams of kind std, by mapping each concrete diagram in D_0 via $\chi_\Delta \circ f \circ \chi_\Gamma^{-1}$. The collection of all such morphisms for all possible choices of φ, determines the arrow map of $\mathfrak{I}_{\mu}^{\langle\mathsf{id}\rangle,\mathsf{std}}$.

As above $\mathfrak{I}_{\mu}^{\langle\mathsf{id}\rangle,\mathsf{std}}$ and $\mathfrak{I}_{\mu}^{\mathsf{std},\langle\mathsf{id}\rangle}$ form a weak equivalence of categories, with $\mathfrak{I}_{\mu}^{\langle\mathsf{id}\rangle,\mathsf{std}}$ being a weak left adjoint to $\mathfrak{I}_{\mu}^{\mathsf{std},\langle\mathsf{id}\rangle}$. Requiring that the lower triangle in Fig. 2 commutes also fixes the properties of the pair $\mathfrak{I}_{\mu}^{\langle\bullet\rangle,\langle\mathsf{id}\rangle}$ and $\mathfrak{I}_{\mu}^{\langle\mathsf{id}\rangle,\langle\bullet\rangle}$, with $\mathfrak{I}_{\mu}^{\langle\bullet\rangle,\langle\mathsf{id}\rangle}$ being a weak left adjoint to $\mathfrak{I}_{\mu}^{\langle\mathsf{id}\rangle,\langle\bullet\rangle}$. We can also see that the rectangle in the left and middle columns of Fig. 2 commutes as we would expect.

We turn to the rectangle in the lower right part of Fig. 2. We observe first the following fact. Suppose in C we have arrows $f : x \to y$, $f' : x' \to y'$, and standard isomorphisms $\sigma(x,x') : x \to x'$, $\sigma(y,y') : y \to y'$, making a commuting square. Let $\tau(x,x') : x \to x'$ be any isomorphism from x to x'. In general there will not be an isomorphism $\tau(y,y') : y \to y'$ making f, f', $\tau(x,x')$, $\tau(y,y')$ commute. However we will assume subsequently that C has enough isomorphisms, in the sense that such a $\tau(y, y')$ can always be found, though it may not be unique. For example Set and Gr have enough isomorphisms.

Now the object map of the feeble functor $\mathfrak{I}_{\mu}^{\mathsf{std},\mathsf{iso}}$ sends an abstract diagram $\mathsf{D}^{\mathsf{std}}$ entirely of kind std to the abstract diagram $\mathsf{D}^{\mathsf{iso}}$ having the same objects, but this time entirely of kind iso. Viewed as a category, D^{iso} has merely acquired more arrows in this process, namely the natural transformations between its concrete diagrams, incorporating at least one nonstandard isomorphism. The arrow map of the feeble functor $\mathfrak{I}_\mu^{\mathsf{std},\mathsf{iso}}$ sends a mediated morphism $c^{\mathsf{std}} : D_0{}^{\mathsf{std}} \to D_1{}^{\mathsf{std}}$ to the set of extensions of c^{std} which cover all the additional natural transformations too. Such extensions will exist by our observation above, but in general they will not be unique.

The object map of the functor $\mathfrak{I}_\mu^{\mathsf{iso},\mathsf{std}}$ likewise sends an abstract diagram D^{iso} entirely of kind iso to the abstract diagram D^{std} having the same objects, but this time entirely of kind std. As a category, D^{iso} is mapped to the subcategory D^{std} having only standard isomorphism natural transformations as arrows. The arrow map of $\mathfrak{I}_\mu^{\mathsf{iso},\mathsf{std}}$ sends a mediated morphism $c^{\mathsf{iso}} : D_0{}^{\mathsf{iso}} \to D_1{}^{\mathsf{iso}}$ to the mediated morphism $c^{\mathsf{std}} : D_0{}^{\mathsf{std}} \to D_1{}^{\mathsf{std}}$ determined as follows. Let Γ be the concrete diagram in $D_0{}^{\mathsf{iso}}$ consisting exclusively of skeleton graphs and morphisms between them. ($D_0{}^{\mathsf{iso}}$ will contain this since it is maximal.) Let $\Xi_{c^{\mathsf{iso}}}(\Gamma)$ be the natural transformation that mediates the morphism c^{iso} at Γ. Let $n : \Gamma \to \gamma$ be an arrow in $D_0{}^{\mathsf{std}}$, and let $\Xi_{c^{\mathsf{iso}}}(\gamma)$ be the corresponding natural transformation at γ. Suppose $n : \Gamma \to \gamma$ is mapped by c^{iso} to the concrete diagram isomorphism $c^{\mathsf{iso}}(n) : c^{\mathsf{iso}}(\Gamma) \to c^{\mathsf{iso}}(\gamma)$. Let χ_m be the collection of isomorphisms such that $\chi_m \circ c^{\mathsf{iso}}(n)$ is a concrete diagram morphism consisting entirely of standard isomorphisms. Then $\chi_m \circ c^{\mathsf{iso}}(n)$ is a morphism of $D_1{}^{\mathsf{std}}$ mediated by $\Xi_{c^{\mathsf{iso}}}(\Gamma)$ at Γ and $\chi_m \circ \Xi_{c^{\mathsf{iso}}}(\gamma)$ at γ. For each γ in $D_0{}^{\mathsf{iso}}$ we replace its subfamily of mediating arrows by the subfamily $\chi_m \circ \Xi_{c^{\mathsf{iso}}}(\gamma)$ so determined. By the properties of standard isomorphisms, all other morphisms $n : \gamma \to \delta$ in $D_0{}^{\mathsf{std}}$ are mapped to morphisms of $D_1{}^{\mathsf{std}}$ which compose properly. This gives the morphism $c^{\mathsf{std}} : D_0{}^{\mathsf{std}} \to D_1{}^{\mathsf{std}}$.

The functors $\mathfrak{I}_\mu^{\langle\mathsf{id}\rangle,\langle\mathsf{iso}\rangle}$ and $\mathfrak{I}_\mu^{\langle\mathsf{iso}\rangle,\langle\mathsf{id}\rangle}$ are similar. The object map of the feeble functor $\mathfrak{I}_\mu^{\langle\mathsf{id}\rangle,\langle\mathsf{iso}\rangle}$ maps the objects via identities — the objects (up to $\mathsf{id} = \mathsf{std}$) being singletons containing concrete diagrams built out of objects and arrows which are equivalence classes of C objects and arrows up to standard isomorphisms. Up to $\mathsf{id} = \mathsf{std}$, abstract diagrams in $\langle C\rangle$ have only the identity automorphism; however up to iso, they in general acquire nontrivial automorphisms. The arrow map of the feeble functor $\mathfrak{I}_\mu^{\langle\mathsf{id}\rangle,\langle\mathsf{iso}\rangle}$ takes a morphism $\{\langle f\rangle_m\} : \{\gamma\} \to \{\delta\}$ between singletons, mediated by a single natural transformation $\langle f\rangle_m$ consisting of isomorphism classes of C arrows up to standard isomorphisms, containing in particular the collection $\{f_m\}$ all of whose domains and codomains are skeleton objects, and maps it as follows. Let Γ and Δ consist exclusively of skeleton objects and arrows, as in the discussion of $\mathfrak{I}_\mu^{\langle\mathsf{id}\rangle,\mathsf{std}}$. Then each nontrivial automorphism of γ (respectively δ) has a unique representative for Γ (respectively Δ). Moreover, each nontrivial automorphism a_Γ of Γ maps via f_m to a nontrivial automorphism a_Δ of Δ, in general in many ways. The equivalence classes up to standard isomorphisms, of the objects and arrows of a_Δ, yield an automorphism of δ which gives a possible action of $\mathfrak{I}_\mu^{\langle\mathsf{id}\rangle,\langle\mathsf{iso}\rangle}$ on the arrow $\{\langle f\rangle_m\}$. The collection of all such possibilities determines the arrow map of $\mathfrak{I}_\mu^{\langle\mathsf{id}\rangle,\langle\mathsf{iso}\rangle}$.

The functor $\mathfrak{I}_\mu^{\langle\mathsf{iso}\rangle,\langle\mathsf{id}\rangle}$ could not be simpler. The action on objects is the identity. On arrows, it is just the restriction to identity automorphisms only, of the action of arrows $\{\langle f\rangle_m\} : \{\gamma\} \to \{\delta\}$ between singletons.

As we had before, the functor pairs $\Im_\mu^{\mathsf{id},\mathsf{iso}}$, $\Im_\mu^{\mathsf{iso},\mathsf{id}}$ and $\Im_\mu^{\langle\mathsf{id}\rangle,\langle\mathsf{iso}\rangle}$, $\Im_\mu^{\langle\mathsf{iso}\rangle,\langle\mathsf{id}\rangle}$ give weak equivalences of categories, in the sense that $\Im_\mu^{\mathsf{id},\mathsf{iso}}$ is a weak left adjoint to $\Im_\mu^{\mathsf{iso},\mathsf{id}}$ and $\Im_\mu^{\langle\mathsf{id}\rangle,\langle\mathsf{iso}\rangle}$ is a weak left adjoint to $\Im_\mu^{\langle\mathsf{iso}\rangle,\langle\mathsf{id}\rangle}$.

Finally we consider $\Im_\mu^{\langle\mathsf{iso}\rangle,\mathsf{iso}}$ and $\Im_\mu^{\mathsf{iso},\langle\mathsf{iso}\rangle}$. The functor $\Im_\mu^{\mathsf{iso},\langle\mathsf{iso}\rangle}$ behaves like the functor $\Im_\mu^{\mathsf{std},\langle\mathsf{id}\rangle}$ except that diagram morphisms must include also the nontrivial automorphisms. Each such nontrivial automorphism of a concrete representative of an abstract diagram of kind iso in $\langle \mathsf{C} \rangle$, is simply mapped to the collection of equivalence classes up to standard isomorphisms in the expected way. Likewise, the feeble functor $\Im_\mu^{\langle\mathsf{iso}\rangle,\mathsf{iso}}$ behaves like $\Im_\mu^{\langle\mathsf{id}\rangle,\mathsf{std}}$ except that again nontrivial automorphisms must be taken into account. These are mapped just like all the other arrows between abstract diagrams of kind iso in $\langle \mathsf{C} \rangle$.

Unsurprisingly the functors $\Im_\mu^{\langle\mathsf{iso}\rangle,\mathsf{iso}}$ and $\Im_\mu^{\mathsf{iso},\langle\mathsf{iso}\rangle}$ form a weak equivalence of categories with $\Im_\mu^{\langle\mathsf{iso}\rangle,\mathsf{iso}}$ being a weak left adjoint to $\Im_\mu^{\mathsf{iso},\langle\mathsf{iso}\rangle}$.

Fig. 2 summarises all of the above by distinguishing the feeble functors from the rest with a small circle. It is worth noting that the feebleness of $\Im_\mu^{\mathsf{std},\mathsf{iso}}$ and of $\Im_\mu^{\langle\mathsf{id}\rangle,\langle\mathsf{iso}\rangle}$ is due to the nonunique way that arbitrary nonstandard isomorphisms translate along arbitrary morphisms, while the feebleness of the other functors is attributable to the many different mediated morphisms of abstract diagrams which map, under equivalence up to standard isomorphisms, to the same morphism of skeleton concrete diagrams. Of course all of this detail gets swept away when one moves to ***MADiag***(C). Also the preceding discussion described the situation when all vertices in the shape of an abstract diagram are of the same kind. In diagrams where the kind varies from vertex to vertex, the facts of the matter may be determined by an easy extrapolation.

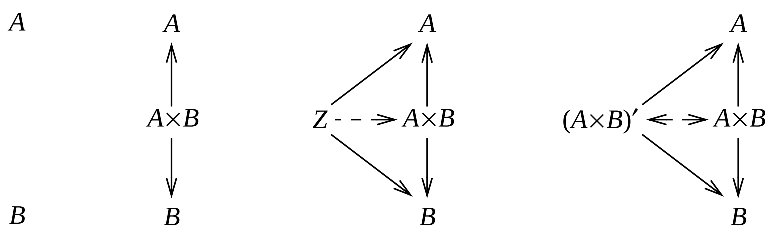
6 Maximality, and the Uniqueness of Diagrammatic Constructions

Diagrammatic reasoning is typically used in category theory in two principal ways. In the first, a diagram displays one or more equalities between compositions of arrows whose existence is already assured. In such cases there is no ambiguity about the relationships that are being stated.

In the second, a diagram displays one or more equalities between compositions of arrows whose existence is claimed. In the such cases, because the existence claim can typically only be made up to isomorphism, the reasoning is more complicated. There is an initial phase in which an explicit construction is given that solves the problem, usually in a canonical way, and this is followed by a second phase in which the universality of the solution up to isomorphism is demonstrated. The latter involves reasoning about a more complicated diagram and showing that a certain arrow is an iso. We reconsider this activity in the light of abstract diagrams. First though, a word about shapes.

Most diagrammatic reasoning takes place without mentioning shapes. It is assumed that the context provides enough clues to render superfluous the explicit definition of the shape and its relation to the substance of the diagram, which usually arises via the implicit geometrical or textual relationships between the diagram and other information appearing on the same page. Thus there is a commonly agreed if unstated shape for each diagram that enters into a particular discourse. The shape, fixed for the duration (though negotiable via isomorphisms of the shape graph, a fact that also excuses the very specific shapes chosen for arrow-diagrams in Section 4), plays the same role in diagrammatic reasoning that natural language plays in most other activities, i.e. the provision of a common framework within which the discourse takes place. If the shape has nontrivial symmetries that the context fails to disambiguate sufficiently, and considerations that break the symmetry enter late into the discourse, the protagonists of the discourse may get a surprise, just as can happen with natural language. Despite the risks, we conform to this standard practice to avoid verbosity below.

We now examine a simple example, the construction of products in a category which supports them. So let C have products. The procedure is illustrated in Fig. 3, and begins with two objects A and B (Fig. 3.(a)), for which we build a product object $A{\times}B$ together with its projections to A and B (Fig. 3.(b)), which enjoy the familiar universal factorisation properties (Fig. 3.(c)). Finally we show that any object $(A{\times}B)'$ enjoying the same factorisation properties is isomorphic to $A{\times}B$ in C (Fig. 3.(d)).



(a) (b) (c) (d)

Fig. 3 Products.

All of Figs. 3.(a)-3.(d) can be viewed just as individual concrete diagrams with obvious shapes. However there is an alternative perspective as follows. Fig. 3.(a) is a concrete diagram that describes the initial situation, while Fig. 3.(b) depicts the result of the construction and is another concrete diagram. Figs. 3.(a) and 3.(b) are the domain and codomain of an obvious arrow of $\mathsf{Diag}(\mathsf{C})$ based on the inclusion of the shape of (a) into that of (b), thus partially characterising the product construction as a family of $\mathsf{Diag}(\mathsf{C})$ morphisms parameterised by A and B, viz. $Prod(A,B) : (A \quad B) \to (A \leftarrow A{\times}B \to B)$, where for convenience we have suppressed the standard $\mathsf{Diag}(\mathsf{C})$ morphism notations.

In a variation on this theme, we note that there are also inclusions which we can succinctly indicate by 3.(a) $\to$ 3.(c) $\leftarrow$ 3.(b). This is a $\mathsf{Diag}(\mathsf{C})$ cospan which provides a more detailed picture of the product construction. Note however that this still fails to describe some of the vital aspects of the construction, such as the universal quantification over Z or the uniqueness of the arrow $Z \to A{\times}B$. Still, this is no worse than what is conveyed by the diagrams in conventional discussions of the product.

Fig. 3.(d) can be regarded as a concrete diagram describing the isomorphism invariance of the product of course, but it is much more in the spirit of this paper to see it as two constituent concrete diagrams of the kinded abstract diagram $A \leftarrow [A{\times}B] \rightarrow B$ (using Notation 5.3), which is up-to-identity at A and B and up-to-arbitrary-isomorphisms at $A{\times}B$. This KAD captures succinctly all the possible concrete products obtainable from A and B.

Moreover noting that C objects which are isomorphic to A and B share the same family of products as A and B themselves, means that there is a more abstract formulation of the product, this time in $\boldsymbol{\mathsf{MADiag}}(\mathsf{C})$, which casts it as a family of inclusion morphisms[2] parameterised by $[A]$ and $[B]$, viz. $\mathsf{A}Prod([A],[B]) : ([A]\ \ [B]) \rightarrow ([A] \leftarrow [A{\times}B] \rightarrow [B])$. And the $\boldsymbol{\mathsf{MADiag}}(\mathsf{C})$ formulation is available even if the the concrete version had not been defined as $\mathsf{Diag}(\mathsf{C})$ morphisms (for example by permitting choice for the introduced object $A{\times}B$ and attendant arrows) — though if it *is* defined using morphisms, then the functor $\boldsymbol{\mathsf{MAbs}}$ carries the concrete formulation into the abstract one. Clearly this is much neater than seeking an abstract formulation via the category $\langle \mathsf{C} \rangle$, where the choice of standard isomorphisms would have intruded unavoidably.

The $\boldsymbol{\mathsf{MADiag}}(\mathsf{C})$ family of inclusion morphisms $\mathsf{A}Prod([A],[B])$ expresses via abstract diagrams and abstract arrow-diagrams, the notion that up to isomorphism, the product yields a unique outcome. We now generalise this to a methodological statement, which in the notation of Section 5, can be tritely expressed as saying that to pass from the usual version of a construction to the $\boldsymbol{\mathsf{MADiag}}(\mathsf{C})$ version, it is sufficient to merely 'put square brackets round everything'.

Definition 6.1 A concrete canonical categorical construction (with values in C) is an argument that establishes the existence of a family of $\mathsf{Diag}(\mathsf{C})$ inclusion morphisms parameterised by some $\mathsf{Diag}(\mathsf{C})$ objects and arrows say $A_1 \ldots A_n$. Thus it can be expressed as $Constr(A_1 \ldots A_n) : \gamma_0(A_1 \ldots A_n) \rightarrow \gamma_1(A_1 \ldots A_n)$, where γ_0 is the premiss concrete diagram and γ_1 is the result concrete diagram.

The 'canonical' qualification here excludes those cases in which the choice of objects and arrows newly introduced during the construction does not depend functionally on the parameters, leading to no unique choice of $\gamma_1(A_1 \ldots A_n)$.

Definition 6.2 An abstract categorical construction (with values in C) is an argument that establishes the existence of a family of $\boldsymbol{\mathsf{MADiag}}(\mathsf{C})$ inclusion morphisms parameterised by some $\boldsymbol{\mathsf{MADiag}}(\mathsf{C})$ objects and arrows say $[A_1] \ldots [A_n]$. Thus it can be expressed as $\mathsf{A}Constr([A_1] \ldots [A_n]) : \mathsf{D}_0([A_1] \ldots [A_n]) \rightarrow \mathsf{D}_1([A_1] \ldots [A_n])$, where D_0 is the premiss abstract diagram and D_1 is the result abstract diagram.

Note that there is no need for a 'canonical' qualification here.

Proposition 6.3 For every concrete canonical categorical construction there is a corresponding abstract construction given by mapping the concrete inclusion morphisms $Constr(A_1 \ldots A_n)$ to the abstract inclusion morphisms $\mathsf{A}Constr([A_1] \ldots [A_n])$ via $\boldsymbol{\mathsf{MAbs}}$.

Proof. See Theorem 4.8. ☺

2. An $\boldsymbol{\mathsf{MADiag}}(\mathsf{C})$ inclusion morphism is an abstract arrow-diagram fashioned from a concrete inclusion arrow-diagram in $\boldsymbol{\mathsf{Diag}}(\mathsf{C})$, itself fashioned from a concrete inclusion in $\mathsf{Diag}(\mathsf{C})$.

The preceding ideas have wide application and we now look at some more examples. Firstly adjunctions. One way of constructing an adjunction $F \dashv U$ is to build a universal arrow for each object A. In Fig. 4 this is characterised by an inclusion from (a) to (b), and the ***MADiag***($\mathcal{C}$) version just puts square brackets round everything in Fig. 4.

Fig. 4 Adjunctions.

Another example, the construction of a cartesian closure, can be characterised by the inclusion of A and B in (a) of Fig. 5 into (b). The new elements, i.e. $(A \Rightarrow B)$ and the product and *eval* arrows, are usually constructed canonically, but only actually needed up to concrete isomorphisms, leading to a KAD version with all of Fig. 5 except A and B 'square bracketed', as well as a fully abstract version in ***MADiag***($\mathcal{C}$) with everything 'square bracketed' including A and B.

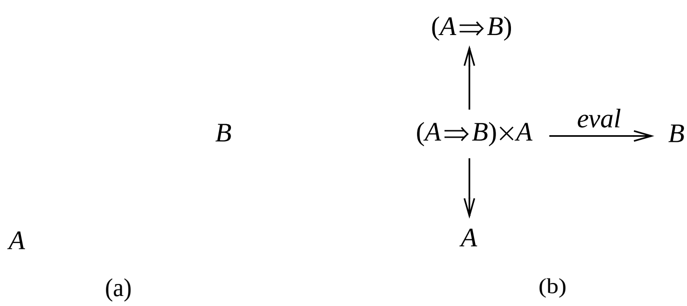


Fig. 5 Cartesian Closure.

A different kind of example arises in the well known Snake Lemma, illustrated by the inclusion of (a) into (b) in Fig. 6. Although the concrete construction is just of an arrow between existing objects, and is thus not ambiguous, the details of the construction involve building a number of additional objects and arrows (not shown), whuch *are* characterised only up to isomorphism. Thus despite appearances, the abstract versions of the construction are more determined than the concrete one.

Our last example is the construction of ends [Mac Lane (1971)]. This corresponds to the inclusion of (a) into (b) in Fig. 7, where $S : \mathcal{X}^{\mathrm{op}} \times \mathcal{X} \to \mathcal{C}$ is a bifunctor and for any E' satisfying the same suite of properties as E, there is a unique arrow $e : E' \to E$ etc. It is clear that the abstract verison of this specifies $[E]$ uniquely, and thus that the construction is analogous to the construction of products or to the CCC construction.

Note that in all of these examples the properties of the abstract diagram were obtained by first considering the properties of the usual concrete construction (i.e. no 'something

(a) (b)

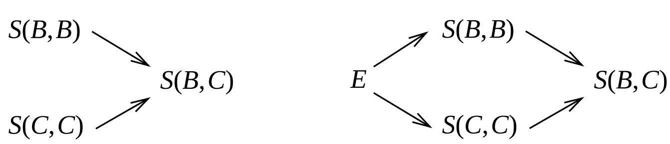
Fig. 6 The Snake Lemma.

for nothing') — in essence we have a fresh way of viewing the conventional answers, one which eg. legitimises the overwhelming temptation to speak of '*the* product' or '*the* CCC arrow object' when given a pair of (isomorphism classes of) objects.

7 Maximal Abstract Diagrams and Semantics

In this section we briefly indicate some of the possibilities in semantics opened up by abstract diagrams.

In the first example we recall that CCCs provide a semantics for the simply typed lambda calculus (STLC), (see [Barendregt (1984)] for the untyped lambda calculus and [Jacobs (1998)] for the typed variant). This is well known, so we do not revisit all the details save to note that the types of the STLC constitute the objects of a CCC and equivalence classes of terms constitute the arrows. Since there are many other CCCs than this one, there are also many other semantic models, and any functor from the standard CCC to such a potential alternative model will yield a semantics. Whether such a potential alternative model is interesting is at least partly determined by how canonical the constituents are. The availability of abstract diagrams now opens up the possibility that some candidate CCCs, viewed as unattractive because their objects and arrows were not canonical, now acquire more attractive abstract counterparts because their abstract objects and arrows can be viewed as abstract diagrams. We could mention CCCs of diagrams themselves, eg. built on the basis of constructions detailed in Ch. 6 of [Barr and Wells (1999)]. For a suggestion as to why such models might be of interest, see the last paragraph in this section.



(a) (b)

Fig. 7 Ends.

A second area we briefly discuss concerns graph transformation by the application of graph rewrite rules, an area where the ambiguities caused by nontrivial graph automorphisms are particularly keenly felt. There are many different ways to design graph rewriting systems; we focus on the technique studied at length in [Rozenberg (1997)] and [Ehrig et al. (1999)]. The basic idea is that rules are spans $L \leftarrow K \rightarrow R$ in Gr. A redex for such a rule is a diagram like Fig. 8.(a), in which there is a graph morphism from L to the object graph G, and the application of the rule succeeds when we can construct the diagram in Fig. 8.(b), i.e. a span morphism, in which the two squares are both pushouts in Gr. Such a setup can describe many operational aspects of computational situations quite cleanly. The point is that since there are no canonical representatives for arbitrary graphs, irritating ambiguities up to isomorphism proliferate at all stages of the theory. This certainly becomes a nuisance when one wishes to study more abstract aspects of the operational semantics. In [Rozenberg (1997), Ehrig et al. (1999)] there are solutions to such questions constructed via concrete diagrams in $\langle \mathsf{Gr} \rangle$, with their reliance on some skeleton category Gr^{K}. However the deployment of MADs gives an approach that is easier, more elegant, and independent of any choices (of skeleton). It is clear that there is a $\mathsf{Diag}(\mathsf{Gr})$ inclusion from (a) to (b) in Fig. 8 (when D and H are constructed canonically, as they usually are), so that abstract rewrites can be simply defined by 'putting square brackets round everything in Fig. 8', thus getting an $\boldsymbol{\mathsf{MADiag}}(\mathsf{Gr})$ inclusion. This gives the starting point for a clean reworking of the theory, which will be explored elsewhere.

$$
\begin{array}{ccccc}
L & \longleftarrow & K & \longrightarrow & R \\
\downarrow & & & & \\
G & & & &
\end{array}
\qquad\qquad
\begin{array}{ccccc}
L & \longleftarrow & K & \longrightarrow & R \\
\downarrow & & \downarrow & & \downarrow \\
G & \longleftarrow & D & \longrightarrow & H
\end{array}
$$

(a) (b)

Fig. 8 Graph Transformation.

One final observation which we do not follow up in this paper is the following. The mathematical study of semantics naturally emphasises convincing abstract models, usually characterised by canonical properties. The practical business of semantics, the implementation of languages on a computer, is unavoidably characterised by pragmatic aspects, such as the choice of specific memory locations etc. Abstract diagrams give us the possibility of bringing the two spheres closer than is usually found, by allowing the development of abstractions that are in effect isomorphism classes of implementations. This is an idea that will be developed in other publications.

8 Relationship to Other Work

To the author's knowledge, the concept of abstract diagram presented here has not been given before, but some related ideas have appeared in the literature. Rather close is the work on anafunctors in [Makkai (1996)]. If $\mathcal{X}$ and $\mathcal{A}$ are categories, then an anafunctor $\mathcal{F} : \mathcal{X} \to \mathcal{A}$ is a class $|F|$ together with maps $\sigma : |F| \to \mathrm{Ob}(\mathcal{X})$ and $\tau : |F| \to \mathrm{Ob}(\mathcal{A})$ such that if $f : \sigma(x) \to \sigma(y)$ is an arrow of $\mathcal{X}$, then there is an arrow $F_{x,y}(f) : \tau(x) \to \tau(y)$ in $\mathcal{A}$. Moreover the correspondence $f \mapsto F_{x,y}(f)$ (parameterised by the x,y) must behave 'functorially'. Thus $\mathrm{id}_{\sigma(x)}$ must correspond to $F_{x,x}(\mathrm{id}_{\sigma(x)}) = \mathrm{id}_{\tau(x)}$, and the composition of $f : \sigma(x) \to \sigma(y)$ and $g : \sigma(y) \to \sigma(z)$ must correspond to the composition of $F_{x,y}(f)$ and $F_{y,z}(g)$. (As a consequence, every $F_{x,y}(\mathrm{id}_{\sigma(x)=\sigma(y)})$ corresponds to an isomorphism of $\tau(x)$.)

For an anafunctor $\mathcal{F} : \mathcal{X} \to \mathcal{A}$, if there is a bijection between elements x in $|F|$ and the collection of pairs $(\sigma(x), \tau(x))$, then up to foundational niceties, and regarding $\mathcal{X}$ as a shape category, it is easy enough to see that $\mathcal{F}$ corresponds to an abstract diagram, namely to a subcategory of $[\mathcal{X}, \mathcal{A}]$. Fixing $\sigma(x)$, the elements x can be seen as labelling the various objects of $\mathcal{A}$ isomorphic to $\tau(x)$. If there is no bijection as described, then the anafunctor contains more data, i.e. the elements x in $|F|$ act as names for the associated pairs $(\sigma(x), \tau(x))$. Assuming for simplicity the correspondence between anafunctors and abstract diagrams, Makkai's saturated anafunctors now correspond to our maximal abstract diagrams. We restrict for the rest of this section to this saturated/maximal case, the one also of most interest to Makkai.

On this assumption we immediately notice that since we constructed abstract diagrams using purely functorial techniques, many of the naturality properties of anafunctors proved directly in [Makkai (1996)] follow immediately in the abstract diagram formulation.

Let us now reconsider products in the two schemes. Makkai's canonical product can be expressed in our kinded notation as the abstract diagram $A \leftarrow [A \times B] \to B$, asymmetric between inputs and outputs as regards our kinds. Our more symmetrical canonical product $[A] \leftarrow [A \times B] \to [B]$ does not appear in [Makkai (1996)], though if the abstract diagram $[A]$ of shape • is regarded as an anafunctor $\mathcal{F} : \bullet \to \mathcal{A}$, then such a formulation can be recovered without great effort, modulo considerations of size.

We close this discussion with some comments on foundational issues which we admittedly neglect in this paper. Our constructions of abstract diagrams have been informed (albeit indirectly) by the Grothendieck Construction[3], including situations which in a Gödel-Bernays formulation would be large. We regard this as innocuous, since given a suitable choice of skeleton for $\mathcal{C}$, the abstract diagram categories become equivalent to locally small ones, as the feeble functors of Section 5 bear witness. This is comparable to a similar remark in [Makkai (1996)], and is about as innocent a use of large categories as one might wish for.

A further litmus test of propriety consists of examining our use of equality during the course of the abstract diagram constructions. In the influential [Bénabou (1985)], the author rails quite energetically against the undisciplined use of equality in category theory. In our paper, the only place equality was used nontrivially was in the consideration of diagram commutativity, which we can always restrict to locally small situations by

3. It is primarily this fact that supports the assertion in the Introduction, that our constructions are purely functorial.

insisting that shapes are suitably small. Aside from that we used identity of domain and codomain objects of concrete arrow-diagrams during the composition of abstract arrow-diagrams. Bénabou does not disparage use of identities and this is something we regard as innocuous.

9 Conclusions

In the previous sections, having motivated the search for a more workable notion of abstract diagram than has been available hitherto, we defined abstract diagrams via functor categories. This conception turned out to be useful in giving diagrams a canonical character, furthermore one that was obtained in the simplest conceivable manner: 'Just put square brackets round everything in the corresponding concrete diagram'. In this sense, the advocated course of action for somethiing like the product construction say, was a little different to that occurring in Makkai's reappraisal of products via anafunctors, though Makkai's anafunctors offer an approach that yields similar results to abstract diagrams in many respects. Beyond this, the more uniform perspective afforded by abstract diagrams suggests their adoption for various semantic purposes for which use of the analogous concrete diagrams is much less attractive due to their non-canonical nature. These aspects in particular merit further exploration.

Acknowledgements

It is a pleasure to thank Peter Aczel and Harold Simmons for comments on this paper. Thanks are also due to Peter Johnstone for comments on earlier work from which this paper arose. Mostly, thanks are due to Andrea Corradini for collaboration on the earlier work, without which, the developments reported here would not have been thought of.

Appendix: Feeble Functors and Weak Adjunctions

The material of this part is adapted from [Krishnan (1981)] and [Kainen (1971)]. A feeble functor $\mathsf{F} : \mathsf{A} \to \mathsf{B}$ maps objects of A to objects of B as usual, but maps arrows $f : a \to b$ of A to nonempty sets of arrows of B via $\mathsf{F}(f : a \to b) \subseteq \hom_{\mathsf{B}}(\mathsf{F}(a), \mathsf{F}(b))$, such that if f and g are composable, then $\mathsf{F}(g) \circ \mathsf{F}(f) \subseteq \mathsf{F}(g \circ f)$ with the obvious overloading of the composition symbol. Given two feeble functors $\mathsf{F}, \mathsf{G} : \mathsf{A} \to \mathsf{B}$, a left-natural transformation $\eta : \mathsf{F} \to \mathsf{G}$ maps each object a in A to a non-empty set of arrows $\eta(a) \subseteq \hom_{\mathsf{B}}(\mathsf{F}(a), \mathsf{G}(a))$ such that for each arrow $f : a \to b$ of A, $\eta(b) \circ \mathsf{F}(f) \supseteq \mathsf{G}(f) \circ \eta(a)$. If the direction of the inclusion is reversed we have a right-natural transformation. A transformation between two feeble functors that is both left-natural and right-natural is called a natural transformation.

Feeble functors are used in weak adjunctions which we now describe. Let $\mathsf{F} : \mathsf{A} \to \mathsf{B}$ be a feeble functor and $\mathsf{G} : \mathsf{B} \to \mathsf{A}$ be a (normal) functor. Then F is a weak left adjoint of G iff there exists a natural transformation $n : \hom_{\mathsf{B}}(\mathsf{F} \times \mathrm{Id}(\mathsf{B})) \to \hom_{\mathsf{A}}(\mathrm{Id}(\mathsf{A}) \times \mathsf{G})$ and a left-natural transformation $m : \hom_{\mathsf{A}}(\mathrm{Id}(\mathsf{A}) \times \mathsf{G}) \to \hom_{\mathsf{B}}(\mathsf{F} \times \mathrm{Id}(\mathsf{B}))$, such that $n \circ m = 1_{\hom_{\mathsf{A}}(\mathrm{Id}(\mathsf{A}) \times \mathsf{G})}$ and $m \circ n \supseteq 1_{\hom_{\mathsf{B}}(\mathsf{F} \times \mathrm{Id}(\mathsf{B}))}$. Here both $\hom_{\_}(-,-)$ notations are being viewed as functors $\mathsf{A}^{\mathrm{op}} \times \mathsf{B} \to \mathsf{Set}_{\mathsf{CF}}$, where $\mathsf{Set}_{\mathsf{CF}}$ is the category of sets and cofull relations between them. Equivalent conditions are given by the following.

Theorem A.1 Let $G : B \to A$ be a functor. Then among the conditions below we have (1) ⇒ (2) ⇒ (3) ⇒ (1) and (1*) ⇒ (2*) ⇒ (3*) ⇒ (1*).

(1) There is a feeble functor $F : A \to B$ which is a weak left adjoint to G.

(1*) In addition to (1), if (n,m) define the weak adjunction then: **(** a is an object of $A \wedge h \in \hom_B(F(a),F(a)) \wedge n(a,F(a)) \wedge G(h) \circ k = k$ **)** ⇒ **(** $h = \mathrm{id}_{F(a)}$ **)**.

(2) There is a feeble functor $F : A \to B$ and a natural transformation $\eta : \mathrm{Id}(A) \to G \circ F$, and for every object b of B a non-empty set $\nu(b) \subseteq \hom_B(F \circ G(b), b)$ such that: (a), $G(\nu(b)) \circ \eta(G(b)) = \mathrm{id}_{G(b)}$, and (b), for every object a of A **(** $f \in \hom_A(a,G(a)) \wedge h \in \nu(b) \circ F(f)$ **)** ⇒ **(** $G(h) \circ \eta(a) = f$ **)**.

(2*) In addition to (2), for every object a of A **(** $h \in \hom_B(F(a),F(a)) \wedge G(h) \circ \eta(a) = \eta(a)$ **)** ⇒ **(** $h = \mathrm{id}_{F(a)}$ **)**.

(3) Every object a of A has a universal arrow $(u : a \to G(b_a), b_a)$ in the sense that for every object b of B and every $f : a \to G(b)$ there is a (not necessarily unique) arrow $g : b_a \to b$ such that $G(g) \circ u = f$.

(3*) In addition to (3), **(** $h : b \to b \wedge G(h) \circ u = u$ **)** ⇒ **(** $h = \mathrm{id}_b$ **)**.

References

Barendregt H. P. (1984); The Lambda Calculus: Its Syntax and Semantics. North-Holland.

Barr M., Wells C. (1999); Category Theory for Computing Science. (3 ed.), Publications CRM, Université de Montréal.

Bénabou J. (1985); Fibered Categories and the Foundations of Naive Category Theory. J.S.L. **50**, 10-37.

Borceux F. (1994); Handbook of Categorical Algebra. Vol 2. Cambridge University Press.

Ehrig H., Kreowski H-J., Montanari U., Rozenberg G. (eds.) (1999); Handbook of Graph Grammars and Computing by Graph Transformations: Concurrency, Parallelism, and Distribution. World Scientific.

Jacobs B. (1998); Categorical Logic and Type Theory. North-Holland.

Kainen P. (1971); Weak Adjoint Functors. Math. Z. **122**, 1-9.

Krishnan V. (1981); An Introduction to Category Theory. North-Holland.

Mac Lane S. (1971); Categories for the Working Mathematician. Springer.

Makkai M. (1996); Avoiding the Axiom of Choice in General Category Theory. J.P.A.A. **108**, 109-173.

Rozenberg G. (ed.) (1997); Handbook of Graph Grammars and Computing by Graph Transformation. World Scientific.