Parametric Distributive Laws: uniform monad composition

Lorenzo Perticone

September 25, 2025

Abstract

Monads play an important role in both the syntax and semantics of modern functional programming languages ([Mog91]). The problem of combining them has been of profound interest at least since the 90s, and different approaches have been employed to tackle it: currently, the most prevalent notion is that of *monad transformers* ([CM93; LHJ95]). We provide a novel abstract framework to describe such "compositions" which we call *parametric distributive laws*.

Our description of this framework hinges on the theory of 2-categories and relies strongly on the construction of a left (1-) adjoint to the construction of monads internal to a 2-category: since distributive laws are monads in monads, we obtain a semi-strict "walking distributive law" by Gray-tensoring the walking monad with itself. By Gray-tensoring again, we can then construct a parametric variant thereof.

We demonstrate the applicability of such a framework by providing two concrete examples (involving the `Writer` and `Either` monads), explicitly describing morphisms of such parametric distributive laws, and showing how (an iterated version of) our construction can be employed to describe (parametric) iterated distributive laws (in the sense of [Che11]), motivating the appearance of the Yang-Baxter equations as coherences between the distributive laws involved.

Introduction

The first structured uses of monads in computer science are due to Moggi [Mog91], with the purpose of modeling different features (e.g. non-termination, side effects, nondeterminism) in the semantics of a programming language. The problem of composing them has been explored long before that, in contexts such as categorical algebra. Different authors have proposed different ways to describe such compositions; the main approaches seem to be *distributive laws* (mainly employed in algebraic contexts) and *monad transformers* (mainly used in computer science, e.g. the Haskell standard library).

A *monad transformer* for a monad $\mathcal{T}$ over the category $\mathcal{C}$ is a pointed endofunctor $(\mathbf{Tr}_{\mathcal{T}}, u)$ over the category of monads over $\mathcal{C}$, such that $\mathbf{Tr}_{\mathcal{T}}(\mathbf{Id}) = \mathcal{T}$. That is, an endofunctor together with a natural transformation

$$u : \mathbf{Id} \to \mathbf{Tr}_{\mathcal{T}} : \mathbf{Mnd}(\mathcal{C}) \to \mathbf{Mnd}(\mathcal{C})$$

In this framework, one considers $\mathbf{Tr}_{\mathcal{T}}(\mathcal{S})$ the "composite monad" of $\mathcal{T}$ and $\mathcal{S}$.

A *distributive law* of a monad $\mathcal{T} = (T, \eta^T, \mu^T)$ over the monad $\mathcal{S} = (S, \eta^S, \mu^S)$ is a natural transformation $\gamma : S \circ T \to T \circ S$, subject to appropriate axioms.

In this framework, one considers $\mathcal{T} \circ \mathcal{S}$ the composite monad (the axioms we glossed over above guarantee that there is a well-defined multiplication for it).

What we propose in this paper is a generalization of the latter notion, allowing for the treatment of a *class* of monads instead of just one such. Intuitively, one should think of that notion as a "coherent family" of distributive laws, in a sense that will be made precise in the next chapter. It also has the pleasant property of being easily generalized to the context of composing n monads (rather than just two), as we'll see later.

For a fixed monad $\mathcal{T}$, our notion will be equivalent to fixing a coherent family of *lifts* $\hat{\mathcal{S}}$ for monads $\mathcal{S}$ on $\mathcal{C}$, obtaining monads on $\mathcal{T}$'s Eilenberg-Moore category $\mathcal{C}^{\mathcal{T}}$, which correspond 1-1 with compositions (by composing the corresponding monadic adjunctions).

1 Monads in 2-category theory

The notion of a *monad over a category* $\mathcal{T} : \mathcal{C} \to \mathcal{C}$ is well-known to mathematicians (dating back to [God58] and [Hub61]; originally called *standard constructions* or *triples*) and computer scientists (the first reference we're aware of in this context is [Mog91]). We'll discuss a direct generalization, *monads in a 2-category* (first introduced in [Bén67], where the name "monad" is first used). We'll start by introducing basic notions, notations (notably, string diagram notation) and results about 2-categories; our treatment will closely follow Bénabou's [Bén67] and Leinster's [Lei98].

Definition 1.1. *2-category*
A 2-category $\mathcal{B}$ is given by

1. A class of *0-cells* or *objects*, denoted $X : \mathcal{B}$,

2. For any objects $X, Y : \mathcal{B}$, a (possibly large) *hom category* $\mathcal{B}[X, Y]$ whose

   - objects are called *1-cells* or *morphisms*

   - morphisms are called *2-cells* or *transformations*

   - composition $\circ$ is called *vertical* composition

3. For any object $X : \mathcal{B}$, a 1-cell $\mathbf{Id}_X : \mathcal{B}[X, X]$ (sometimes denoted by X)

4. For any objects $X, Y, Z : \mathcal{B}$ a *horizontal composition* functor

$$\cdot : \mathcal{B}[Y, Z] \times \mathcal{B}[X, Y] \to \mathcal{B}[X, Z]$$

5. For any objects $X, Y : \mathcal{B}$, *left* and *right unitors* natural isomorphisms

$$\lambda : (Y \cdot -) \to \mathbf{Id}_{\mathcal{B}[X,Y]} : \mathcal{B}[X, Y] \to \mathcal{B}[X, Y]$$

$$\rho : (- \cdot X) \to \mathbf{Id}_{\mathcal{B}[X,Y]} : \mathcal{B}[X, Y] \to \mathcal{B}[X, Y]$$

6. For any objects $W, X, Y, Z : \mathcal{B}$, an *associator* natural isomorphism

$$\alpha : ((- \cdot -) \cdot -) \to (- \cdot (- \cdot -)) : \mathcal{B}[Y, Z] \times \mathcal{B}[X, Y] \times \mathcal{B}[W, X] \to \mathcal{B}[W, Z]$$

satisfying the triangle and pentagon identities (whenever well-typed)

These axioms are to be read as ordinary commutative diagrams in the appropriate hom categories.

We say that a 2-category is *strict* if its unitors and associator are identities (so that horizontal composition is unital and associative *on the nose*). It is also worth stressing that what we call 2-categories are sometimes called *bicategories*, in which case the name *2-categories* is reserved for the strict ones (or, in select cases, to some semi-strict notion).

Remark 1.2. *Diagrammatic and string diagram notation*
There are two widespread ways of depicting cells and their compositions in a 2-category $\mathcal{B}$: *diagrammatic notation* and *string diagram notation*. We first give the example of a 2-cell $\phi : f \to g : X \to Y : \mathcal{C}$: its depiction would then be (diagrammatic notation is on the left, string diagram notation on the right)

In diagrammatic notation 0-cells are points, 1-cells are arrows between the points, and 2-cells are (directed) areas enclosed by arrows. String diagram notation is the dual (in the sense of Poincarè duality) of diagrammatic notation: 0-cells are areas, 1-cells are strings, and 2-cells are points. Moreover, identity 1-cells, unitors, and associators are (almost) never depicted in string diagram notation[1].

There are a lot of 2-categories which people are (more or less) used to dealing with:

Example 1.3. *Examples of 2-categories*

1. Most notions of a "structured category" yield a 2-category: **Cat** is the 2-category of categories, functors and natural transformations; **MonCat** is the 2-category of monoidal categories, monoidal functors and monoidal natural transformations; **BrMonCat** for braided such; **SymMonCat** for symmetric such and **CartMonCat** for cartesian monoidal categories[2]. We also have the 2-categories of (co-) complete categories, (co-) continuous functors and natural transformation, and so on. All these are examples of *strict* 2-categories.

2. Every category $\mathcal{C}$ can be seen as a strict 2-category **LDisc**($\mathcal{C}$) (the *locally discrete* 2-category generated by $\mathcal{C}$) with the same objects as $\mathcal{C}$, and whose hom-categories are the discrete categories on the hom-sets of $\mathcal{C}$; meaning **LDisc**($\mathcal{C}$)$[X, Y] := $ **Disc**($\mathcal{C}[X, Y]$). More concretely, this 2-category has the same objects and morphisms of the original category, and only identity 2-cells. From now on, we will abuse notation and omit any mention of **LDisc**, treating every category as the corresponding locally discrete 2-category.

3. Given a monoidal category $\mathbb{C} = (\mathcal{C}, \otimes, I, \dots)$ with (an explicit choice for) all binary coequalizers, and such that the tensor product preserves (the choices of) such coequalizers in each argument separately, we can define a 2-category **BiMod**($\mathbb{C}$) whose 0-cells are monoid objects in $\mathcal{C}$, 1-cells are bimodules, 2-cells are bimodule homomorphisms, identities are the monoids themselves seen as bimodules, and composition of 1-cells is given by tensoring bimodules along the monoid in the middle. The associators can be defined using the universal property of coequalizers, while unitors witness the statement that tensoring a bimodule by a monoid (seen as a bimodule over itself) is an equivalence. Under appropriate assumptions, the same can be done using comonoids and bi-comodules (assuming the existence of all binary equalizers and that the tensor product preserves them).

4. Given a cosmos $\mathbb{V}$ (in the sense of Bénabou [Str74]), there is a strict 2-category $\mathbb{V}$**Cat** with objects $\mathbb{V}$-enriched categories, hom-categories the

---

[1] When they are, they are depicted with dotted strings.

[2] Each of these comes in four distinct incarnation, using strict, weak, lax and colax monoidal functors as morphisms.

(ordinary, i.e. *not* **V**-enriched) categories of **V**-enriched functors and **V**-enriched natural transformations (with the obvious choices for identity 1-cells, unitors and associators). In the case $\mathbb{V} = \mathbf{Set}$ this recovers the familiar 2-category $\mathbf{Cat} = \mathbf{SetCat}$ of categories, functors and natural transformations.

5. Given a monoidal category $(\mathcal{C}, I, \otimes, \lambda, \rho, \alpha)$, we can construct a 2-category $\mathbb{B}\mathcal{C}$ with just one object $* : \mathbb{B}\mathcal{C}$, such that $\mathbb{B}\mathcal{C}[*, *] := \mathcal{C}$: 1-cells are objects in $\mathcal{C}$, 2-cells are morphisms in $\mathcal{C}$, the identity is $I : \mathcal{C}$ and horizontal composition is the tensor product $\otimes$. We call it the *delooping* of the monoidal category $\mathcal{C}$, and it is strict precisely when the monoidal structure on $\mathcal{C}$ is. In particular, the terminal category $\mathbb{1}$ admits a strict monoidal structure. Its delooping, which we still denote with $\mathbb{1} := \mathbb{B}\mathbb{1}$, is the terminal (strict) 2-category.

6. Given a category $\mathcal{C}$ with binary pullbacks, we can define the 2-category $\mathbf{Span}(\mathcal{C})$ whose 0-cells are $\mathcal{C}$'s objects, whose 1-cells $X \to Y$ are spans $X \leftarrow Z \to Y$, 2-cells are morphisms between the vertices of the related spans, identities given by identity spans $X \xleftarrow{\mathbf{Id}} X \xrightarrow{\mathbf{Id}} X$ and horizontal composition given by pulling back the spans.

7. Given a 2-category $\mathcal{B}$, we can define 2-categories $\mathcal{C}^{co}, \mathcal{C}^{op}, \mathcal{C}^{coop}$ by (resp.) reversing 2-cells, 1-cells or both. These are strict precisely when $\mathcal{B}$ is.

Remark 1.4. *Loop space of a pointed 2-category*
   Given an object X in a 2-category $\mathcal{B}$, the hom category $\mathcal{B}[X, X]$ is monoidal.

Definition 1.5. *Lax 2-functor*
   Given 2-categories $\mathcal{B}, \mathcal{C}$, a lax 2-functor $\mathcal{F} : \mathcal{B} \to \mathcal{C}$ is given by

1. For any object $X : \mathcal{B}$, an object $\mathcal{F}(X) : \mathcal{C}$,

2. For any objects $X, Y : \mathcal{B}$, a functor $\mathcal{F} : \mathcal{B}[X, Y] \to \mathcal{C}[\mathcal{F}(X), \mathcal{F}(Y)]$,

3. For any object $X : \mathcal{B}$, a 2-cell

$$\mathcal{F}_u : \mathbf{Id}_{\mathcal{F}(X)} \to \mathcal{F}(\mathbf{Id}_X)$$

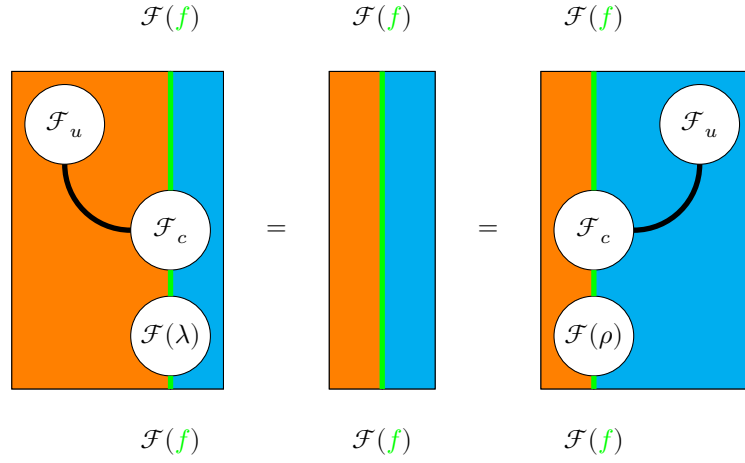4. For any $X \xrightarrow{f} Y \xrightarrow{g} Z : \mathcal{B}$, a 2-cell (natural in f, g)

$$\mathcal{F}_c : \mathcal{F}(g) \cdot \mathcal{F}(f) \to \mathcal{F}(g \cdot f)$$

satisfying the following axioms:

1. Unitality: for $f : X \to Y : \mathcal{B}$

5

$$\mathbf{Id}_{\mathcal{F}(Y)} \cdot \mathcal{F}(f) \xrightarrow{\ \lambda\ } \mathcal{F}(f) \xleftarrow{\ \rho\ } \mathcal{F}(f) \cdot \mathbf{Id}_{\mathcal{F}(X)}$$

$$\mathcal{F}_u \cdot \mathbf{Id}_{\mathcal{F}(f)} \downarrow \qquad\qquad \Big\| \qquad\qquad \downarrow \mathbf{Id}_{\mathcal{F}(f)} \cdot \mathcal{F}_u$$

$$\mathcal{F}(\mathbf{Id}_Y) \cdot \mathcal{F}(f) \qquad\qquad\qquad \mathcal{F}(f) \cdot \mathcal{F}(\mathbf{Id}_X)$$

$$\mathcal{F}_c \downarrow \qquad\qquad\qquad\qquad \downarrow \mathcal{F}_c$$

$$\mathcal{F}(\mathbf{Id}_Y \cdot f) \xrightarrow[\mathcal{F}(\lambda)]{} \mathcal{F}(f) \xleftarrow[\mathcal{F}(\rho)]{} \mathcal{F}(f \cdot \mathbf{Id}_X)$$

Or, in string diagram notation (where the black line is $\mathcal{F}(\mathbf{Id})$ and we don't depict identity 1-cells or coherences),

2. Associativity: for $W \xrightarrow{f} X \xrightarrow{g} Y \xrightarrow{h} Z : \mathcal{B}$

$$(\mathcal{F}(h) \cdot \mathcal{F}(g)) \cdot \mathcal{F}(f) \xrightarrow{\ \alpha\ } \mathcal{F}(h) \cdot (\mathcal{F}(g) \cdot \mathcal{F}(f))$$

$$\mathcal{F}_c \cdot \mathbf{Id}_{\mathcal{F}(f)} \downarrow \qquad\qquad\qquad \downarrow \mathbf{Id}_{\mathcal{F}(h)} \cdot \mathcal{F}_c$$

$$\mathcal{F}(h \cdot g) \cdot \mathcal{F}(f) \qquad\qquad \mathcal{F}(h) \cdot \mathcal{F}(g \cdot f)$$

$$\mathcal{F}_c \downarrow \qquad\qquad\qquad\qquad \downarrow \mathcal{F}_c$$

$$\mathcal{F}((h \cdot g) \cdot f) \xrightarrow[\mathcal{F}(\alpha)]{} \mathcal{F}(h \cdot (g \cdot f))$$

Which corresponds to the following string diagram equation

We say that

- a *strict 2-functor* $\mathcal{F} : \mathcal{B} \to \mathcal{C}$ is a lax 2-functor where $\mathcal{F}_u$ and $\mathcal{F}_c$ are all identities;

- a *(pseudo) 2-functor* $\mathcal{F} : \mathcal{B} \to \mathcal{C}$ is a lax 2-functor where $\mathcal{F}_u$ and $\mathcal{F}_c$ are all isomorphisms (we will omit the "pseudo" part of the name most of the times);

- a *colax 2-functor* $\mathcal{F} : \mathcal{B} \to \mathcal{C}$ is a lax 2-functor $\mathcal{F} : \mathcal{B} \to \mathcal{C}^{co}$ (i.e. if $\mathcal{F}_u$ and $\mathcal{F}_c$ have their sources and target flipped).

Example 1.6. *Examples of 2-functors*

1. Every ordinary functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ can be seen as a strict 2-functor **LDisc**$(\mathcal{F}) : \mathcal{C} \to \mathcal{D}$ in the obvious way: we map objects to objects as $\mathcal{F}$ did, and we do the same for 1-cells. There is no choice for 2-cells, and 2-functoriality follows immediately from $\mathcal{F}$'s functoriality. We will abuse notation and write $\mathcal{F}$ for **LDisc**$(\mathcal{F})$.

2. Monoidal functors preserving binary coequalizers[3] $\mathbb{F} : \mathbb{C} \to \mathbb{D}$ between monoidal categories as in point 3 in 1.3 induce 2-functors

   $$\mathbf{BiMod}(\mathbb{F}) : \mathbf{BiMod}(\mathbb{C}) \to \mathbf{BiMod}(\mathbb{D})$$

   which are as strict as $\mathbb{F}$ is.

3. Any lax monoidal functor $\mathcal{F} : \mathbb{V} \to \mathbb{W}$ between cosmoi induces a strict "change-of-base" 2-functor $\mathcal{F}_*$ (for the details, see prop. 6.4.3 in [Bor94], vol. 2).

4. Every monoidal functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ can be delooped to a 2-functor $\mathbb{B}\mathcal{F} : \mathbb{B}\mathcal{C} \to \mathbb{B}\mathcal{D}$ in the obvious way; it will be as strict as $\mathcal{F}$ is.

Definition 1.7. *Lax 2-natural transformation*

Given (strict, ordinary, lax or colax) 2-functors $\mathcal{F}, \mathcal{G} : \mathcal{B} \to \mathcal{C}$, a lax 2-natural transformation $\Phi : \mathcal{F} \to \mathcal{G}$ is given by

1. For any $X : \mathcal{B}$, a 1-cell $\Phi_X : \mathcal{F}(X) \to \mathcal{G}(X)$,

2. For any $f : X \to Y : \mathcal{B}$ an 2-cell $\Phi_f : \mathcal{G}(f) \cdot \Phi_X \to \Phi_Y \cdot \mathcal{F}(f)$ naturally in f, meaning that for every $\phi : f \to g : X \to Y : \mathcal{B}$ the following diagram is to be commutative in $\mathcal{C}[\mathcal{F}(X), \mathcal{G}(Y)]$

$$
\begin{array}{ccc}
\mathcal{G}(f) \cdot \Phi_X & \xrightarrow{\mathcal{G}(\phi) \cdot \Phi_X} & \mathcal{G}(g) \cdot \Phi_X \\
{\scriptstyle \Phi_f} \downarrow & & \downarrow {\scriptstyle \Phi_g} \\
\Phi_Y \cdot \mathcal{F}(f) & \xrightarrow{\Phi_Y \cdot \mathcal{F}(\phi)} & \Phi_Y \cdot \mathcal{F}(g)
\end{array}
$$

satisfying the following axioms:

1. Compatibility with unitors:

---

[3]It seems that if $\mathcal{F}$ is lax monoidal, there is no need for it to preserve coequalizer at all in order for it to induce a lax 2-functor.

$$\mathbf{Id}_{\mathcal{G}(X)} \cdot \Phi_X \xrightarrow{\mathcal{G}_u \cdot \mathbf{Id}_{\Phi_X}} \mathcal{G}(\mathbf{Id}_X) \cdot \Phi_X$$

$$\downarrow^{\lambda} \qquad\qquad\qquad \downarrow^{\Phi_{\mathbf{Id}_X}}$$

$$\Phi_X$$

$$\downarrow^{\rho^{-1}}$$

$$\Phi_X \cdot \mathbf{Id}_{\mathcal{F}(X)} \xrightarrow{\mathbf{Id}_{\Phi_X} \cdot \mathcal{F}_u} \Phi_X \cdot \mathcal{F}(\mathbf{Id}_X)$$

2. Compatibility with associators:

$$(\mathcal{G}(g) \cdot \mathcal{G}(f)) \cdot \Phi_X \xrightarrow{\mathcal{G}_c \cdot \mathbf{Id}_{\Phi_X}} \mathcal{G}(g \cdot f) \cdot \Phi_X$$

$$\downarrow^{\alpha}$$

$$\mathcal{G}(g) \cdot (\mathcal{G}(f) \cdot \Phi_X)$$

$$\downarrow^{\mathbf{Id}_{\mathcal{G}(g)} \cdot \Phi_f}$$

$$\mathcal{G}(g) \cdot (\Phi_Y \cdot \mathcal{F}(f))$$

$$\downarrow^{\alpha^{-1}} \qquad\qquad\qquad \downarrow^{\Phi_{g \cdot f}}$$

$$(\mathcal{G}(g) \cdot \Phi_Y) \cdot \mathcal{F}(f)$$

$$\downarrow^{\Phi_g \cdot \mathbf{Id}_{\mathcal{F}(f)}}$$

$$(\Phi_Z \cdot \mathcal{F}(g)) \cdot \mathcal{F}(f)$$

$$\downarrow^{\alpha}$$

$$\Phi_Z \cdot (\mathcal{F}(g) \cdot \mathcal{F}(f)) \xrightarrow{\mathbf{Id}_{\Phi_Z} \cdot \mathcal{F}_c} \Phi_Z \cdot \mathcal{F}(g \cdot f)$$

We say that

- a *strict 2-natural transformation* $\Phi : \mathcal{F} \to \mathcal{G}$ is a lax 2-natural transformation where all the Φ_f are identities;

- a *2-natural transformation* $\Phi : \mathcal{F} \to \mathcal{G}$ is a lax 2-natural transformation where all the Φ_f are isomorphisms;

- a *colax 2-natural transformation* $\Phi : \mathcal{F} \to \mathcal{G}$ is a lax 2-natural transformation $\Phi : \mathcal{F} \to \mathcal{G} : \mathcal{B} \to \mathcal{C}^{co}$ (i.e. if Φ_f has its source and target flipped).

Example 1.8. *Examples of 2-natural transformations*

- Every ordinary natural transformation $\phi : \mathcal{F} \to \mathcal{G} : \mathcal{C} \to \mathcal{D}$ can be seen as a strict 2-natural transformation $\mathbf{LDisc}(\phi) : \mathbf{LDisc}(\mathcal{F}) \to \mathbf{LDisc}(\mathcal{G})$ in the obvious way.

- Given monoidal categories $\mathbb{C}, \mathbb{D}$ with all binary coequalizers and functors $\mathbb{F}, \mathbb{G} : \mathbb{C} \to \mathbb{D}$ preserving them, a monoidal natural transformation $\phi : \mathbb{F} \to \mathbb{G}$ extends to a strict 2-transformation

$$\mathbf{BiMod}(\phi) : \mathbf{BiMod}(\mathbb{F}) \to \mathbf{BiMod}(\mathbb{G})$$

- Any monoidal natural transformation $\phi : \mathcal{F} \to \mathcal{G} : \mathcal{C} \to \mathcal{D}$ between (lax) monoidal functors induces a strict 2-natural transformation between the induced change-of-base strict 2-functors (see 4.3.1 in [Cru08] for the details).

- Any monoidal natural transformation between (strict, weak, co/lax) monoidal functors gives rise to a 2-natural transformation between the corresponding deloopings.

Definition 1.9. *Modification*

Given 2-natural transformations $\Phi, \Psi : \mathcal{F} \to \mathcal{G} : \mathcal{B} \to \mathcal{C}$, a modification

$$\Gamma : \Phi \to \Psi : \mathcal{F} \to \mathcal{G} : \mathcal{B} \to \mathcal{C}$$

consists of the datum of a 2-cell Γ_X for any $X : \mathcal{B}$, such that for any $f : X \to Y : \mathcal{B}$, the following diagram commutes

$$
\begin{array}{ccc}
\mathcal{G}(f) \cdot \Phi_X & \xrightarrow{\;\;\Phi_f\;\;} & \Phi_Y \cdot \mathcal{F}(f) \\
{\scriptstyle \mathbf{Id}_{\mathcal{G}(f)} \cdot \Gamma_X} \Big\downarrow & & \Big\downarrow {\scriptstyle \Gamma_Y \cdot \mathbf{Id}_{\mathcal{F}(f)}} \\
\mathcal{G}(f) \cdot \Psi_X & \xrightarrow[\;\;\Psi_f\;\;]{} & \Psi_Y \cdot \mathcal{F}(f)
\end{array}
$$

As one might have guessed, all these gadgets (2-functors, 2-natural transformations, 2-modifications) make up 2-categories and (sometimes) compose appropriately. We will not get into the details, just stating the following result without any proofs.

Proposition 1.10. *2-functor 2-categories*

Given 2-categories $\mathcal{B}, \mathcal{C}$ there are 2-categories

1. Given $\mathbf{St_{St}}[\mathcal{B}, \mathcal{C}]$ made of strict 2-functors, strict 2-natural transformations and modifications,

2. $\mathbf{St_{Ps}}[\mathcal{B}, \mathcal{C}]$ made of strict 2-functors, 2-natural transformations and modifications,

3. $\mathbf{St_{Lax}}[\mathcal{B}, \mathcal{C}]$ made of strict 2-functors, lax 2-natural transformations and modifications,

4. $\mathbf{St_{CoLax}}[\mathcal{B}, \mathcal{C}]$ made of strict 2-functors, colax 2-natural transformations and modifications,

5. $\mathbf{Ps_{Ps}}[\mathcal{B}, \mathcal{C}]$ made of 2-functors, 2-natural transformations and modifications,

6. $\mathbf{Ps_{Lax}}[\mathcal{B}, \mathcal{C}]$ made of 2-functors, lax 2-natural transformations and modifications,

7. $\mathbf{Ps_{CoLax}}[\mathcal{B}, \mathcal{C}]$ made of 2-functors, colax 2-natural transformations and modifications,

8. $\mathbf{Lax}[\mathcal{B}, \mathcal{C}]$ made of lax 2-functors, lax 2-natural transformations and modifications,

9. $\mathbf{CoLax}[\mathcal{B}, \mathcal{C}]$ made of colax 2-functors, colax 2-natural transformations and modifications.

We also get composition 2-functors, but only in cases (1, 2, 5). Moreover, the above list is *not* complete: we can consider, e.g., lax 2-functors and 2-natural transformations, or colax 2-functors and lax 2-natural transformations, etc. Finally, a note about notation: the use of $\mathbf{Ps}$ is motivated by the widespread terminology calling pseudofunctors what we just call 2-functors.

In what follows we'll mostly be interested in (1, 3, 8).

Definition 1.11. *Monad in a 2-category*

Given a 2-category $\mathcal{B}$, a monad in it is given by the data of a 0-cell X and a monoid object in $\mathcal{B}[X, X]$. More explicitly, this means a 4-tuple

$$(X, t : X \to X, \eta : \mathbf{Id}_X \to t, \mu : t \cdot t \to t)$$

such that the following diagrams commute in $\mathcal{B}[X, X]$, encoding unitality and associativity (resp.)

Analogously, we can define a comonad in a 2-category $\mathcal{B}$ as a monad in $\mathcal{B}^{co}$. It's easy to see that 2-functors interact nicely with monads and comonads

Proposition 1.12. *Lax 2-functors preserve monads*

Let $\mathcal{F} : \mathcal{B} \to \mathcal{C}$ be a lax 2-functor, and (X, t, η, μ) be a monad in $\mathcal{B}$. Then we get a monad in $\mathcal{C}$

$$\mathcal{F}(X, t, \eta, \mu) = (\mathcal{F}(X), \mathcal{F}(t), \mathcal{F}(\eta) \circ \mathcal{F}_u, \mathcal{F}(\mu) \circ \mathcal{F}_c)$$

Proof. Start with unitality:

10

The pentagons on the top commute by the unitality assumption on the 2-functor $\mathcal{F}$, the squares at the extreme left and right by naturality of $\mathcal{F}_c$, while the remaining two squares are the unitality axioms on the monad (X, t, η, μ).

$$
\begin{array}{ccc}
(\mathcal{F}(t) \cdot \mathcal{F}(t)) \cdot \mathcal{F}(t) & \xrightarrow{\quad\alpha\quad} & \mathcal{F}(t) \cdot (\mathcal{F}(t) \cdot \mathcal{F}(t))
\end{array}
$$

The hexagon on top is the associativity assumption on the 2-functor $\mathcal{F}$, the squares on the left and right commute by naturality of $\mathcal{F}_c$ and the remaining pentagon is the associativity axiom on the monad. $\qquad\square$

An analogous result holds for comonads and colax 2-functors

Proposition 1.13. *Colax 2-functors preserve comonads* Let $\mathcal{F} : \mathcal{B} \to \mathcal{C}$ a colax 2-functor, and (X, t, ϵ, Δ) a comonad in $\mathcal{B}$. Then we get a comonad in $\mathcal{C}$

$$
\mathcal{F}(X, t, \epsilon, \Delta) := (\mathcal{F}(X), \mathcal{F}(t), \mathcal{F}_u \circ \mathcal{F}(\epsilon), \mathcal{F}_c \circ \mathcal{F}(\Delta))
$$

Proof. Identical to that in 1.12. $\qquad\square$

More is true: every monad can be encoded as a lax 2-functor

Proposition 1.14. *Monads as lax 2-functors*

Given a monad (X, t, η, μ) in a 2-category $\mathcal{B}$, there is a lax 2-functor that sends the unique monad $(*, \mathbf{Id}_*, \mathbf{Id}_{\mathbf{Id}_*}, \mathbf{Id}_{\mathbf{Id}_*})$ in $\mathbb{1}$ to (X, t, η, μ).

Proof. Clearly the lax 2-functor $\mathcal{F} : \mathbb{1} \to \mathcal{B}$ will send $*$ to $\mathcal{F}(*) = X$ and $\mathbf{Id}_*$ to $\mathcal{F}(\mathbf{Id}_*) = t$. In order for it to be a lax functor, we need 1-cells

$$
\mathcal{F}_u : \mathbf{Id}_{\mathcal{F}(*)} := \mathbf{Id}_X \xrightarrow{\eta} t =: \mathcal{F}(\mathbf{Id}_*)
$$

$$
\mathcal{F}_c : \mathcal{F}(\mathbf{Id}_*) \cdot \mathcal{F}(\mathbf{Id}_*) := t \cdot t \xrightarrow{\mu} t =: \mathcal{F}(\mathbf{Id}_*) = \mathcal{F}(\mathbf{Id}_* \cdot \mathbf{Id}_*)
$$

where the last equality follows from the fact that $\mathbb{1}$ is a strict 2-category. Since the only 1-cell in $\mathbb{1}$ is $\mathbf{Id}_*$, the lax 2-functor axioms become

$$
\begin{array}{ccccc}
\mathbf{Id}_X \cdot t & \xrightarrow{\lambda} & t & \xleftarrow{\rho^{-1}} & t \cdot \mathbf{Id}_X \\
{\scriptstyle \eta \cdot \mathbf{Id}_t}\Big\downarrow & & \Big\| & & \Big\downarrow{\scriptstyle \mathbf{Id}_t \cdot \eta} \\
t \cdot t & & & & t \cdot t \\
{\scriptstyle \mu}\Big\downarrow & & & & \Big\downarrow{\scriptstyle \mu} \\
t & = & t & = & t
\end{array}
\qquad
\begin{array}{ccc}
(t \cdot t) \cdot t & \xrightarrow{\alpha} & t \cdot (t \cdot t) \\
{\scriptstyle \mu \cdot \mathbf{Id}_t}\Big\downarrow & & \Big\downarrow{\scriptstyle \mathbf{Id}_t \cdot \mu} \\
t \cdot t & & t \cdot t \\
{\scriptstyle \mu}\Big\downarrow & & \Big\downarrow{\scriptstyle \mu} \\
t & = & t
\end{array}
$$

which are exactly the monad axioms. □

In total analogy, we can prove the following

Proposition 1.15. *Comonads as colax 2-functors*

Given a comonad (X, t, ϵ, Δ) in a 2-category $\mathcal{B}$ there's a colax 2-functor that sends the unique comonad $(*, \mathbf{Id}_*, \mathbf{Id}_{\mathbf{Id}_*}, \mathbf{Id}_{\mathbf{Id}_*})$ in $\mathbb{1}$ to (X, s, ϵ, Δ).

Proof. Identical to that in 1.14. □

These results suggest the following

Definition 1.16. *2-categories of monads and comonads*

Let $\mathcal{B}$ be a 2-category. We define the 2-categories

$$\mathbf{Mnd}(\mathcal{B}) := \mathbf{Lax}[\mathbb{1}, \mathcal{B}]$$

$$\mathbf{CoMnd}(\mathcal{B}) := \mathbf{CoLax}[\mathbb{1}, \mathcal{B}]$$

This is worth unpacking, and we'll take the chance to introduce a notation we'll use throughout the rest of this section: string diagrams for 2-categories. In string diagram notation 0-cells are areas, 1-cells are strings and 2-cells are nodes; this comes with two big advantages: there's no way to depict identities, unitors and associators (so there will be no coherence issues), and it'll allow us to reason graphically about complex commutative diagrams.

1. A monad $\mathcal{T} = (X, t, \eta, \mu)$ would be depicted as

   and the axioms would translate to

2. A 1-cell $\Phi = (f, \phi) : (X, t^X, \eta^X, \mu^X) \to (Y, t^Y, \eta^Y, \mu^Y)$ in $\mathbf{Mnd}(\mathcal{B})$ is given by a 1-cell $f : X \to Y$ and a 2-cell

$$
\begin{array}{ccc}
X & \xrightarrow{t^X} & X \\
{\scriptstyle f}\downarrow & {\scriptstyle \phi}\nearrow & \downarrow{\scriptstyle f} \\
Y & \xrightarrow{t^Y} & Y
\end{array}
$$

12

depicted as
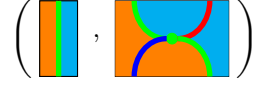
and since there's no non-trivial 2-cells in $\mathbb{1}$, naturality just means $\phi = \phi$; the two axioms translate to the commutativity of the following diagram

$$
\begin{array}{ccccc}
Y \cdot f & \xrightarrow{\eta^Y \cdot f} & t^Y \cdot f & \xleftarrow{\mu^Y \cdot f} & (t^Y \cdot t^Y) \cdot f \\
& & & & \downarrow{\scriptstyle \alpha} \\
& & & & t^Y \cdot (t^Y \cdot f) \\
\lambda \downarrow & & & & \downarrow{\scriptstyle t^Y \cdot \phi} \\
& & & & t^Y \cdot (f \cdot t^X) \\
f & & \phi & & \downarrow{\scriptstyle \alpha^{-1}} \\
& & & & (t^Y \cdot f) \cdot t^X \\
\rho^{-1} \downarrow & & & & \downarrow{\scriptstyle \phi \cdot t^X} \\
& & & & (f \cdot t^X) \cdot t^X \\
& & & & \downarrow{\scriptstyle \alpha} \\
f \cdot X & \xrightarrow[f \cdot \eta^X]{} & f \cdot t^X & \xleftarrow[f \cdot \mu^X]{} & f \cdot (t^X \cdot t^X)
\end{array}
$$

that is
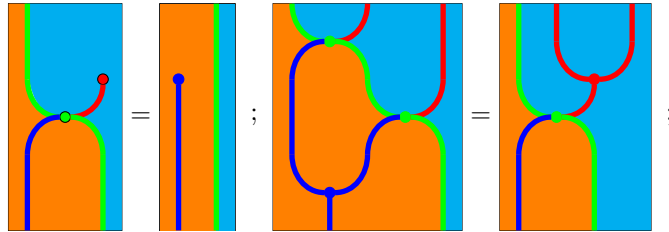
3. A 2-cell $\gamma : (f, \phi) \to (g, \psi) : (X, t^X, \eta^X, \mu^X) \to (Y, t^Y, \eta^Y, \mu^Y) : \mathbf{Mnd}(\mathcal{B})$ consists of a 2-cell $\gamma : f \to g : X \to Y : \mathcal{B}$, depicted as

13

such that

$$t^Y \cdot f \xrightarrow{\phi} f \cdot t^X$$
$$t^Y \cdot \gamma \downarrow \qquad \qquad \downarrow \gamma \cdot t^X$$
$$t^Y \cdot g \xrightarrow{\psi} g \cdot t^X$$

that is
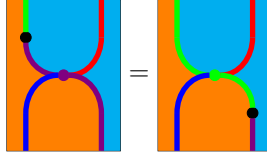
Another central notion in the theory of monads is that of a distributive law. This too can be internalized in an arbitrary 2-category

Definition 1.17. *Distributive law*

Let $\mathcal{B}$ be a 2-category and $(X, t_1, \eta_1, \mu_1), (X, t_2, \eta_2, \mu_2)$ be two monads on the same 0-cell. A distributive law

$$\gamma : (X, t_1, \eta_1, \mu_1) \rightsquigarrow (X, t_2, \eta_2, \mu_2)$$

is given by a 2-cell $\gamma : t_2 \cdot t_1 \to t_1 \cdot t_2$, depicted[4]
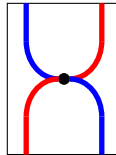
making the following diagrams commute

$$
\begin{array}{ccccc}
X \cdot t_1 & \xrightarrow{\eta_2 \cdot t_1} & t_2 \cdot t_1 & \xleftarrow{t_2 \cdot \eta_1} & t_2 \cdot X \\
\lambda^{-1} \uparrow & & \downarrow & & \uparrow \rho^{-1} \\
t_1 & & \gamma & & t_2 \\
\rho^{-1} \downarrow & & \downarrow & & \downarrow \lambda^{-1} \\
t_1 \cdot X & \xrightarrow[t_1 \cdot \eta_2]{} & t_1 \cdot t_2 & \xleftarrow[\eta_1 \cdot t_2]{} & X \cdot t_2
\end{array}
$$

---

[4]Anytime there's only one 0-cell involved, we'll depict it as a white background.

14

$$t_2 \cdot (t_1 \cdot t_1) \xrightarrow{t_2 \cdot \mu_1} t_2 \cdot t_1 \xleftarrow{\mu_2 \cdot t_1} (t_2 \cdot t_2) \cdot t_1$$

$$\alpha^{-1} \downarrow \qquad\qquad\qquad\qquad\qquad\qquad \downarrow \alpha$$

$$(t_2 \cdot t_1) \cdot t_1 \qquad\qquad\qquad\qquad t_2 \cdot (t_2 \cdot t_1)$$

$$\gamma \cdot t_1 \downarrow \qquad\qquad\qquad\qquad\qquad\qquad \downarrow t_2 \cdot \gamma$$

$$(t_1 \cdot t_2) \cdot t_1 \qquad\qquad\qquad\qquad t_2 \cdot (t_1 \cdot t_2)$$

$$\alpha \downarrow \qquad\qquad\qquad \gamma \qquad\qquad\qquad \downarrow \alpha^{-1}$$

$$t_1 \cdot (t_2 \cdot t_1) \qquad\qquad\qquad\qquad (t_2 \cdot t_1) \cdot t_2$$

$$t_1 \cdot \gamma \downarrow \qquad\qquad\qquad\qquad\qquad\qquad \downarrow \gamma \cdot t_2$$

$$t_1 \cdot (t_1 \cdot t_2) \qquad\qquad\qquad\qquad (t_1 \cdot t_2) \cdot t_2$$

$$\alpha^{-1} \downarrow \qquad\qquad\qquad\qquad\qquad\qquad \downarrow \alpha$$

$$(t_1 \cdot t_1) \cdot t_2 \xrightarrow[\mu_1 \cdot t_2]{} t_1 \cdot t_2 \xleftarrow[t_1 \cdot \mu_2]{} t_1 \cdot (t_2 \cdot t_2)$$

that is (the first two equalities encode the first diagram, the second two the second)
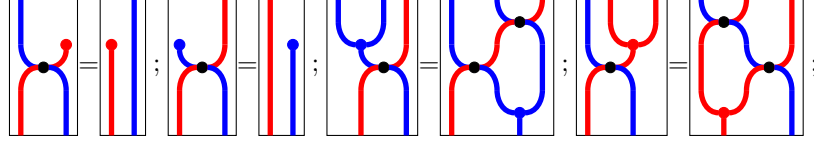
In the case of monads over categories (that is, monads in the 2-category of categories), it's well-known that distributive laws correspond 1-to-1 to lifts of one monad to the Eilenberg-Moore category for the other, allowing one to compose the two monads (for the precise statement, see the first proposition in [Bec69]). This fails in general 2-categories[5], but one of the two implications still holds in full generality

Proposition 1.18. *Distributive laws allow for monad composition*

In a 2-category $\mathcal{B}$, fix a distributive law between monads

$$\gamma : (X, t_1, \eta_1, \mu_1) \rightsquigarrow (X, t_2, \eta_2, \mu_2)$$

We can define a new monad $\mathcal{T} = (X, t, \eta, \mu)$ with $t = t_1 \cdot t_2$.

Proof. First, we define the monad $\mathcal{T}$

---
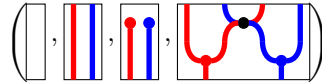
[5]Though under the appropriate assumptions, that is existence of Eilenberg-Moore objects, the same theorem still holds: the proof is almost identical to that in [Bec69]

Showing $\mathcal{T}$ is a monad amounts to proving unitality and associativity: using string diagrams this is straightforward. The following string diagram calculation proves unitality
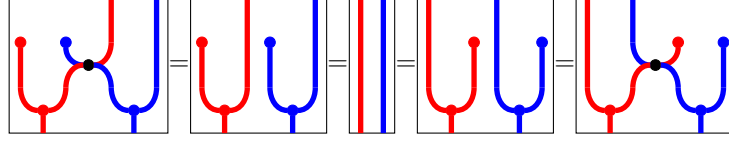
The first and last equalities follow from the axioms of a distributive law, the middle ones are unitality for the two monads. The next string diagram calculation proves associativity
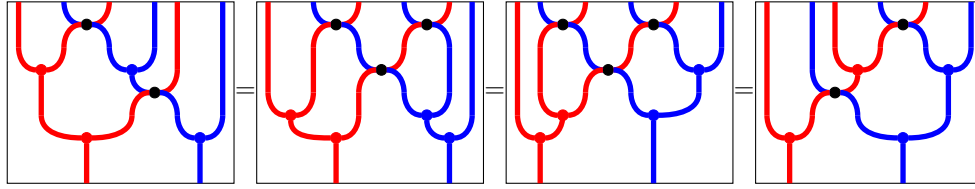
The first and last equalities follow from the axioms of a distributive law, the middle one is associativity for the two monads.

$\square$

As it turns out, distributive laws can be encoded as monads too

Proposition 1.19. *Distributive laws as monads*

The data of a distributive law $\gamma : (X, t, \eta^t, \mu^t) \rightsquigarrow (X, s, \eta^S, \mu^s)$ between two monads $\mathcal{T} = (X, t, \eta^t, \mu^t), \mathcal{S} = (X, s, \eta^s, \mu^s)$ over the same 0-cell $X : \mathcal{B}$ is equivalently encoded as a monad

$$\Gamma = ((X, t, \eta^t, \mu^t), (s, \gamma), \eta^s, \mu^s) : \mathbf{Mnd}(\mathbf{Mnd}(\mathcal{B}))$$

in the 2-category $\mathbf{Mnd}(\mathcal{B}) \simeq \mathbf{Lax}[\mathbb{1}, \mathcal{B}]$

Proof. The proof consists in just writing down the axioms for both and matching them:

1. the fact that (X, s, η^S, μ^S) is a monad follows from the fact that (s, γ) is a monad on (X, t, η^t, μ^t) with unit η^s and multiplication μ^s;

2. the fact that γ is a distributive law follows from the assumptions that (s, γ) is a 1-cell and η^s, μ^s are 2-cells in $\mathbf{Mnd}(\mathcal{B})$.

$\square$

This suggests we *define* the 2-category of distributive laws as "monads in monads"

Definition 1.20. *2-category of distributive laws*

We define the 2-category $\mathbf{Dist}(\mathcal{B}) := \mathbf{Mnd}(\mathbf{Mnd}(\mathcal{B}))$.

As we did for the 2-category $\mathbf{Mnd}(\mathcal{B})$, it is worth giving a concrete description of $\mathbf{Dist}(\mathcal{B})$

Remark 1.21. *2-category of distributive laws, explicitly*

Fix a (strict, for the sake of convenience) 2-category $\mathcal{B}$. The 2-category $\mathbf{Dist}(\mathcal{B})$ can be presented[6] in the following way:

1. its 0-cells are 8-tuples $(X, t, s, \eta^t, \eta^s, \mu^t, \mu^s, \sigma)$, with

   - $X : \mathcal{B}$
   - $t, s : X \to X : \mathcal{B}$
   - $\eta^t : \mathbf{Id}_X \to t : X \to X : \mathcal{B}$
   - $\eta^s : \mathbf{Id}_X \to s : X \to X : \mathcal{B}$
   - $\mu^t : t \cdot t \to t : X \to X : \mathcal{B}$
   - $\mu^s : s \cdot s \to s : X \to X : \mathcal{B}$
   - $\sigma : s \cdot t \to t \cdot s : X \to X : \mathcal{B}$

   satisfying the following axioms:

   (a) $(X, t, \eta^t, \mu^t) : \mathbf{Mnd}(\mathcal{B})$ is a monad:

$$
\begin{array}{ccccc}
X \cdot t & =\!=\!= & t & =\!=\!= & t \cdot X \\
{\scriptstyle \eta^t \cdot t}\downarrow & & \| & & \downarrow{\scriptstyle t \cdot \eta^t} \\
t \cdot t & \xrightarrow{\ \mu^t\ } & t & \xleftarrow{\ \mu^t\ } & t \cdot t
\end{array}
\qquad
\begin{array}{ccc}
t \cdot t \cdot t & \xrightarrow{t \cdot \mu^t} & t \cdot t \\
{\scriptstyle \mu^t \cdot t}\downarrow & & \downarrow{\scriptstyle \mu^t} \\
t \cdot t & \xrightarrow{\ \mu^t\ } & t
\end{array}
$$

   (b) $(X, s, \eta^s, \mu^s) : \mathbf{Mnd}(\mathcal{B})$ is a monad:

$$
\begin{array}{ccccc}
X \cdot s & =\!=\!= & s & =\!=\!= & s \cdot X \\
{\scriptstyle \eta^s \cdot s}\downarrow & & \| & & \downarrow{\scriptstyle s \cdot \eta^s} \\
s \cdot s & \xrightarrow{\ \mu^s\ } & s & \xleftarrow{\ \mu^s\ } & s \cdot s
\end{array}
\qquad
\begin{array}{ccc}
s \cdot s \cdot s & \xrightarrow{s \cdot \mu^s} & s \cdot s \\
{\scriptstyle \mu^s \cdot s}\downarrow & & \downarrow{\scriptstyle \mu^s} \\
s \cdot s & \xrightarrow{\ \mu^s\ } & s
\end{array}
$$

   (c) $\sigma : (X, t, \eta^t, \mu^t) \rightsquigarrow (X, s, \eta^s, \mu^s)$ is a distributive law:

$$
\begin{array}{ccccc}
X \cdot t & \xrightarrow{\eta^s \cdot t} & s \cdot t & \xleftarrow{s \cdot \eta^t} & s \cdot X \\
\| & & \downarrow{\scriptstyle \sigma} & & \| \\
t & & & & s \\
\| & & \downarrow & & \| \\
t \cdot X & \xrightarrow{t \cdot \eta^s} & t \cdot s & \xleftarrow{\eta^t \cdot s} & X \cdot s
\end{array}
\qquad
\begin{array}{ccccc}
s \cdot s \cdot t & \xrightarrow{\mu^s \cdot t} & s \cdot t & \xleftarrow{s \cdot \mu^t} & s \cdot t \cdot t \\
{\scriptstyle s \cdot \sigma}\downarrow & & & & \downarrow{\scriptstyle \sigma \cdot t} \\
s \cdot t \cdot s & & \downarrow{\scriptstyle \sigma} & & t \cdot s \cdot t \\
{\scriptstyle \sigma \cdot s}\downarrow & & & & \downarrow{\scriptstyle t \cdot \sigma} \\
t \cdot s \cdot s & \xrightarrow{t \cdot \mu^s} & t \cdot s & \xleftarrow{\mu^t \cdot s} & t \cdot t \cdot s
\end{array}
$$

---

[6]Meaning $\mathbf{Dist}(\mathcal{B})$ it is strictly equivalent to the 2-category being defined: there's strict 2-functors back-and-forth (defined in the obvious way), which compose to the identity 2-functor on the nose.

2. its 1-cells are triples $(f, \phi, \psi) : (X, t, s, \eta^t, \eta^s, \mu^t, \mu^s, \sigma) \to (Y, r, u, \eta^r, \eta^u, \mu^r, \mu^u, \rho)$ where

- $f : X \to Y : \mathcal{B}$
- $\phi : f \cdot t \to r \cdot f$
- $\psi : f \cdot s \to u \cdot f$

satisfying the following axioms:

(a) $(f, \phi) : (X, t, \eta^t, \mu^t) \to (Y, r, \eta^r, \mu^r) : \mathbf{Mnd}(\mathcal{B})$

$$
\begin{array}{ccccc}
f \cdot X & \xrightarrow{\ f \cdot \eta^t\ } & f \cdot t & \xleftarrow{\ f \cdot \mu^t\ } & f \cdot t \cdot t \\[4pt]
\| & & \Big\downarrow \phi & & \Big\downarrow \phi \cdot t \\[4pt]
f & & & & r \cdot f \cdot t \\[4pt]
\| & & & & \Big\downarrow r \cdot \phi \\[4pt]
Y \cdot f & \xrightarrow{\ \eta^r \cdot f\ } & r \cdot f & \xleftarrow{\ \mu^r \cdot f\ } & r \cdot r \cdot f
\end{array}
$$

(b) $(f, \psi) : (X, s, \eta^s, \mu^s) \to (Y, u, \eta^u, \mu^u) : \mathbf{Mnd}(\mathcal{B})$

$$
\begin{array}{ccccc}
f \cdot X & \xrightarrow{\ f \cdot \eta^s\ } & f \cdot s & \xleftarrow{\ f \cdot \mu^s\ } & f \cdot s \cdot s \\[4pt]
\| & & \Big\downarrow \psi & & \Big\downarrow \psi \cdot s \\[4pt]
f & & & & u \cdot f \cdot s \\[4pt]
\| & & & & \Big\downarrow u \cdot \psi \\[4pt]
Y \cdot f & \xrightarrow{\ \eta^u \cdot f\ } & u \cdot f & \xleftarrow{\ \mu^u \cdot f\ } & u \cdot u \cdot f
\end{array}
$$

(c) $\psi : (f, \phi) \cdot (s, \sigma) \to (u, \rho) \cdot (f, \phi)$ is a 2-cell in $\mathbf{Mnd}(\mathcal{C})$

$$
\begin{array}{ccc}
f \cdot s \cdot t & \xrightarrow{\ f \cdot \sigma\ } & f \cdot t \cdot s \\[4pt]
\Big\downarrow \psi \cdot t & & \Big\downarrow \phi \cdot s \\[4pt]
u \cdot f \cdot t & & r \cdot f \cdot s \\[4pt]
\Big\downarrow u \cdot \phi & & \Big\downarrow r \cdot \psi \\[4pt]
u \cdot r \cdot f & \xrightarrow{\ \rho \cdot f\ } & r \cdot u \cdot f
\end{array}
$$

3. given composable 1-cells (f_1, ϕ_1, ψ_1), (f_2, ϕ_2, ψ_2) their composition is given by

$$(f_2, \phi_2, \psi_2) \cdot (f_1, \phi_1, \psi_1) := (f_2 \cdot f_1, (\phi_2 \cdot f_1) \circ (f_2 \cdot \phi_1), (\psi_2 \cdot f_1) \circ (f_2 \cdot \psi_1))$$

4. its 2-cells $\gamma : (f, \phi_1, \psi_1) \to (g, \phi_2, \psi_2)$ are 2-cells $\gamma : f \to g : X \to Y : \mathcal{B}$ satisfying the following axioms:

(a) $\gamma : (f, \phi_1) \to (g, \phi_2) : (X, t, \eta^t, \mu^t) \to (Y, r, \eta^r, \mu^r) : \mathbf{Mnd}(\mathcal{B})$

$$f \cdot t \xrightarrow{\phi_1} r \cdot f$$
$$\downarrow{\gamma \cdot t} \qquad \downarrow{r \cdot \gamma}$$
$$g \cdot t \xrightarrow{\phi_2} r \cdot g$$

(b) $\gamma : (f, \psi_1) \to (g, \psi_2) : (X, s, \eta^s, \mu^s) \to (Y, u, \eta^u, \mu^u) : \mathbf{Mnd}(\mathcal{B})$

$$f \cdot s \xrightarrow{\psi_1} u \cdot f$$
$$\downarrow{\gamma \cdot s} \qquad \downarrow{u \cdot \gamma}$$
$$g \cdot s \xrightarrow{\psi_2} u \cdot g$$

5. given vertically composable 2-cells γ_1, γ_2 their vertical composite is given by $\gamma_2 \circ \gamma_1$

6. given horizontally composable 2-cells γ_1, γ_2 their horizontal composite is given by $\gamma_2 \cdot \gamma_1$

Proof. Immediate. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

We can therefore define 2-functors $\mathbf{Dist}(\mathcal{B}) \to \mathbf{Mnd}(\mathcal{B})$, encoding the recovery of the underlying monads (as well as the composite we described in 1.18):

Remark 1.22. *Monads from distributive laws*

Given a (strict, for the sake of conciseness) 2-category $\mathcal{B}$, we have strict 2-functors

$$\mathcal{U}_1, \mathcal{U}_2, \mathcal{C} : \mathbf{Dist}(\mathcal{B}) \to \mathbf{Mnd}(\mathcal{B})$$

defined as follows:

$$\mathcal{U}_1((a, t, \eta^t, \mu^t), (s, \sigma), \eta^s, \mu^s) = (a, t, \eta^t, \mu^t)$$

$$\mathcal{U}_2((a, t, \eta^t, \mu^t), (s, \sigma), \eta^s, \mu^s) = (a, s, \eta^s, \mu^s)$$

$$\mathcal{C}((a, t, \eta^t, \mu^t), (s, \sigma), \eta^s, \mu^s) = (a, s \cdot t, \eta^s \cdot \eta^t, (s \cdot \sigma \cdot t) \circ (\mu^s \cdot \mu^t))$$

Proof. We need to complete the definition of $\mathcal{U}_1, \mathcal{U}_2, \mathcal{C}$ to 2-functors: give the actions on 1- and 2-cells, and prove that makes them into 2-functors. Fix a 1-cell in $\mathbf{Dist}(\mathcal{B})$

$$((f, \phi), \psi) : ((a_1, t_1, \eta_1^t, \mu_1^t), (s_1, \sigma_1), \eta_1^s, \mu_1^s) \to ((a_2, t_2, \eta_2^t, \mu_2^t), (s_2, \sigma_2), \eta_2^s, \mu_2^s)$$

We then define $\mathcal{U}_1((f, \phi), \psi) := (f, \phi)$, $\mathcal{U}_2((f, \phi), \psi) = (f, \psi)$ and $\mathcal{C}((f, \phi), \psi) = (f, (s_2 \cdot \phi) \circ (\psi \cdot t_1))$. Similarly, fix a 2-cell $\gamma : ((f_1, \phi_1), \psi_1) \to ((f_2, \phi_2), \psi_2)$: we (unsurprisingly) define $\mathcal{U}_1(\gamma) = \mathcal{U}_2(\gamma) = \mathcal{C}(\gamma) = \gamma$. Moreover, the images of 1-cells (resp. 2-cells) are 1-cells (2-cells) in $\mathbf{Mnd}(\mathcal{B})$: the argument for $\mathcal{U}_i$ follows immediately from our previous remark 1.21 (2.a and 4.a for $\mathcal{U}_1$, 2.b and 4.b for $\mathcal{U}_2$ and 2.c), while that for $\mathcal{C}$ follows from 2.a - 2.c and 4.a - 4.b.

That $\mathcal{U}_1$ and $\mathcal{U}_2$ are strict 2-functors is an immediate consequence of our previous remark, 1.21. We will sketch the proof for $\mathcal{U}_1$, as the one for $\mathcal{U}_2$ is identical. Fix composable 1-cells

$$T_1 \xrightarrow{F_1=((f_1,\phi_1),\psi_1)} T_2 \xrightarrow{F_2=((f_2,\phi_2),\psi_2)} T_3$$

We want to show $\mathcal{U}_1(F_2 \cdot F_1) = \mathcal{U}_1(F_2) \cdot \mathcal{U}_1(F_2)$. Recall that

$$F_2 \cdot F_1 = ((f_2 \cdot f_1, (\phi_2 \cdot f_1) \circ (f_2 \cdot \phi_1)), (\psi_2 \cdot f_1) \circ (f_2 \cdot \psi_1))$$

so that $\mathcal{U}_1(F_2 \cdot F_1) = (f_2 \cdot f_1, (\phi_2 \cdot f_1) \circ (f_2 \cdot \phi_1))$. On the other hand, we have $\mathcal{U}_1(F_1) = (f_1, \phi_1)$ and $\mathcal{U}_1(F_2) = (f_2, \phi_2)$, implying

$$\mathcal{U}_1(F_2) \cdot \mathcal{U}_1(F_1) = (f_2, \phi_2) \cdot (f_1, \phi_1) = (f_2 \cdot f_1, (\phi_2 \cdot f_1) \circ (f_2 \cdot \phi_1))$$

Similarly, we can see that it sends identity 1-cells to identity 1-cells. Likewise, it is trivial to check that it preserves identity and (vertical, horizontal) composites of 2-cells: $\mathcal{U}_1$ is a strict 2-functor.

The case for $\mathcal{C}$ is slightly less immediate: consider composable 1-cells F_1, F_2 as before. We can compute

$$\mathcal{C}(F_2) \cdot \mathcal{C}(F_1) = (f_2, (s_3 \cdot \phi_2) \circ (\psi_2 \cdot t_2)) \cdot (f_1, (s_2 \cdot \phi_1) \circ (\psi_1 \cdot t_1))$$
$$= (f_2 \cdot f_1, (s_2 \cdot \phi_2 \cdot f_1) \circ (\psi_2 \cdot t_2 \cdot f_1) \circ (f_2 \cdot s_2 \cdot \phi_1) \circ (f_2 \cdot \psi_1 \cdot t_1))$$

$$\mathcal{C}(F_2 \cdot F_1) = \mathcal{C}((f_2 \cdot f_1, (\phi_2 \cdot f_1) \circ (f_2 \cdot \phi_1)), (\psi_2 \cdot f_1) \circ (f_2 \cdot \psi_1))$$
$$= (f_2 \cdot f_1, (s_2 \cdot \phi_2 \cdot f_1) \circ (s_2 \cdot f_2 \cdot \phi_1) \circ (\psi_2 \cdot f_1 \cdot t_1) \circ (f_2 \cdot \psi_1 \cdot t_1))$$

The first components are clearly equal; the second components are equal because of interchange. Showing that identity 1-cells are also preserved on the nose is immediate, as is showing identity and (vertical, horizontal) composites of 2-cells are preserved. Hence $\mathcal{C}$ is a strict 2-functor. $\qquad\square$

This allows us to define what a parametric distributive law is:

Definition 1.23. *Parametric Distributive Law*

Given 2-categories $\mathcal{B}$ and $\mathcal{C}$, the category of $\mathcal{C}$-parametric distributive laws in $\mathcal{B}$ is the 2-functor 2-category

$$\mathbf{PDist}(\mathcal{C}, \mathcal{B}) := \mathbf{St_{St}}[\mathcal{C}, \mathbf{Dist}(\mathcal{B})]$$

An immediate consequence of 1.19 is then

$$\mathbf{PDist}(\mathcal{C}, \mathcal{B}) = \mathbf{St_{St}}[\mathcal{C}, \mathbf{Lax}[\mathbb{1}, \mathbf{Lax}[\mathbb{1}, \mathcal{B}]]]$$

The absence of a well-behaved left 2-adjoint to the **Lax** 2-functor makes this somewhat annoying to deal with; the main purpose of the next section is to alleviate this issue.

2 The lax functor classifier

In this section we will present a construction that is known to experts in the field[7], but for which we are not aware of any reference: the "lax functor classifier". Given a 2-category $\mathcal{B}$, its lax functor classifier $\widehat{\mathcal{B}}$ is a 2-category such that, for any other 2-category $\mathcal{C}$, we have an equivalence

$$\mathbf{Lax}[\mathcal{B}, \mathcal{C}] \simeq \mathbf{Ps_{Lax}}[\widehat{\mathcal{B}}, \mathcal{C}]$$

It is reminiscent of a related construction, sometimes called the "lax morphism classifier", introduced by Blackwell, Kelly and Power in [BKP89]. In order to give a definition that is not too combinatorial in nature, we shall first recall the definition of the (augmented) simplex and interval categories, together with the duality that relates them.

Remark 2.1. *Simplex and interval categories*

Recall that the category of preorders **PreOrd** is a full subcategory of the category of categories **Cat**:

$$\mathbf{PreOrd} \hookrightarrow \mathbf{Cat}$$

The augmented simplex category, denoted by $\Delta_+ \hookrightarrow \mathbf{PreOrd}$ is the full subcategory of finite, totally ordered sets and the simplex category, denoted by $\Delta \hookrightarrow \Delta_+$ is the full subcategory of *non-empty* such. We will employ the following, widespread conventions:

1. $\forall n \in \mathbb{N}$, write $[n-1] := \{0 \leq ... \leq n-1\} : \Delta_+$ for the[8] poset with n elements;

2. $\forall n \in \mathbb{N}$ and $0 \leq k \leq n$, write $\delta_k^n : [n-1] \to [n]$ for the unique injective, monotonic function such that $\delta_k^n(k) = k+1$;

3. $\forall n \in \mathbb{N}$ and $0 \leq k \leq n$, write $\sigma_k^n : [n+1] \to [n]$ for the unique surjective, monotonic function such that $\sigma_k^n(k) = k = \sigma_k^n(k+1)$.

Furthermore the augmented interval category, denoted by $\mathbf{I}_+ \hookrightarrow \mathbf{PreOrd}$ is the subcategory of finite intervals (that is, totally ordered sets with *not necessarily distinct* top and bottom element) and interval maps (monotonic functions preserving the top and bottom elements), and the interval category, denoted by $\mathbf{I} \hookrightarrow \mathbf{I}_+$ is for the full subcategory on objects with *distinct* top and bottom elements.

Finally, recall that $\Delta_+^{op} \simeq \mathbf{I}_+$, the equivalence being induced by the hom-functors

$$\Delta_+[-, [1]] : \Delta_+^{op} \to \mathbf{I}_+$$

$$\mathbf{I}_+[-, [1]] : \mathbf{I}_+^{op} \to \Delta_+$$

---

[7]E.g. as described in T. Johnson-Freyd's post [httb]

[8]Unique up to isomorphism in **PreOrd**

Given $k \in \{-1, 0, \dots\}$, the poset structure on $\Delta_+[[k], [1]] = \{f : [k] \to [1] : \Delta_+\}$ is the usual, pointwise preorder (which happens to be a total order because the codomain is $[1]$), with top and bottom elements being the constant functions on 1 and 0 respectively. Analogously, given $k \in \mathbb{N}$, the poset structure on $\mathbf{I}_+[[k], [1]] = \{f : [k] \to [1] : \mathbf{I}_+\}$ is pointwise. Finally, recall that this equivalence restricts to an equivalence $\Delta^{op} \simeq \mathbf{I}$.

In what follows we will keep writing $[k]$ for both the (posetal) category and its corresponding locally discrete 2-category.

Definition 2.2. *The lax functor classifier*

The lax functor classifier $\widehat{\mathcal{B}}$ of a 2-category $\mathcal{B}$ is defined as follows:

1. Its 0-cells are the same as those of $\mathcal{B}$;

2. Its 1-cells $(k, \mathcal{F}) : a \to b : \widehat{\mathcal{B}}$ are pairs where $k \in \mathbb{N}$, $\mathcal{F} : [k] \to \mathcal{B}$ is a semistrict pseudofunctor (meaning it preserves identity 1-cells *on the nose*), and $\mathcal{F}(0) = a$, $\mathcal{F}(k) = b$;

3. Identity 1-cells $\mathbb{1}_a := (0, !_a) : a \to a : \widehat{\mathcal{B}}$ are given by the strict 2-functors $!_a : [0] \to \mathcal{B}$ with $!_a(0) = a$ (and hence $!_a(\mathbb{1}_0) = \mathbb{1}_a$);

4. Its 2-cells $(\phi, \alpha) : (m, \mathcal{F}) \to (n, \mathcal{G}) : a \to b : \widehat{\mathcal{B}}$ are pairs where $\alpha : [n] \to [m] : \mathbf{I}_+$, $\mathcal{F} \circ \alpha$ and $\mathcal{G}$ agree on objects and $\phi : \mathcal{F} \circ \alpha \to \mathcal{G}$ is an ICON;

5. Identity 2-cells $\mathbb{1}_{(m, \mathcal{F})} : (m, \mathcal{F}) \to (m, \mathcal{F})$ are the 2-cells $(\mathbb{1}_{\mathcal{F}}, \mathbb{1}_{[m]})$

6. Given $(\phi, \alpha) : (m, \mathcal{G}) \to (n, \mathcal{H})$ and $(\psi, \beta) : (l, \mathcal{F}) \to (m, \mathcal{G})$, their vertical composite is defined by

$$(\phi, \alpha) \circ (\psi, \beta) := (\phi \circ (\psi \alpha), \beta \circ \alpha) : (l, \mathcal{F}) \to (n, \mathcal{H})$$

where $\psi \alpha : \mathcal{F} \circ \beta \circ \alpha \to \mathcal{G} \circ \alpha$ denotes the ICON obtained by whiskering ψ on the right by the (strict) 2-functor α.

7. Given $(n, \mathcal{G}) : b \to c$ and $(m, \mathcal{F}) : a \to b$, their (horizontal) composite is defined by

$$(n, \mathcal{G}) \cdot (m, \mathcal{F}) := (n + m, \mathcal{G} \star \mathcal{F})$$

where $\mathcal{G} \star \mathcal{F}$ is given by 2-categorical Day convolution with respect to (the dual of) the monoidal structure $(- \vee -)$ on $\mathbf{I}_+$ obtained by transferring ordinal sum over the equivalence $\Delta_+^{op} \simeq \mathbf{I}_+$. Explicitly, define on objects

$$(\mathcal{G} \star \mathcal{F})(k) := \begin{cases} \mathcal{F}(k) & \text{if } k \leq m \\ \mathcal{G}(k - m) & \text{if } m \leq k \end{cases}$$

(which is well-defined because $\mathcal{F}(k) = b = \mathcal{G}(0)$) and on morphisms $f : i \to j : [m + n]$

$$(\mathcal{G} \star \mathcal{F}) := \begin{cases} \mathcal{F}(f) & \text{if } j < m \\ \mathcal{G}(f) & \text{if } m < i \\ \mathcal{F}(f_1) \cdot \mathcal{G}(f_2) & \text{if } i \leq m \leq j, \text{ hence } f : i \xrightarrow{f_2} m \xrightarrow{f_1} j \end{cases}$$

22

8. Given
$$(\psi, \beta) : (n_1, \mathcal{G}_1) \to (n_2, \mathcal{G}_2) : b \to c$$
$$(\phi, \alpha) : (m_1, \mathcal{F}_1) \to (m_2, \mathcal{F}_2) : a \to b$$

define their horizontal composite by $(\psi, \beta) \cdot (\phi, \alpha) := (\psi \star \phi, \beta \vee \alpha)$ where $(- \vee -)$ is the aforementioned monoidal structure on $\mathbf{I}_+$, and $\psi \star \phi$ is Day convolution of ICONs. Explicitly, define $\beta \vee \alpha$ as:

$$(\beta \vee \alpha)(k) := \begin{cases} \alpha(k) & \text{if } k \leq m_1 \\ \beta(k - m_1) + n_1 & \text{if } m_1 \leq k \end{cases}$$

(which is well-defined since $\alpha(m_1) = n_1$ and $\beta(0) = 0$) and $\psi \star \phi$, for any $f : i \to j : [m_1 + m_2]$

$$(\psi \star \phi)_f := \begin{cases} \phi_f & \text{if } j < m_1 \\ \psi_f & \text{if } m_1 < i \\ \phi_{f_1} \cdot \psi_{f_2} & \text{if } i \leq m_1 \leq j, \text{ hence } f : i \xrightarrow{f_2} m_1 \xrightarrow{f_1} j \end{cases}$$

9. The left and right unitors are defined as follows

$$\lambda_{(k,\mathcal{F})} := (\lambda^{\mathcal{F}}, \mathbb{1}_{[k]}) : \mathbb{1}_b \cdot (k, \mathcal{F}) \to (k, \mathcal{F})$$

$$\rho_{(k,\mathcal{F})} := (\rho^{\mathcal{F}}, \mathbb{1}_{[k]}) : (k, \mathcal{F}) \cdot \mathbb{1}_a \to (k, \mathcal{F})$$

with $\lambda^{\mathcal{F}}, \rho^{\mathcal{F}}$ being the ICONs with, for all $f : i \to j : [k]$:

$$\lambda_f^{\mathcal{F}} := \begin{cases} \mathbb{1}_{\mathcal{F}(f)} & \text{if } j < k \\ \lambda_{\mathcal{F}(f)} & \text{if } j = k \end{cases} \quad ; \quad \rho_f^{\mathcal{F}} := \begin{cases} \rho_{\mathcal{F}(f)} & \text{if } i = 0 \\ \mathbb{1}_{\mathcal{F}(f)} & \text{if } i > 0 \end{cases}$$

10. The associator

$$\alpha_{(l,\mathcal{F}),(m,\mathcal{G}),(n,\mathcal{H})} : ((l, \mathcal{F}) \cdot (m, \mathcal{G})) \cdot (n, \mathcal{H}) \to (l, \mathcal{F}) \cdot ((m, \mathcal{G}) \cdot (n, \mathcal{H}))$$

is defined as
$$\alpha_{(l,\mathcal{F}),(m,\mathcal{G}),(n,\mathcal{H})} := (\alpha^{\mathcal{F},\mathcal{G},\mathcal{H}}, \mathbb{1}_{[n+m+l]})$$

where $\alpha^{\mathcal{F},\mathcal{G},\mathcal{H}}$ is the ICON with, for any $f : i \to j : [n + m + l]$,

$$\alpha_f^{\mathcal{F},\mathcal{G},\mathcal{H}} := \alpha_{\mathcal{F}(f_1),\mathcal{G}(f_2),\mathcal{H}(f_3)} \text{ whenever } i \leq n \leq n + m \leq j$$

(after factoring f appropriately), and the appropriate identity 2-cell in all other cases.

Proof. It is clear that, for any $a, b : \mathcal{B}$, the 1-cells and 2-cells defined above make up a category: checking that (vertical) composition is associative and unital is a simple exercise.

It is equally straightforward (though slightly less immediate) to see that horizontal composition of 1-cells satisfies the usual triangle and pentagon identities. We are only left to check that horizontal composition is functorial, that is, interchange: given the following

$$(m_1, \mathcal{F}_1) \xrightarrow{(\phi_1, \alpha_1)} (m_2, \mathcal{F}_2) \xrightarrow{(\phi_2, \alpha_2)} (m_3, \mathcal{F}_3) : b \to c$$

$$(n_1, \mathcal{G}_1) \xrightarrow{(\psi_1, \beta_1)} (n_2, \mathcal{G}_2) \xrightarrow{(\psi_2, \beta_2)} (n_3, \mathcal{G}_3) : a \to b$$

we need to check that

$$((\phi_2, \alpha_2) \circ (\phi_1, \alpha_1)) \cdot ((\psi_2, \beta_2) \circ (\psi_1, \beta_1)) = ((\phi_2, \alpha_2) \cdot (\psi_2, \beta_2)) \circ ((\phi_1, \alpha_1) \cdot (\psi_1, \beta_1))$$

It is easy to check that the left-hand side equals

$$((\phi_2 \circ (\phi_1 \alpha_2)) \star (\psi_2 \circ (\psi_1 \beta_2)), (\alpha_1 \circ \alpha_2) \vee (\beta_1 \circ \beta_2))$$

and the right-hand side equals

$$((\phi_2 \star \psi_2) \circ ((\phi_1 \star \psi_1)(\alpha_2 \vee \beta_2)), (\alpha_1 \vee \beta_1) \circ (\alpha_2 \vee \beta_2))$$

One can explicitly check that $(\alpha_1 \circ \alpha_2) \vee (\beta_1 \circ \beta_2) = (\alpha_1 \vee \beta_1) \circ (\alpha_2 \vee \beta_2)$, so the second components agree. For the first component, notice how

1. $(\phi_1 \star \psi_1)(\alpha_2 \vee \beta_2) = (\phi_1 \alpha_2) \star (\psi_1 \beta_2)$

2. $(\phi_2 \circ \theta_1) \star (\psi_2 \circ \theta_2) = (\phi_2 \star \psi_2) \circ (\theta_1 \star \theta_2)$

These two facts imply that the first components agree, concluding the proof. $\square$

We can now prove the main result in this section:

Theorem 2.3. *The lax functor classifier classifies lax functors*
Given 2-categories $\mathcal{B}, \mathcal{C}$, there is a biequivalence

$$\mathbf{Lax}[\mathcal{B}, \mathcal{C}] \simeq \mathbf{Ps_{Lax}}[\widehat{\mathcal{B}}, \mathcal{C}]$$

Proof. We start by defining 2-functors

$$(\widehat{-}) : \mathbf{Lax}[\mathcal{B}, \mathcal{C}] \to \mathbf{Ps_{Lax}}[\widehat{\mathcal{B}}, \mathcal{C}]$$

$$(\widetilde{-}) : \mathbf{Ps_{Lax}}[\widehat{\mathcal{B}}, \mathcal{C}] \to \mathbf{Lax}[\mathcal{B}, \mathcal{C}]$$

then show they are each other's inverses.
Start with $\widehat{(-)}$: we define

1. On objects (that is, lax functors $\mathcal{F} : \mathcal{B} \to \mathcal{C}$) define $\widehat{\mathcal{F}}$ as

   - On objects $a : \widehat{\mathcal{B}}$, define $\widehat{\mathcal{F}}(a) := \mathcal{F}(a)$,

- On morphisms $f = (k_f, \mathcal{H}_f) : a \to b : \widehat{\mathcal{B}}$, define $\widehat{\mathcal{F}}(f) := \mathcal{H}_{\widehat{\mathcal{F}}(f)}(0 \leq k_f)$ where

$$\mathcal{H}_{\widehat{\mathcal{F}}(f)}(i) := \mathcal{F}(\mathcal{H}(i)) \ \ \forall i : [k_f]$$

$$\mathcal{H}_{\widehat{\mathcal{F}}(f)}(i \leq j) := \begin{cases} \mathbb{1} & \text{if } j = i \\ \mathcal{F}(\mathcal{H}(i \leq i+1)) & \text{if } j = i+1 \\ \mathcal{H}_{\widehat{\mathcal{F}}(f)}(i+1 \leq j) \cdot \mathcal{H}_{\widehat{\mathcal{F}}(f)}(i \leq i+1) & \text{if } j > i+1 \end{cases}$$

- On 2-cells $\Gamma = (\phi_\Gamma, \alpha_\Gamma) : f \to g : a \to b : \widehat{\mathcal{B}}$ we further distinguish two cases:

  - If ϕ_Γ is an identity, we decompose[9] α. We define $\widehat{\mathcal{F}}(\mathbb{1}, \epsilon) := \mathcal{F}_u$ and $\widehat{\mathcal{F}}(\mathbb{1}, \delta) := \mathcal{F}_c$.

  - If α_Γ is an identity, we define $\widehat{\mathcal{F}}(\Gamma) := (\phi_{\widehat{\mathcal{F}}(\Gamma)})_{0 \leq k_f}$ where

$$(\widehat{\phi_\Gamma, \mathbb{1}}) := \begin{cases} \mathbb{1} & \text{if } j = i \\ \mathcal{F}((\phi_\Gamma)_{i \leq i+1}) & \text{if } j = i+1 \\ (\phi_{\widehat{\mathcal{F}}(\Gamma)})_{i+1 \leq j} \cdot (\phi_{\widehat{\mathcal{F}}(\Gamma)})_{i \leq i+1} & \text{if } j > i+1 \end{cases}$$

- The unit constraint $\widehat{\mathcal{F}}_u$ is the identity,

- The composition constraint $\widehat{\mathcal{F}}_c$ is the appropriate composition of associators.

Showing that $\widehat{F}$ is a pseudofunctor is a simple exercise.

2. On morphisms (that is, lax 2-transformations $\Phi : \mathcal{F} \to \mathcal{G}$) define $\widehat{\Phi}$ as

- on objects $\widehat{\Phi}_a := \Phi_a$
- on morphisms $f = (k_f, \mathcal{H}_f) : a \to b$,

$$\widehat{\Phi}_f := (\Phi_{f_{k_f}} \cdot \mathcal{H}_{\widehat{\mathcal{F}}(f)}(0 \leq k_f - 1)) \cdot ... \cdot (\mathcal{H}_{\widehat{\mathcal{F}}(f)}(1 \leq k_f) \cdot \Phi_{f_1})$$

(where the horizontal compositions are to be associated to the left)

As for $\widehat{\mathcal{F}}$, showing that $\widehat{\Phi}$ is a lax 2-transformation is a simple exercise.

3. On 2-cells (that is, modifications : $\Gamma : \Phi \to \Psi : \mathcal{F} \to \mathcal{G}$) we define $\widehat{\Gamma}_a := \Gamma_a$. Showing it is a modification is, once more, a simple exercise.

Showing that $\widehat{(-)} : \mathbf{Lax}[\mathcal{B}, \mathcal{C}] \to \mathbf{Ps_{Lax}}[\widehat{\mathcal{B}}, \mathcal{C}]$ is also a straightforward calculation.

Similarly, we define $\widetilde{(-)}$:

---

[9]Recall how $I_+ \simeq \Delta_+$, hence its morphisms are generated under composition and ordinal sum by the comonoid $([1], \epsilon, \delta)$

1. On objects (that is, pseudofunctors $\mathcal{F} : \widehat{\mathcal{B}} \to \mathcal{C}$), we define $\widetilde{\mathcal{F}} : \mathcal{B} \to \mathcal{C}$ by

    - On objects $a : \mathcal{C}$, as $\widetilde{\mathcal{F}}(a) = \mathcal{F}(a)$,
    - On morphisms $f : a \to b$, as $\widetilde{\mathcal{F}}(f) := \mathcal{F}(!_f)$, where $!_f : [1] \to \mathcal{C}$ is the only semistrict pseudofunctor such that $(!_f)(0 \leq 1) = f$
    - On 2-cells $\phi : f \to g : a \to b$, as $\widetilde{\mathcal{F}}(\phi) := \mathcal{F}(!_\phi, \mathbb{1})$ where $!_\phi : !_f \to !_g$ is the only ICON with $(!_\phi)_{0 \leq 1} = \phi$.
    - The unit constraint $\widetilde{\mathcal{F}}_u$ is given by $\mathcal{F}(\mathbb{1}, \epsilon)$
    - The composition constraint $\widetilde{\mathcal{F}}$ is given by $\mathcal{F}(\mathbb{1}, \delta)$.

    Once again, showing this defines a lax functor is a simple exercise.

2. On morphisms (that is, lax 2-transformations $\Phi : \mathcal{F} \to \mathcal{G}$), we define $\widetilde{\Phi}$ by

    - On objects, $\widetilde{\Phi}_a := \Phi_a$,
    - On morphisms, $\widetilde{\Phi}_f := \Phi_{!_f}$

    Which can be easily seen as a lax 2-transformation.

3. On 2-cells (that is, modifications $\Gamma : \Phi \to \Psi : \mathcal{F} \to \mathcal{G}$), we just stipulate $\tilde{\Gamma}_a := \Gamma_a$. This clearly defines a modification.

What's left to check here is that the functors $(\widetilde{})$ and $(\widehat{})$ define a biequivalence (in fact, this turns out to be an equivalence of 2-categories). This is easily verified to be the case. $\qquad\square$

This construction is closely related to (and heavily inspired by) the notion of a *lax morphism classifier*, introduced by Blackwell, Kelly and Power [BKP89]. The following lemma is immediate:

Lemma 2.4. *Lax functor classifier as lax morphism classifier*

Given a monoidal category $\mathbb{C} := (\mathcal{C}, \otimes, I, \dots)$, the delooping of the lax morphism classifier is biequivalent to the lax functor classifier of the delooping:

$$\mathbb{B}\widehat{\mathbb{C}} \simeq \widehat{\mathbb{B}\mathbb{C}}$$

As an immediate corollary, we get a description of the lax functor classifier of the terminal 2-category:

$$\widehat{\ast} = \widehat{\mathbb{B}\ast} \simeq \mathbb{B}\widehat{\ast} \simeq \mathbb{B}\Delta_+$$

From this it's easy to see that

$$\mathbf{Mnd}(\mathcal{B}) \simeq \mathbf{Lax}[\mathbb{1}, \mathcal{B}] \simeq \mathbf{St_{Lax}}[\mathbb{B}\Delta_\alpha, \mathcal{B}]$$

$$\mathbf{Dist}(\mathcal{B}) := \mathbf{Mnd}(\mathbf{Mnd}(\mathcal{B})) \simeq \mathbf{St_{Lax}}[\mathbb{B}\Delta_\alpha, \mathbf{St_{Lax}}[\mathbb{B}\Delta_\alpha, \mathcal{B}]]$$

It would be convenient to "curry" this presentation of distributive laws, so that we could characterize $\mathbf{Mnd}^n(\mathcal{B})$ as a 2-functor category with codomain $\mathcal{B}$ (rather than an *iterated* 2-functor category). We can do something very close by employing (the lax version of) the Gray tensor product.

26

3   The Gray tensor product

Consider the *category* **2Cat** of strict 2-categories and strict 2-functors. As it turns out (cft. [Gra74] theorems I.4.9 and I.4.14), this category can be endowed with the structure of a monoidal closed category in such a way that $\mathbf{Ps}[\mathcal{C}, \mathcal{D}]$ is its internal hom: its left adjoint is then called the "Gray tensor product", usually written as just $\otimes$. There's also variants[10] that make **Lax**, or **CoLax**, the internal hom; we'll refer to them as $\otimes_l$ and $\otimes_c$.

Spelling out the adjunction in the lax case, we get

$$\mathbf{2Cat}[\mathcal{B}, \mathbf{Lax}[\mathcal{C}, \mathcal{D}]] \simeq \mathbf{2Cat}[\mathcal{B} \otimes_l \mathcal{C}, \mathcal{D}]$$

It is important to remember that what we get is *not* an equivalence of 2-categories: **2Cat** is a category, and the above construction is known to not be amenable to be extended to a 2-functorial (or 3-functorial) one[11]. What follows is a sketch of the explicit construction for the lax Gray tensor product:

Theorem 3.1. *The lax Gray tensor product*
Given 2-categories $\mathcal{B}$, $\mathcal{C}$, we define $\mathcal{B} \otimes_l \mathcal{C}$ to be the 2-category whose

1. Objects $(b, c) : \mathcal{B} \otimes_l \mathcal{C}$ are pairs of objects $b : \mathcal{B}$, $c : \mathcal{C}$;

2. 1-cells $h : (b, c) \to (b', c') : \mathcal{B} \otimes_l \mathcal{C}$ are composable sequences of pairs of 1-cells $h = (f_1, g_2) \cdot \ldots \cdot (f_n, g_n)$ where for each $i \in \{1, \ldots, n\}$ either $f = \mathbf{Id}$ or $g = \mathbf{Id}$, subject to the relations

$$(f_1, \mathbf{Id}) \circ (f_2, \mathbf{Id}) = (f_1 \circ f_2, \mathbf{Id})$$

$$(\mathbf{Id}, g_1) \circ (\mathbf{Id}, g_2) = (\mathbf{Id}, g_1 \circ g_2)$$

3. 2-cells are generated (under horizontal and vertical composition) by the following three families:

$$(\phi, \mathbf{Id}) : (f_1, \mathbf{Id}) \to (f_2, \mathbf{Id})$$

$$(\mathbf{Id}, \psi) : (\mathbf{Id}, g_1) \to (\mathbf{Id}, g_2)$$

$$\gamma_{f,g} : (f, \mathbf{Id}) \circ (\mathbf{Id}, g) \to (\mathbf{Id}, g) \circ (f, \mathbf{Id})$$

subject to the following relations, where we write $\circ$ for vertical and $\cdot$ for horizontal composition:

(a) for any $f_3 \xrightarrow{\phi_2} f_2 \xrightarrow{\phi_1} f_1 : b_1 \to b_2 : \mathcal{B}$ and $c : \mathcal{C}$,

$$(\phi_1, \mathbf{Id}) \circ (\phi_2, \mathbf{Id}) = (\phi_1 \circ \phi_2, \mathbf{Id}) : (f_3, \mathbf{Id}) \to (f_1, \mathbf{Id})$$

---

[10]The lax variant is the one that was originally discussed in Gray's paper, thm I.4.9 in [Gra74]

[11]See this post by T. Champion for a counterexample: [htta]

(b) for any $g_3 \xrightarrow{\psi_2} g_2 \xrightarrow{\psi_1} g_1 : c_1 \to c_2 : \mathcal{C}$ and $b : \mathcal{B}$,

$$(\mathbf{Id}, \psi_1) \circ (\mathbf{Id}, \psi_2) = (\mathbf{Id}, \psi_1 \circ \psi_2) : (\mathbf{Id}, g_3) \to (\mathbf{Id}, g_1) :$$

(c) for any $f_1 \xrightarrow{\phi_1} f_2 : b_2 \to b_3 : \mathcal{B}$, $f_3 \xrightarrow{\phi_2} f_4 : b_1 \to b_2 : \mathcal{B}$ and $c : \mathcal{C}$,

$$(\phi_1, \mathbf{Id}) \cdot (\phi_2, \mathbf{Id}) = (\phi_1 \cdot \phi_2, \mathbf{Id}) : (f_1 \cdot f_3, \mathbf{Id}) \to (f_2 \cdot f_4, \mathbf{Id})$$

(d) for any $g_1 \xrightarrow{\psi_1} g_2 : c_2 \to c_3 : \mathcal{C}$, $g_3 \xrightarrow{\psi_2} g_4 : c_1 \to c_2 : \mathcal{C}$ and $b : \mathcal{B}$,

$$(\mathbf{Id}, \psi_1) \cdot (\mathbf{Id}, \psi_2) = (\mathbf{Id}, \psi_1 \cdot \psi_2) : (\mathbf{Id}, g_1 \cdot g_3) \to (\mathbf{Id}, g_2 \cdot g_4)$$

(e) for all $f : b_1 \to b_2 : \mathcal{B}$,

$$\gamma_{f,\mathbf{Id}} = \mathbf{Id} : (f, \mathbf{Id}) \to (f, \mathbf{Id})$$

(f) for all $g : c_1 \to c_2 : \mathcal{C}$,

$$\gamma_{\mathbf{Id},g} = \mathbf{Id} : (\mathbf{Id}, g) \to (\mathbf{Id}, g)$$

(g) for any $b_3 \xrightarrow{f_2} b_2 \xrightarrow{f_1} b_1 : \mathcal{B}$ and $g : c_1 \to c_2 : \mathcal{C}$,

$$\gamma_{(f_1 \cdot f_2), g} = (\gamma_{f_1, g} \cdot (f_2, \mathbf{Id})) \circ ((f_1, \mathbf{Id}) \cdot \gamma_{f_2, g}) : (f_1 \cdot f_2, \mathbf{Id}) \cdot (\mathbf{Id}, g) \to (\mathbf{Id}, g) \cdot (f_1 \cdot f_2, \mathbf{Id})$$

(h) for any $c_3 \xrightarrow{g_2} c_2 \xrightarrow{g_1} c_1 : \mathcal{C}$ and $f : b_1 \to b_2 : \mathcal{B}$,

$$\gamma_{f,(g_1 \cdot g_2)} = ((\mathbf{Id}, g_1) \cdot \gamma_{f, g_2}) \circ (\gamma_{f, g_1} \cdot (\mathbf{Id}, g_2)) : (f, \mathbf{Id}) \cdot (\mathbf{Id}, g_1 \cdot g_2) \to (\mathbf{Id}, g_1 \cdot g_2) \cdot (f, \mathbf{Id})$$

(i) for any $f_1 \xrightarrow{\phi} f_2 : b_1 \to b_2 : \mathcal{B}$ and $g_1 \xrightarrow{\psi} g_2 : c_1 \to c_2 : \mathcal{C}$,

$$\gamma_{f_2, g_2} \circ ((\phi, \mathbf{Id}) \cdot (\mathbf{Id}, \psi)) = ((\mathbf{Id}, \psi) \cdot (\phi, \mathbf{Id})) \circ \gamma_{f_1, g_1} : (f_1, \mathbf{Id}) \cdot (\mathbf{Id}, g_1) \to (\mathbf{Id}, g_2) \cdot (f_2, \mathbf{Id})$$

With this tool, we can get a (1-)categorical description of parametric monads and parametric distributive laws. If we replace the 2-category of parametric distributive laws

$$\mathbf{PDist}(\mathcal{B}, \mathcal{C}) := \mathbf{St_{St}}[\mathcal{B}, \mathbf{Dist}(\mathcal{C})]$$

with a less structured *category* of parametric distributive laws

$$PDist(\mathcal{C}, \mathcal{D}) := \mathbf{2Cat}[\mathcal{C}, \mathbf{Dist}(\mathcal{D})]$$

then we can apply the tensor-hom adjunction for the lax Gray tensor product, obtaining the following characterization:

$$PDist(\mathcal{B}, \mathcal{C}) \simeq \mathbf{2Cat}[(\mathcal{B} \otimes_l \mathbb{B}\Delta_\alpha) \otimes_l \mathbb{B}\Delta_\alpha, \mathcal{C}]$$

Something completely analogous can be done for parametric monads $PMnd$ (replacing the previously defined **PMnd**)

$$PMnd(\mathcal{B}, \mathcal{C}) \simeq \mathbf{2Cat}[\mathcal{B} \otimes_l \mathbb{B}\Delta_\alpha, \mathcal{C}]$$

and (as we will discuss later) we can do the same for iterated distributive laws. This suggests defining operations on 2-categories

$$\mathbf{mnd}(\mathcal{C}) := \mathcal{C} \otimes_l \mathbb{B}\Delta_\alpha$$

$$\mathbf{dist}(\mathcal{C}) := \mathbf{mnd}^2(\mathcal{C}) = (\mathcal{C} \otimes_l \mathbb{B}\Delta_\alpha) \otimes_l \mathbb{B}\Delta_\alpha$$

so that

$$PMnd(\mathcal{C}, \mathcal{B}) \simeq \mathbf{2Cat}[\mathbf{mnd}(\mathcal{C}), \mathcal{B}]$$

$$PDist(\mathcal{C}, \mathcal{B}) \simeq \mathbf{2Cat}[\mathbf{dist}(\mathcal{C}), \mathcal{B}]$$

Before moving on, we want to state the following explicit characterization of iterated Gray tensor products (which will be useful later)

Theorem 3.2. *Iterated lax Gray tensor products*

Given n 2-categories $\mathcal{C}_1, \ldots, \mathcal{C}_n$, we define $\bigotimes_{i=1}^{n} \mathcal{C}_i$ by $\bigotimes_{i=1}^{1} := \mathcal{C}_1$ and $\bigotimes_{i=1}^{k+1} \mathcal{C}_i := (\bigotimes_{i=1}^{k} \mathcal{C}_i) \otimes_l \mathcal{C}_{k+1}$.

1. the 0-cells in $\bigotimes_{i=1}^{n} \mathcal{C}_i$ are n-tuples of 0-cells $(a_1, \ldots, a_n)$ with $a_i : \mathcal{C}_i$;

2. the 1-cells $(a_1, \ldots, a_n) \to (b_1, \ldots, b_n)$ are generated (under composition) by $\{f\}_i : (a_1, \ldots, a_i, \ldots, a_n) \to (a_1, \ldots, a_i', \ldots, a_n)$ for $f : a_i \to a_i' : \mathcal{C}_i$, subject to the relation $\{f\}_i \cdot \{g\}_i = \{f \cdot g\}_i$ for all choices of f, g, i.

3. the 2-cells are generated, under horizontal and vertical composition, by the following two families:

   - $\{\phi\}_i : \{f\}_i \to \{g\}_i$ for $\phi : f \to g$ in $\mathcal{C}_i$
   - $\gamma_{f,g}^{i,j} : \{f\}_i \cdot \{g\}_j \to \{g\}_j \cdot \{f\}_i$ for any $i < j$

   subject to the following relations (whenever well-typed):

   (a) compatibility with vertical and horizontal composition: for each i

   $$\{\phi\}_i \circ \{\psi\}_i = \{\phi \circ \psi\}_i \ , \ \ \{\phi\}_i \cdot \{\psi\}_i = \{\phi \cdot \psi\}_i$$

   (b) compatibility with identities: for each $i < j$

   $$\gamma_{f,\mathbf{Id}}^{i,j} = \{f\}_i \ , \ \ \gamma_{\mathbf{Id},f}^{i,j} = \{f\}_j$$

   (c) compatibility with composition of 1-cells: for each $i < j$

   $$\gamma_{(f_1 \cdot f_1),g}^{i,j} = (\gamma_{f_1,g}^{i,j} \cdot \{f_2\}_i) \circ (\{f_1\}_i \cdot \gamma_{f_2,g}^{i,j})$$

   $$\gamma_{f,(g_1 \cdot g_2)}^{i,j} = (\{g_1\}_j \cdot \gamma_{f,g_2}^{i,j}) \circ (\gamma_{f,g_1}^{i,j} \cdot \{g_2\}_j)$$

(d) naturality: for each $i < j$

$$(\{\psi\}_j \cdot \{\phi\}_i) \circ \gamma_{f_1,g_1}^{i,j} = \gamma_{f_2,g_2}^{i,j} \circ (\{\phi\}_i \cdot \{\psi\}_j)$$

(e) the Yang-Baxter equation(s): for each $i < j < k$

$$(\{h\}_k \cdot \gamma_{f,g}^{i,j}) \circ (\gamma_{f,h}^{i,k} \cdot \{g\}_j) \circ (\{f\}_i \cdot \gamma_{g,h}^{j,k}) = (\gamma_{g,h}^{j,k} \cdot \{f\}_i) \circ (\{g\}_j \cdot \gamma_{f,h}^{i,k}) \circ (\gamma_{f,g}^{i,j} \cdot \{h\}_k)$$

Proof. By induction on n; the base case is trivial. Indeed, for $n = 2$ conditions (a - d) are exactly the ones spelled out in 3.1. Notice the shift in notation: we write $\{f\}_1 := (f, \mathbf{Id})$ and $\{f\}_2 := (\mathbf{Id}, f)$ for 1- and 2-cells, and $\gamma_{f,g}^{1,2} := \gamma_{f,g}$. Condition (e) doesn't apply, since it requires at least $n = 3$.

Consider now the inductive case $n + 1$ (for $n \geq 2$):

$$\bigotimes_{i=1}^{n+1} \mathcal{C}_i = (\bigotimes_{i=1}^{n} \mathcal{C}_i) \otimes_l \mathcal{C}_{n+1}$$

By inductive hypothesis, the characterization in question applies to $\bigotimes_{i=1}^{n} \mathcal{C}_i$. For $i < j \leq n$ we write $\{f\}_i := (\{f\}_i, \mathbf{id})$ and $\{f\}_{n+1} := (\mathbf{Id}, f)$ for 1- and 2-cells, and $\gamma_{f,g}^{i,j} := (\gamma_{f,g}^{i,j}, \mathbf{Id})$. Moreover, we write $\gamma_{f,g}^{i,n+1} := \gamma_{\{f\}_i,g}$.

The characterization for 1-cells holds trivially; we will only discuss the relations imposed on 2-cells. Clearly, if all indices are strictly less than $n + 1$ the relations hold by inductive hypothesis;

(a) holds by definition for $i = n + 1$;

(b) for $i \leq n$, we want to prove

$$\gamma_{f,\mathbf{Id}}^{i,n+1} := \gamma_{\{f\}_1,\mathbf{Id}} = (\{f\}_i, \mathbf{Id}) =: \{f\}_i$$

$$\gamma_{\mathbf{Id},f}^{i,n+1} := \gamma_{\mathbf{Id},f} = (\mathbf{Id}, f) =: \{f\}_{n+1}$$

both of which hold by the definition of Gray tensor product 3.1, conditions (e) and (f) respectively.

(c) for $i \leq n$, we want to prove

$$\gamma_{(f_1 \cdot f_2),g}^{i,n+1} = (\gamma_{f_1,g}^{i,n+1} \cdot \{f_2\}_i) \circ (\{f_1\}_i \cdot \gamma_{f_2,g}^{i,n+1})$$

$$\gamma_{f,(g_1 \cdot g_2)}^{i,n+1} = (\{g_1\}_{n+1} \cdot \gamma_{f,g_2}^{i,n+1}) \circ (\gamma_{f,g_1}^{i,n+1} \cdot \{g_2\}_{n+1})$$

Start with the first equation: the left hand side is defined as

$$\gamma_{(f_1 \cdot f_2),g}^{i,n+1} := \gamma_{\{f_1 \cdot f_2\}_i,g} = \gamma_{(\{f_1\}_i \cdot \{f_2\}_i),g}$$

while the right hand side is

$$(\gamma_{f_1,g}^{i,n+1} \cdot \{f_2\}_i) \circ (\{f_1\}_i \cdot \gamma_{f_2,g}^{i,n+1}) := (\gamma_{\{f_1\}_i,g} \cdot (\{f_2\}_i, \mathbf{Id})) \circ ((\{f_1\}_i, \mathbf{Id}) \cdot \gamma_{\{f_2\}_i,g})$$

So our claim reduces to

$$\gamma_{(\{f_1\}_i \cdot \{f_2\}_i),g} = (\gamma_{\{f_1\}_i,g} \cdot (\{f_2\}_i, \mathbf{Id})) \circ ((\{f_1\}_i, \mathbf{Id}) \cdot \gamma_{\{f_2\}_i,g})$$

which is exactly point (g) in 3.1. The proof for the other equation is completely analogous.

30

(d) for $i \leq n$ we want to prove

$$(\{\psi\}_{n+1} \cdot \{\phi\}_i) \circ \gamma_{f_1,g_1}^{i,n+1} = \gamma_{f_2,g_2}^{i,n+1} \circ (\{\phi\}_i \cdot \{\psi\}_{n+1})$$

Again, we unroll our notation obtaining

$$((\mathbf{Id}, \psi) \cdot (\{\phi\}_i, \mathbf{Id})) \circ \gamma_{\{f_1\}_i,g_1} = \gamma_{\{f_2\}_i,g_2} \circ ((\{\phi\}_i, \mathbf{Id}) \cdot (\mathbf{Id}, \psi))$$

which is exactly condition (i) in 3.1.

(e) for $i < j \leq n$, we want to prove

$$(\{h\}_{n+1} \cdot \gamma_{f,g}^{i,j}) \circ (\gamma_{f,h}^{i,n+1} \cdot \{g\}_j) \circ (\{f\}_i \cdot \gamma_{g,h}^{j,n+1}) = (\gamma_{g,h}^{j,n+1} \cdot \{f\}_i) \circ (\{g\}_j \cdot \gamma_{f,h}^{i,n+1}) \circ (\gamma_{f,g}^{i,j} \cdot \{h\}_{n+1})$$

Unrolling our notation one last time, we get

$$((\mathbf{Id}, h) \cdot (\gamma_{f,g}^{i,j}, \mathbf{Id})) \circ (\gamma_{\{f\}_i,h} \cdot (\{g\}_i, \mathbf{Id})) \circ ((\{f\}_i, \mathbf{Id}) \cdot \gamma_{\{g\}_j,h})$$

for the left hand side, and

$$(\gamma_{\{g\}_j,h} \cdot (\{f\}_i, \mathbf{Id})) \circ ((\{g\}_j, \mathbf{Id}) \cdot \gamma_{\{f\}_i,h}) \circ ((\gamma_{f,g}^{i,j}, \mathbf{Id}) \cdot (\mathbf{Id}, h))$$

for the right hand side. As is easy to see, we can apply point (g) in 3.1 on both sides and get

$$((\mathbf{Id}, h) \cdot (\gamma_{f,g}^{i,j}, \mathbf{Id})) \circ \gamma_{(\{f\}_i \cdot \{g\}_j),h}$$

for the left hand side, and

$$\gamma_{(\{g\}_j \cdot \{f\}_i),h} \circ ((\gamma_{f,g}^{i,j}, \mathbf{Id}) \cdot (\mathbf{Id}, h))$$

for the right hand side. We can now invoke point (i) in 3.1 to conclude that they must equal each other.

$\square$

As a corollary of this, whenever we perform two or more applications of the Gray tensor product, we are going to introduce some Yang-Baxter equations.

4   Parametric distributive laws, explicitly

In this section, we give a concrete description of the data that goes into a parametric distributive law. In order to do so, we will specifically look at $\mathcal{C}$-parametric distributive laws in the case where the 2-category of parameters is the terminal 2-category $\mathcal{C} = \mathbb{1}$.

As is easy to compute $\mathbb{1} \otimes_l \mathcal{C} = \mathcal{C}$, hence $\mathbb{1} \otimes_l \mathbb{B}\Delta_\alpha = \mathbb{B}\Delta_\alpha$. This means that the category of $\mathbb{1}$-parametric distributive laws[12]

$$PDist(\mathbb{1}, \mathcal{B}) \simeq \mathbf{2Cat}[\mathbb{B}\Delta_\alpha \otimes_l \mathbb{B}\Delta_\alpha, \mathcal{B}]$$

---

[12]Notice how this is different from the 2-category of distributive laws $\mathbf{Dist}(\mathcal{B}) \simeq \mathbf{St_{Lax}}[\mathbb{B}\Delta_\alpha, \mathbf{St_{lax}}[\mathbb{B}\Delta_\alpha, \mathcal{B}]]$: while 0-cells are the same, the 1-cells differ already! Moreover, the former is "just" a category while the latter is a fully fledged 2-category.

It is worth spelling out the data that goes into such a parametric distributive law more explicitly, stressing what pieces of data correspond to the two monads and what to the actual distributive law.

Remark 4.1. *Distributive laws as strict functors*

A $\mathbb{1}$-parametric distributive law $\mathcal{D} : PDist(\mathbb{1}, \mathcal{B})$ is given by the following data

1. A 0-cell $a : \mathcal{B}$,

2. 1-cells $s, t : a \to a : \mathcal{B}$

3. 2-cells $\eta^s : \mathbf{Id}_a \to s$, $\mu^s : s \cdot s \to s$, $\eta^t : \mathbf{Id}_a \to t$, $\mu^t : t \cdot t \to t$ and $\sigma : s \cdot t \to t \cdot s$

subject to the appropriate axioms: (s, η^s, μ^s) and (t, η^t, μ^t) are required to be monads, and σ is required to be a distributive law in the sense of J. Beck ([Bec69]).

Proof. Recall how Δ_α is the free monoidal category with a monoid object[13]. This implies that

1. Its objects are generated (under its tensor product, sometimes called ordinal sum $\oplus$) by the singleton ordinal $[0]$: the empty ordinal $[-1]$ would then be its 0-fold ordinal sum; and in general we have that $[n] = [0]^{\oplus n+1}$.

2. Its morphisms are generated (under both composition and ordinal sum) by the unit and multiplication of the monoid structure on $[0]$: $\eta : [-1] \to [0]$ and $\mu : [1] \to [0]$.

Hence its delooping admits an almost identical characterization for its 0-, 1- and 2-cells. From the construction of the Gray tensor product, we can then notice that in order to define a strict 2-functor $\mathbb{B}\Delta_\alpha \otimes_l \mathbb{B}\Delta_\alpha \to \mathcal{B}$ we only need to specify where the two copies of the data we just listed end up, together with the "swap" 2-cell $\gamma_{[0],[0]}$. Explicitly, we only need to care about

1. The 0-cell $(*, *)$, whose image $\mathcal{D}(*, *)$ we can call a;

2. the 1-cells $([0], \mathbf{Id})$ and $(\mathbf{Id}, [0])$, whose image we can call s and t respectively;

3. the 2-cells $(\eta, \mathbf{Id})$, $(\mu, \mathbf{Id})$, $(\mathbf{Id}, \eta)$, $(\mathbf{Id}, \mu)$ and $\gamma_{[0],[0]}$. We can call their images $\eta^s, \mu^s, \eta^t, \mu^t, \sigma$ respectively.

$\square$

---

[13]more precisely: it is the initial object in the 2-category whose 0-cells are monoidal categories with a specified monoid object, whose 1-cells are "pointed" monoidal functors, and whose 2-cells are monoidal natural transformations

This should not be surprising: we set up our notion of parametric distributive laws exactly for this result to follow. What might turn out to be (at least slightly) surprising is that we can now produce a similarly explicit description of parametric iterated distributive laws too, which we define just as

$$PDist^n(\mathcal{C}, \mathcal{B}) := \mathbf{2Cat}[\mathcal{C}, \mathbf{Mnd}^n(\mathcal{B})] \simeq \mathbf{2Cat}[\mathbf{mnd}^n(\mathcal{C}), \mathcal{B}]$$

Notice how, as an artifact of our counting convention, a 1-fold parametric distributive law is a parametric monad; and a 2-fold such is a parametric distributive law. The definition produces something new for $n > 2$.

Remark 4.2. *3-fold distributive laws as strict functors*

A $\mathbb{1}$-parametric 3-fold distributive law $\mathcal{D} : PDist^3(\mathbb{1}, \mathcal{B})$ is given by the data of

1. A 0-cell $a : \mathcal{B}$,

2. 1-cells $t_1, t_2, t_3 : a \to a$,

3. 2-cells $\eta^i : \mathbf{Id}_a \to t_i$ and $\mu^i : t_i \cdot t_i \to t_i$ for $i \in \{1, 2, 3\}$, as well as $\sigma_{i,j}$ as before for $i < j \in \{1, 2, 3\}$

subject to the appropriate axioms: (t_i, η^i, μ^i) are required to be monads, $\gamma_{i,j}$ to be distributive laws, and (as spelled out by E. Cheng in [Che11]) such distributive laws are required to satisfy the Yang-Baxter equation.

Proof. This is a corollary of the characterization of $\mathbb{1}$-parametric distributive laws in 4.1 and our remarks about the universal property of Δ_α therein, together with 3.2. $\qquad\square$

The previous result immediately generalizes to $\mathbb{1}$-parametric n-fold distributive laws, without any significant change to the proof:

Remark 4.3. *n-fold distributive laws as strict functors*

A $\mathbb{1}$-parametric n-fold distributive law $\mathcal{D} : PDist^n(\mathbb{1}, \mathcal{B})$ is given by the following data:

1. A 0-cell $a : \mathcal{B}$,

2. 1-cells $t_1, \dots, t_n : a \to a$,

3. 2-cells η^i and μ^i for $1 \leq i \leq n$ making t_i a monad on a

4. 2-cells $\sigma_{i,j}$ for $1 \leq i < j \leq n$ distributive laws, satisfying the Yang-Baxter equation when appropriate, i.e. for all choices of $1 \leq i < j < k \leq n$.

5 Example: the Writer monad

In this section, we will show how the "Writer" monad admits a parametric distributive law. We will achieve this by explicitly constructing a lift of an arbitrary (appropriately structured) monad to its Eilenberg-Moore category.

First, some definitions:

Definition 5.1. *The Writer monad* Let $\mathcal{C}$ be a category with a terminal object $*$ and all binary products $A \times B$. Given a monoid object $\mathbf{M} = (M, u, m) : \mathbf{Mon}(\mathcal{C})$, we can define a functor

$$\mathbf{Writer_M} : \mathcal{C} \to \mathcal{C}$$

$$\mathbf{Writer_M}(A \xrightarrow{f} B) := (M \times A) \xrightarrow{M \times f} (M \times B)$$

As well as natural transformations

$$\eta : \mathbf{Id} \to \mathbf{Writer_M}$$

$$\mu : \mathbf{Writer_M^2} \to \mathbf{Writer_M}$$

by setting

$$\eta_A = A \simeq * \times A \xrightarrow{u \times A} M \times A$$

$$\mu_A : M \times (M \times A) \simeq (M \times M) \times A \xrightarrow{m \times A} M \times A$$

Naturality is immediate to verify. This choice makes $(\mathbf{Writer_M}, \eta, \mu)$ into a monad.

Proof. We need to show that η and μ satisfy unitality and associativity. For unitality, consider

The triangles on the top left and right commute by the coherence theorem, the two squares on the sides by interchange. The remaining cells in the center are exactly the unitality assumption on the monoid $\mathbf{M} = (M, u, m)$ tensored by A. A similar argument works for associativity: consider the following, where juxtaposition is to be read as the cartesian product $\times$:

$$M(M(MA)) \xrightarrow{\quad M\alpha^{-1} \quad} M((MM)A) \xrightarrow{M(mA)} M(MA)$$

The two square cells commute by interchange, the pentagon on the top left is the pentagon axiom for monoidal categories. The remaining pentagon is the associativity condition on $\mathbf{M} = (M, u, m)$ tensored by A. $\qquad\square$

There's a more compact way of describing the above definition: it is the monad sending an object $A : \mathcal{C}$ to the (carrier of the) corresponding *free* $\mathbf{M}$-*module*

$$\mathbf{M} \times A := (M \times A, m \times A)$$

More explicitly, we have the category $\mathbf{Mod_M}(\mathcal{C})$ of $\mathbf{M}$-modules[14] whose objects are pairs (A, ϕ) with $\phi : M \times A \to A$ satisfying appropriate unitality and associativity conditions; morphisms $f : (A, \phi) \to (B, \psi)$ are the morphisms in the base category $f : A \to B : \mathcal{C}$ that make the obvious square commute.

The category $\mathbf{Mod_M}(\mathcal{C})$ always comes equipped with a forgetful functor $\mathcal{U} : \mathbf{Mod_M}(\mathcal{C}) \to \mathcal{C}$, which turns out to have a left adjoint; the associated monad is exactly $\mathbf{Writer_M}$.

In the following, we will assume $\mathcal{C}$ to be a Cartesian Closed Category, implying that it is enriched over itself (though somewhat milder assumptions would suffice). We will also assume fixed a monoid object $\mathbf{M} = (M, u, m) : \mathbf{Mon}(\mathcal{C})$.

Lemma 5.2. *Strong and enriched functors*

Given an endofunctor $\mathcal{F} : \mathcal{C} \to \mathcal{C}$, the following are equivalent

- A tensorial strength $\tau_{A,B} : A \times \mathcal{F}(B) \to \mathcal{F}(A \times B)$ interacting appropriately with the unitors and associators of $\mathcal{C}$;

- A $\mathcal{C}$-enrichment for $\mathcal{F}$, meaning a natural family of morphisms interacting appropriately with identities and compositions

$$\mathcal{F}_{A,B} : [A, B]_{\mathcal{C}} \to [\mathcal{F}(A), \mathcal{F}(B)]_{\mathcal{C}}$$

Proof. The result is standard, so we only sketch the construction back-and-forth (for a full proof, see sections 3.2 and 3.3 of [Rat13]).

---

[14]We don't need to choose any handedness for these modules, since the cartesian monoidal structure is always symmetric; for the sake of concreteness however, whenever such a choice is to be made, we'll work with left modules.

First, assume we have a tensorial strength $\tau_{A,B} : A \times \mathcal{F}(B) \to \mathcal{F}(A \times B)$. We can define

$$[A, B]_\mathcal{C} \times \mathcal{F}(A) \xrightarrow{\tau} \mathcal{F}([A, B]_\mathcal{C} \times A) \xrightarrow{Ev} \mathcal{F}(B)$$

where Ev is the evaluation map, the counit of the tensor-hom adjunction. The morphism we just defined is the mate (under the tensor-hom adjunction in $\mathcal{C}$) of the desired enrichment

$$\mathcal{F}_{A,B} : [A, B]_\mathcal{C} \to [\mathcal{F}(A), \mathcal{F}(B)]_\mathcal{C}$$

Second, assume we have an enrichment

$$\mathcal{F}_{A,B} : [A, B]_\mathcal{C} \to [\mathcal{F}(A), \mathcal{F}(B)]_\mathcal{C}$$

We can define

$$A \times \mathcal{F}(B) \to [B, (A \times B)]_\mathcal{C} \times \mathcal{F}(B) \to [\mathcal{F}(B), \mathcal{F}(A \times B)]_\mathcal{C} \times \mathcal{F}(B) \to \mathcal{F}(A \times B)$$

$\square$

where the first map is the unit of the tensor-hom adjunction (tensored by $\mathcal{F}(B)$), the second is the enrichment, followed by the evaluation map.

An analogous statement holds for monads: the notions of enriched monad and strong monad coincide (see proposition 5.8 in [MU22]).

We can now state our main result:

Theorem 5.3. *Enriched monads lift to modules*

Given an enriched monad $(\mathcal{T}, \eta, \mu)$ on $\mathcal{C}$, there's a corresponding monad $(\hat{\mathcal{T}}, \hat{\eta}, \hat{\mu})$ on $\mathbf{Mod_M}(\mathcal{C})$ such that

$$\mathcal{U} \circ \hat{\mathcal{T}} = \mathcal{T} \circ \mathcal{U}$$

Proof. We start by defining the monad $\hat{\mathcal{T}} : \mathbf{Mod_M}(\mathcal{C}) \to \mathbf{Mod_M}(\mathcal{C})$

On objects, we can just define $\hat{\mathcal{T}}(A, \phi) := (\mathcal{T}(A), \hat{\phi})$ with

$$M \times \mathcal{T}(A) \xrightarrow{\sigma} \mathcal{T}(M \times A) \xrightarrow{\mathcal{T}(\phi)} \mathcal{T}(A)$$

where σ is the tensorial strength of $\mathcal{T}$.

To see that $\hat{\mathcal{T}}(A, \phi) : \mathbf{Mod_M}(\mathcal{C})$, consider the following:

$$
\begin{array}{ccccccc}
* \times \mathcal{T}(A) & \xrightarrow{u \times \mathcal{T}(\phi)} & M \times \mathcal{T}(A) & \xleftarrow{M \times \mathcal{T}(\phi)} & M \times \mathcal{T}(M \times A) & \xleftarrow{M \times \sigma} & M \times (M \times \mathcal{T}(A)) \\
\downarrow{\sigma} & & \downarrow{\sigma} & & \sigma\downarrow & & \alpha^{-1}\downarrow \\
\mathcal{T}(* \times A) & \xrightarrow{\mathcal{T}(u \times A)} & \mathcal{T}(M \times A) & \xleftarrow{\mathcal{T}(M \times \phi)} & \mathcal{T}(M \times (M \times A)) & & (M \times M) \times \mathcal{T}(A) \\
\downarrow{\mathcal{T}(\sim)} & & \downarrow{\mathcal{T}(\phi)} & & & \mathcal{T}(\alpha^{-1})\searrow & \sigma\downarrow \\
\mathcal{T}(A) & =\!=\!= & \mathcal{T}(A) & \xleftarrow{\mathcal{T}(\phi)} & \mathcal{T}(M \times A) & \xleftarrow{\mathcal{T}(m \times A)} & \mathcal{T}((M \times M) \times A)
\end{array}
$$

The two squares on the top (left and center) commute by naturality of σ, the pentagon on the right involving the associator twice by the fact that $(\mathcal{T}, \sigma)$ is a strong functor. The two remaining cells commute by the assumption that (A, ϕ) is an **M**-module.

Let now $f : (A, \phi) \to (B, \psi) : \mathbf{Mod_M}(\mathcal{C})$. We define $\hat{\mathcal{T}}(f) := \mathcal{T}(f)$; to see that it is a morphism of **M**-modules, consider

$$
\begin{array}{ccc}
M \times \mathcal{T}(A) & \xrightarrow{M \times \mathcal{T}(f)} & M \times \mathcal{T}(B) \\
\downarrow{\scriptstyle \sigma} & & \downarrow{\scriptstyle \sigma} \\
\mathcal{T}(M \times A) & \xrightarrow{\mathcal{T}(M \times f)} & \mathcal{T}(M \times B) \\
\downarrow{\scriptstyle \mathcal{T}(\phi)} & & \downarrow{\scriptstyle \mathcal{T}(\psi)} \\
\mathcal{T}(A) & \xrightarrow{\mathcal{T}(f)} & \mathcal{T}(B)
\end{array}
$$

The top square commutes by naturality of σ, the bottom one commute by the assumption that f is a morphism of **M**-modules.

All that remains is defining $\hat{\eta}$ and $\hat{\mu}$, in a way such that $(\hat{\mathcal{T}}, \hat{\eta}, \hat{\mu})$ is a monad on $\mathbf{Mod_M}(\mathcal{C})$; we define $\hat{\eta}_{(A,\phi)} := \eta_A$ and $\hat{\mu}_{(A,\phi)} := \mu_A$.

We need to show that the components $\hat{\eta}_{(A,\phi)}$, $\hat{\mu}_{(A,\phi)}$ are morphisms of **M**-modules. For $\hat{\eta}$, consider:

$$
\begin{array}{ccc}
M \times A & \xrightarrow{M \times \eta_A} & M \times \mathcal{T}(A) \\
 & \searrow{\scriptstyle \eta_{M \times A}} & \downarrow{\scriptstyle \sigma} \\
\downarrow{\scriptstyle \phi} & & \mathcal{T}(M \times A) \\
 & & \downarrow{\scriptstyle \mathcal{T}(\phi)} \\
A & \xrightarrow{\eta_A} & \mathcal{T}(A)
\end{array}
$$

The triangle commutes by the assumption that σ makes $(\mathcal{T}, \eta, \mu)$ into a strong monad, the square by naturality of η. As for $\hat{\mu}$, consider

$$
\begin{array}{ccc}
M \times \mathcal{T}(\mathcal{T}(A)) & \xrightarrow{M \times \mu_A} & M \times \mathcal{T}(A) \\
\downarrow{\scriptstyle \sigma} & & \\
\mathcal{T}(M \times \mathcal{T}(A)) & & \downarrow{\scriptstyle \sigma} \\
\downarrow{\scriptstyle \mathcal{T}(\sigma)} & & \\
\mathcal{T}(\mathcal{T}(M \times A)) & \xrightarrow{\mu_{M \times A}} & \mathcal{T}(M \times A) \\
\downarrow{\scriptstyle \mathcal{T}(\mathcal{T}(\phi))} & & \downarrow{\scriptstyle \mathcal{T}(\phi)} \\
\mathcal{T}(\mathcal{T}(A)) & \xrightarrow{\mu_A} & \mathcal{T}(A)
\end{array}
$$

The pentagon commutes by σ being the tensorial strength of the monad $(\mathcal{T}, \eta, \mu)$, the square by naturality of μ.

It only remains to prove that $(\hat{\mathcal{T}}, \hat{\eta}, \hat{\mu})$ is a monad; the proof follows immediately by the assumption that $(\mathcal{T}, \eta, \mu)$ is. $\qquad\square$

This allows us to construct a parametric distributive law with parameters given by the 2-category of enriched monads over $\mathcal{C}$:

Theorem 5.4. *Parametric distributive law for the Writer monad*

Define $\mathcal{B} := \mathbb{B}\mathbf{Cat}_{\mathcal{C}}[\mathcal{C}, \mathcal{C}]$ to be the delooping of the (monoidal) category of $\mathcal{C}$-enriched endofunctors of $\mathcal{C}$, and $\mathcal{D} := \mathbf{Mnd}(\mathcal{C})$.

There is an object in $PDist(\mathcal{D}, \mathcal{C})$, that is a strict 2-functor

$$(\mathbf{Mnd}(\mathbb{B}\mathbf{Cat}_{\mathcal{C}}[\mathcal{C}, \mathcal{C}]) \otimes_l \mathbb{B}\Delta_\alpha) \otimes_l \Delta_\alpha \to \mathbb{B}\mathbf{Cat}_{\mathcal{C}}[\mathcal{C}, \mathcal{C}]$$

such that the underlying monads, encoded as strict 2-functors

$$\mathbf{Mnd}(\mathbb{B}\mathbf{Cat}_{\mathcal{C}}[\mathcal{C}, \mathcal{C}]) \to \mathbf{Mnd}(\mathbb{B}\mathbf{Cat}_{\mathcal{C}}[\mathcal{C}, \mathcal{C}])$$

are, respectively, constantly the writer monad and the identity.

Proof. Such a 2-functor would correspond to a 2-functor

$$\mathbf{DWriter_M} : \mathbf{Mnd}(\mathbb{B}\mathbf{Cat}_{\mathcal{C}}[\mathcal{C}, \mathcal{C}]) \to \mathbf{Dist}(\mathbb{B}\mathbf{Cat}_{\mathcal{C}}[\mathcal{C}, \mathcal{C}])$$

which we will describe in detail.

On objects $(*, T, \eta^T, \mu^T)$, the functor acts as follows (under the identification between distributive laws and monads in monads):

$$\mathbf{DWriter_M}(*, T, \eta^T, \mu^T) := ((*, \mathbf{Writer_M}, \eta^{\mathbf{M}}, \mu^{\mathbf{M}}), (T, \sigma^T), \eta^T, \mu^T)$$

On 1-cells of monads $(\Phi, \phi) : (*, T, \eta^T, \mu^T) \to (*, S, \eta^S, \mu^S)$, the functor acts as

$$\mathbf{DWriter_M}(\Phi, \phi) := ((\Phi, \sigma^\Phi), \phi)$$

where σ^Φ is the tensorial strength the functor Φ is canonically endowed with.

Finally, on 2-cells $\gamma : (\Phi, \phi) \to (\Psi, \psi)$ the functor acts as

$$\mathbf{DWriter_M}(\gamma) = \gamma$$

Checking that this is indeed a 2-functor is a long exercise. $\qquad\square$

6 Example: the Exception monad

In this section we will show that the monad computer scientists call "Either" admits a parametric distributive law. Our strategy will be similar to what we did for the Writer monad in the section above: we will directly show that any other monad lifts to its Eilenberg-Moore category; everything else will trivially follow.

We start by defining the writer monad and proving an equivalent definition that will prove extremely useful:

Definition 6.1. *The Either monad*

Let $\mathcal{C}$ be a category with an initial object $\emptyset$ and all binary coproducts $A+B$. Then for any object $X : \mathcal{C}$ we can define a functor

$$\mathbf{Either}_X : \mathcal{C} \to \mathcal{C}$$

$$\mathbf{Either}_X(A \xrightarrow{f} B) := (X + A) \xrightarrow{X+f} (x + B)$$

We also define natural transformations

$$\eta : \mathbf{Id} \to \mathbf{Either}_X$$

$$\mu : \mathbf{Either}_X^2 \to \mathbf{Either}_X$$

by setting

$$\eta_A = r : A \to X + A$$

$$\mu_A = \langle \ell, id \rangle : X + (X + A) \to X + A$$

Naturality for η and μ is trivial to verify. This makes the triple $(\mathbf{Either}_x, \eta, \mu)$ into a monad; this is the either monad.

Proof. We need to prove that η and μ indeed satisfy unitality and associativity. We start with unitality:

$$
\begin{array}{ccccc}
& & X + A & & \\
& {\scriptstyle \langle \ell, r \circ r \rangle} \swarrow & \| & \searrow {\scriptstyle r} & \\
X + (X + A) & & \| & & X + (X + A) \\
& {\scriptstyle \langle \ell, id \rangle} \searrow & \| & \swarrow {\scriptstyle \langle \ell, id \rangle} & \\
& & X + A & &
\end{array}
$$

We want to prove that

$$\langle \ell, id \rangle \circ \langle \ell, r \circ r \rangle = id = \langle \ell, id \rangle \circ r$$

Recall that, by the universal property of coproducts, for any $\phi : B + C \to D : \mathcal{C}$, we get

$$\phi = \langle (\phi \circ \ell), (\phi \circ r) \rangle$$

So we can just compute:

$$\langle \ell, id \rangle \circ r = id$$

showing right unitality, and

$$\langle \ell, id \rangle \circ \langle \ell, (r \circ r) \rangle \circ \ell = \langle \ell, r \rangle \circ \ell = \ell$$

$$\langle \ell, id \rangle \circ \langle \ell, (r \circ r) \rangle \circ r = \langle \ell, id \rangle \circ r \circ r = id \circ r = r$$

showing left unitality (since $id = \langle \ell, r \rangle$).

The argument for associativity is similar:

$$X + (X + (X + A)) \xrightarrow{\langle \ell, r \circ \langle \ell, id \rangle \rangle} X + (X + A)$$

$$\downarrow {\scriptstyle \langle \ell, id \rangle} \qquad\qquad\qquad\qquad \downarrow {\scriptstyle \langle \ell, id \rangle}$$

$$X + (X + A) \xrightarrow[\langle \ell, id \rangle]{} X + A$$

We want to show the two morphisms are equal, and by the aforementioned universal property it's enough to show they are equal when composed with the coprojections. So we can compute:

$$\langle \ell, id \rangle \circ \langle \ell, id \rangle \circ \ell = \langle \ell, id \rangle \circ \ell = \ell$$

$$\langle \ell, id \rangle \circ \langle \ell, r \circ \langle \ell, id \rangle \rangle \circ \ell = \langle \ell, id \rangle \circ \ell = \ell$$

and similarly

$$\langle \ell, id \rangle \circ \langle \ell, id \rangle \circ r = \langle \ell, id \rangle \circ id = \langle \ell, id \rangle$$

$$\langle \ell, id \rangle \circ \langle \ell, r \circ \langle \ell, id \rangle \rangle \circ r = \langle \ell, id \rangle \circ r \circ \langle \ell, id \rangle = id \circ \langle \ell, id \rangle = \langle \ell, id \rangle$$

$\square$

There is a more compact way of defining this monad, as the monad whose algebras are modules over monoids in the cocartesian monoidal structure $(\mathcal{C}, +, \emptyset, \dots)$ (which is the monoidal structure induced by the initial object $\emptyset$ and coproducts $+$). We'll go through the details below, as this characterization will turn out very useful.

Definition 6.2. *Cocartesian monoidal categories*

Let $\mathcal{C}$ admit an initial object $\emptyset$ and binary coproducts $+$. Then we can endow $\mathcal{C}$ with a monoidal structure $(\mathcal{C}, +, \emptyset, \lambda, \rho, \alpha)$, called the cocartesian monoidal structure.

Proof. We need to define the unitors λ, ρ and the associator α, as well as to prove the triangle and pentagon identities.

1. $\lambda_A : \emptyset + A \xrightarrow{\langle \epsilon_A, id \rangle} A$, where $\epsilon_A : \emptyset \to A$ is the initial morphism;

2. $\rho_A : A + \emptyset \xrightarrow{\langle id, \epsilon_A \rangle} A$ with the same definition for ϵ_A;

3. $\alpha_{A,B,C} : (A + B) + C \xrightarrow{\langle \langle \ell, r \circ \ell \rangle, r \circ r \rangle} A + (B + C)$

each with the obvious inverse. The triangle identity is then equivalently

$$\rho + B = (A + \lambda) \circ \alpha : (A + \emptyset) + B \to A + B$$

The two morphisms are equal if and only if they are equal when composed with all the canonical injections

$$\ell \circ \ell : A \to (A + \emptyset) + B \; ; \; \ell \circ r : \emptyset \to (A + \emptyset) + B \; ; \; r : B \to (A + \emptyset) + B$$

Before proceeding with the proof, recall that given $f : X \to X'$, $g : Y \to Y'$, we have

$$(f + g) := \langle \ell \circ f, r \circ g \rangle : X + Y \to X' + Y'$$

We can now compute:

1. $\ell \circ \ell$:

$$(\rho + b) \circ \ell \circ \ell = \langle \ell \circ \rho, r \rangle \circ \ell \circ \ell = \ell \circ \rho \circ \ell = \ell \circ \langle id, \epsilon \rangle \circ \ell = \ell \circ id = \ell$$

$$(A+\lambda) \circ \alpha \circ \ell \circ \ell = (A+\lambda) \circ \langle \langle \ell, r \circ \ell \rangle, r \circ r \rangle \circ \ell \circ \ell = (A+\lambda) \circ \langle \ell, r \circ \ell \rangle \circ \ell = (A+\lambda) \circ \ell = \langle \ell, r \circ \lambda \rangle \circ \ell = \ell$$

2. $\ell \circ r$: since the domain is $\emptyset$, the claim follows from initiality

3. r:
$$(\rho + B) \circ r = \langle \ell \circ \rho, r \rangle \circ r = r$$

$$(A+\lambda) \circ \alpha \circ r = (A+\lambda) \circ \langle \langle \ell, r \circ \ell \rangle, r \circ r \rangle \circ r = (A+\lambda) \circ r \circ r = \langle \ell, r \circ \lambda \rangle \circ r \circ r = r \circ \lambda \circ r = r \circ \langle \epsilon, id \rangle \circ r = r \circ id = r$$

A similar argument allows us to prove the pentagon identity; we omit the explicit details for the sake of brevity. $\square$

Lemma 6.3. *Monoids in cocartesian monoidal categories*

Let $(\mathcal{C}, +, \emptyset)$ be a cocartesian monoidal category. There is an isomorphism of categories

$$\mathbf{Mon}(\mathcal{C}, +, \emptyset, ...) \simeq \mathcal{C}$$

Proof. We will only define a functor $\mathcal{F} : \mathcal{C} \to \mathbf{Mon}(\mathcal{C}, +, \emptyset)$ which is inverse to the usual forgetful $\mathcal{U} : \mathbf{Mon}(\mathcal{C}, +, \emptyset) \to \mathcal{C}$.

$$\mathcal{F}(A) := (A, \epsilon, \langle id, id \rangle)$$

is clearly seen to be a monoid object: for unitality

$$
\begin{array}{ccccc}
\emptyset + A & \xleftarrow{\ r\ } & A & \xrightarrow{\ \ell\ } & A + \emptyset \\
{\scriptstyle \epsilon+A}\downarrow & & \| & & \downarrow{\scriptstyle A+\epsilon} \\
A + A & \xrightarrow[\langle id,id \rangle]{} & A & \xleftarrow[\langle id,id \rangle]{} & A + A
\end{array}
$$

we can compute the following:

$$\langle id, id \rangle \circ (\epsilon + A) \circ r = \langle id, id \rangle \circ \langle \ell \circ \epsilon, r \rangle \circ r = \langle id, id \rangle \circ r = id$$

$$\langle id, id \rangle \circ (A + \epsilon) \circ \ell = \langle id, id \rangle \circ \langle \ell, r \circ \epsilon \rangle \circ \ell = \langle id, id \rangle \circ \ell = id$$

The argument for proving associativity is similar (compose with the canonical injections, then compute) but quite long; we'll omit it. The same applies to the proof that $\mathcal{F}(f) := f$ is a well-defined morphism of monoid objects, for any morphism f in $\mathcal{C}$.

We now need to prove that the functor $\mathcal{F}$ we just defined is the inverse to the usual forgetful functor $\mathcal{U}$. That $\mathcal{U} \circ \mathcal{F} = \mathbf{Id}$ is trivial; we need to prove $\mathcal{F} \circ \mathcal{U} = \mathbf{Id}$.

On objects, this means that for any given monoid $(A, u, m) : \mathbf{Mon}(\mathcal{C}, +, \emptyset, \dots)$ we have $u = \epsilon$ and $m = \langle id, id \rangle$: the first statement follows from initiality of $\emptyset$, while the second statement is a consequence of unitality for (A, ϵ, m). Indeed, we know that

$$m = \langle m \circ \ell, m \circ r \rangle : A + A \to A$$

and by our previous computation, $(\epsilon + A) \circ r = r$ and $(A + \epsilon) \circ \ell = \ell$; from this and unitality, it follows that

$$m \circ (\epsilon + A) \circ r = m \circ r = id = m \circ \ell = m \circ (A + \epsilon) \circ \ell$$

which implies $m = \langle id, id \rangle$.

Finally, the statement that $\mathcal{F}(\mathcal{U}(f)) = f$ is trivial. $\qquad\square$

It should now be clear that the **Either** monad on a given object is the one obtained by "tensoring" (i.e. performing the coproduct) with that object (with the monad structure given by the unique monoid structure that object is endowed with).

Lemma 6.4. *Modules over monoids in cocartesian monoidal categories*

Let $\mathcal{C}$ a cocartesian monoidal category, and let $A : \mathcal{C}$ be any object (so that $(A, \epsilon, \langle id, id \rangle)$ is a monoid). Then we have an isomorphism of categories

$$\mathbf{Mod}_{(A, \epsilon, \langle id, id \rangle)}(\mathcal{C}, +, \emptyset, \dots) \simeq A \downarrow \mathcal{C}$$

between the category of modules over a monoid and the coslice category.

Proof. We define functors

$$\mathcal{F} : \mathbf{Mod}_{(A, \epsilon, \langle id, id \rangle)}(\mathcal{C}, +, \emptyset, \dots) \to A \downarrow \mathcal{C}$$

$$\mathcal{G} : A \downarrow \mathcal{C} \to \mathbf{Mod}_{(A, \epsilon, \langle id, id \rangle)}(\mathcal{C}, +, \emptyset, \dots)$$

and prove they are inverses of each other.

For $\mathcal{F}$, let first $(X, \sigma) : \mathbf{Mod}_{(A, \epsilon, \langle id, id \rangle)}(\mathcal{C}, +, \emptyset, \dots)$. This means that $X : \mathcal{C}$ and $\sigma : A + X \to X$, so there exist morphisms $\phi : A \to X$ and $\psi : X \to X$ in $\mathcal{C}$ such that $\sigma = \langle \phi, \psi \rangle$. The fact that $(X, \sigma) : \mathbf{Mod}_{(A, \epsilon, \langle id, id \rangle)}(\mathcal{C}, +, \emptyset, \dots)$ means that the following two diagrams commute in $\mathcal{C}$:

$$
\begin{array}{ccccc}
\emptyset + X & \xrightarrow{\epsilon + X} & A + X & \xleftarrow{A + \sigma} & A + (A + X) \\
& {\scriptstyle\lambda}\searrow & \downarrow{\scriptstyle\sigma} & & \downarrow{\scriptstyle\alpha} \\
& & & & (A + A) + X \\
& & & & \downarrow{\scriptstyle\langle id, id\rangle + X} \\
& & X & \xleftarrow{\sigma} & A + X
\end{array}
$$

42

Commutativity of the triangle on the left means that, in particular

$$\sigma \circ (\epsilon + X) \circ r = \lambda \circ r$$

and a routine computation shows that

$$\sigma \circ (\epsilon + X) \circ r = \psi$$

$$\lambda \circ r = id$$

so that $\psi = id$. Notably, $\sigma \circ (\epsilon + X) \circ \ell = \lambda \circ \ell$ is trivial because of $\emptyset$'s initiality. Similarly, the pentagon on the right is also always commutative: (X, σ) is an object of $\mathbf{Mod}_{(A, \epsilon, \langle id, id \rangle)}(\mathcal{C}, +, \emptyset, \dots)$ if and only if $\sigma = \langle \phi, id \rangle$. This allows us to define

$$\mathcal{F}(X, \sigma) := (X, \sigma \circ \ell) = (X, \phi)$$

which is an object of $A \downarrow \mathcal{C}$.

Given $(X, \sigma), (Y, \gamma) : \mathbf{Mod}_{(A, \epsilon, \langle id, id \rangle)}(\mathcal{C}, +, \emptyset, \dots)$ and $f : X \to Y : \mathcal{C}$, the condition for f to be a morphism $f : (X, \sigma) \to (Y, \gamma)$ is

$$
\begin{array}{ccc}
A + X & \xrightarrow{A+f} & A + Y \\
\sigma \downarrow & & \downarrow \gamma \\
X & \xrightarrow{\quad f \quad} & Y
\end{array}
$$

Since we proved that $\sigma = \langle \phi, id \rangle$ (and similarly, $\gamma = \langle \psi, id \rangle$), the above condition reduces to $f \circ \phi = \psi$. This means that $f : (X, \phi) \to (Y, \psi) : A \downarrow \mathcal{C}$. This completes the definition of $\mathcal{F}$.

It's easy to see that the definition of $\mathcal{G}$ is similarly straightforward: we can define $\mathcal{G}(X, \phi) := (X, \langle \phi, id \rangle)$ and $\mathcal{G}(f) = f$. Indeed, the only non-trivial condition on $\mathcal{G}(X, \phi)$ is unitality (which is satisfied because the action is given by $\langle \phi, id \rangle$), and the condition on $\mathcal{G}(f)$ is equivalent to the condition f satisfies as a morphism in the coslice $A \downarrow \mathcal{C}$.

Checking that the above two functors are (strict) inverses of each other is now completely trivial. $\qquad \square$

This means, since $\mathbf{Either}_A$ is the monad for modules over A (or more precisely, the canonical monoid structure A is endowed with), that its Eilenberg-Moore category is given by the coslice in the previous lemma.

Lemma 6.5. *Lifting pointed endofunctors*

Let $A : \mathcal{C}$ be an object in a cocartesian monoidal category, and $\mathcal{U} : A \downarrow \mathcal{C} \to \mathcal{C}$ the forgetful functor. Let also $(\mathcal{F}, u)$ be a pointed endofunctor on $\mathcal{C}$, meaning

$$u : \mathbf{Id} \to \mathcal{F} : \mathcal{C} \to \mathcal{C}$$

There is a pointed endofunctor $(\hat{\mathcal{F}}, \hat{u})$ on $A \downarrow \mathcal{C}$ such that $\mathcal{U} \circ \hat{\mathcal{F}} = \mathcal{F} \circ \mathcal{U}$.

Proof. We first define $\hat{\mathcal{F}} : A \downarrow \mathcal{C} \to A \downarrow \mathcal{C}$, by defining for any $(X, \phi) : A \downarrow \mathcal{C}$ the object

$$\hat{\mathcal{F}}(X, \phi) := (\mathcal{F}(X), \mathcal{F}(\phi) \circ u_A)$$

Given a morphism $f : (X, \phi) \to (Y, \psi) : A \downarrow \mathcal{C}$, (that is, a morphism $f : X \to Y : \mathcal{C}$ such that $f \circ \psi = \phi$) we get a morphism

$$\hat{\mathcal{F}}(f) := \mathcal{F}(f) : \hat{\mathcal{F}}(X, \phi) \to \hat{\mathcal{F}}(Y, \psi) : A \downarrow \mathcal{C}$$

since

$$\mathcal{F}(f) \circ \mathcal{F}(\phi) \circ u_A = \mathcal{F}(f \circ \phi) \circ u_A = \mathcal{F}(\psi) \circ u_A$$

The fact $\hat{\mathcal{F}}$ is a functor (i.e. preserves identities and compositions) follows trivially from the fact that $\mathcal{F}$ is a functor.

We can now define $\hat{u} : \mathbf{Id} \to \hat{\mathcal{F}}$ as $\hat{u}_{(X, \phi)} := u_X$. To check that $\hat{u}_{(X, \phi)} : (X, \phi) \to \mathcal{F}(X, \phi)$, we need to prove that the following square commutes for any $(X, \phi) : A \downarrow \mathcal{C}$:

$$
\begin{array}{ccc}
A & \xrightarrow{u_A} & \mathcal{F}(A) \\
\phi \downarrow & & \downarrow \mathcal{F}(\phi) \\
X & \xrightarrow{u_X} & \mathcal{F}(X)
\end{array}
$$

This follows directly from u being a natural transformation. The proof that $\hat{u}$ is indeed a natural transformation is completely analogous. Hence, $(\hat{\mathcal{F}}, \hat{u})$ is a pointed endofunctor of $A \downarrow \mathcal{C}$; furthermore, unwinding the definitions of $\mathcal{U}$ and $\hat{\mathcal{F}}$ proves that

$$\mathcal{U} \circ \hat{\mathcal{F}} = \mathcal{F} \circ \mathcal{U}$$

concluding the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 6.6. *Lifting monads*

With $\mathcal{C}$, A, $\mathcal{U}$ as in the above lemma, let $\mathcal{T} = (T, \eta, \mu)$ be a monad on $\mathcal{C}$. There exists a monad $\hat{\mathcal{T}} = (\hat{T}, \hat{\eta}, \hat{\mu})$ on $A \downarrow \mathcal{C}$ such that $\mathcal{U} \circ \hat{T} = T \circ \mathcal{U}$.

Proof. Since (T, η, μ) is a monad, (T, η) is a pointed endofunctor of $\mathcal{C}$: we can invoke 6.5 and construct a pointed endofunctor $(\hat{T}, \hat{\eta})$ of $A \downarrow \mathcal{C}$ satisfying the equation $\mathcal{U} \circ \hat{T} = T \circ \mathcal{U}$; we just need to enhance $(\hat{T}, \hat{\eta})$ to a monad $(\hat{T}, \hat{\eta}, \hat{\mu})$.

Start by defining $\hat{\mu}_{(X, \phi)} := \mu_X$; to see that this defines a morphism in $A \downarrow \mathcal{C}$, we need to check that

$$\mu_X \circ T^{\circ 2}(\phi) \circ T(\eta_A) \circ \eta_A = T(\phi) \circ \eta_A$$

that is, the following diagram commutes

$$
\begin{array}{ccccc}
A & \xrightarrow{\eta_A} & T(A) & \xrightarrow{T(\phi)} & T(X) \\
\eta_A \downarrow & \diagup\diagup & \uparrow \mu_A & & \uparrow \mu_X \\
T(A) & \xrightarrow[T(\eta_A)]{} & T^{\circ 2}(A) & \xrightarrow[T^{\circ 2}(\phi)]{} & T^{\circ 2}(X)
\end{array}
$$

44

The triangle on the upper left corner commutes by definition, the other one by unitality for (T, η, μ). The square commutes by naturality of μ.

Finally, the proof that $(\hat{T}, \hat{\eta}, \hat{\mu})$ is a monad trivially reduces to the statement that (T, η, μ) is a monad, which we assumed. $\qquad\square$

As noted in section 2, such lifts correspond 1-1 to distributive laws; we now unroll the constructions and describe such distributive laws explicitly, as well as showing how they can be made "parametric"

Theorem 6.7. *Parametric distributive law for Either*

Let $\mathcal{C}$ be a cocartesian monoidal category, and $A : \mathcal{C}$. Define $\mathcal{B} := \mathbb{B}\mathbf{Cat}_*[\mathcal{C}, \mathcal{C}]$ to be the 2-category obtained by delooping the monoidal category of pointed endofunctors of $\mathcal{C}$, and $\mathcal{D} := \mathbf{Mnd}(\mathcal{B})$ (which are just monads in $\mathcal{C}$, since the point is required to agree with the monad's unit).

There is an object in $PDist(\mathcal{D}, \mathcal{B})$, that is a strict 2-functor

$$(\mathbf{Mnd}(\mathbb{B}\mathbf{Cat}_*[\mathcal{C}, \mathcal{C}]) \otimes_l \mathbb{B}\Delta_\alpha) \otimes_l \mathbb{B}\Delta_\alpha \to \mathbb{B}\mathbf{Cat}_*[\mathcal{C}, \mathcal{C}]$$

such that the underlying monads, encoded as strict 2-functors

$$\mathbf{Mnd}(\mathbb{B}\mathbf{Cat}_*[\mathcal{B}, \mathcal{B}]) \to \mathbf{Mnd}(\mathbb{B}\mathbf{Cat}_*[\mathcal{B}, \mathcal{B}])$$

Proof. Such a 2-functor corresponds to a 2-functor

$$\mathbf{DEither}_X : \mathbf{Mnd}(\mathbb{B}\mathbf{Cat}[\mathcal{C}, \mathcal{C}]) \to \mathbf{Dist}(\mathbb{B}\mathbf{Cat}[\mathcal{C}, \mathcal{C}])$$

defined in a very similar way to 5.4: we employ lemma 6.6 for its action on objects and lemma 6.5 for its action on 1-cells; the action on 2-cells is trivial. $\quad\square$

References

[Bec69] Jon Beck. "Distributive laws". In: *Seminar on Triples and Categorical Homology Theory.* Ed. by B. Eckmann. Springer Berlin Heidelberg, 1969, pp. 119–140. ISBN: 978-3-540-36091-9.

[Bén67] Jean Bénabou. "Introduction to bicategories". In: *Reports of the Midwest Category Seminar.* Berlin, Heidelberg: Springer Berlin Heidelberg, 1967, pp. 1–77. ISBN: 978-3-540-35545-8.

[BKP89] R. Blackwell, G.M. Kelly, and A.J. Power. "Two-dimensional monad theory". In: *Journal of Pure and Applied Algebra* 59.1 (1989), pp. 1–41. ISSN: 0022-4049.

[Bor94] Francis Borceux. *Handbook of Categorical Algebra.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1994.

[Che11] Eugenia Cheng. "Iterated distributive laws". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 150.3 (2011), pp. 459–487. DOI: 10.1017/S0305004110000599.

[CM93]   Pietro Cenciarelli and Eugenio Moggi. "A syntactic approach to modularity in denotational semantics". In: *CTCS 1993* (1993).

[Cru08]   Geoff S. H. Cruttwell. "Normed Spaces and the Change of Base for Enriched Categories". PhD thesis. Dalhousie University, 2008.

[God58]   Roger Godement. "Topologie algébrique et théorie des faisceaux". In: (1958).

[Gra74]   John Walker Gray. *Formal category theory: adjointness for 2-categories*. Springer Berlin, Heidelberg, 1974. ISBN: 978-3-540-06830-3.

[htta]   Tim Campion (https://mathoverflow.net/users/2362/tim-campion). *3-functoriality of the lax Gray tensor product*. MathOverflow. URL:https://mathoverflow.net/q/45936 (version: 2023-11-29). eprint: `https://mathoverflow.net/q/459360`. URL: `https://mathoverflow.net/q/459360`.

[httb]   Theo Johnson-Freyd (https://mathoverflow.net/users/78/theo-johnson-freyd). *Does every bicategory have a "delaxing object"?* MathOverflow. URL:https://mathoverflow.net/q/227590 (version: 2016-01-04). eprint: `https://mathoverflow.net/q/227590`. URL: `https://mathoverflow.net/q/227590`.

[Hub61]   Peter J Huber. "Homotopy theory in general categories". PhD thesis. ETH Zurich, 1961.

[Lei98]   Tom Leinster. *Basic Bicategories*. 1998. arXiv: `math/9810017 [math.CT]`. URL: `https://arxiv.org/abs/math/9810017`.

[LHJ95]   Sheng Liang, Paul Hudak, and Mark Jones. "Monad transformers and modular interpreters". In: *Proceedings of the 22nd ACM SIGPLAN-SIGACT symposium on Principles of programming languages*. 1995, pp. 333–343.

[Mog91]   Eugenio Moggi. "Notions of computation and monads". In: *Information and computation* 93.1 (1991), pp. 55–92.

[MU22]   Dylan McDermott and Tarmo Uustalu. "What Makes a Strong Monad?" In: *Electronic Proceedings in Theoretical Computer Science* 360 (June 2022), pp. 113–133. ISSN: 2075-2180. DOI: `10.4204/eptcs.360.6`. URL: `http://dx.doi.org/10.4204/EPTCS.360.6`.

[Rat13]   Kruna Segrt Ratkovic. *Morita theory in enriched context*. 2013. arXiv: `1302.2774 [math.CT]`. URL: `https://arxiv.org/abs/1302.2774`.

[Str74]   Ross Street. "Elementary cosmoi I". In: *Category Seminar*. Ed. by Gregory M. Kelly. Berlin, Heidelberg: Springer Berlin Heidelberg, 1974, pp. 134–180. ISBN: 978-3-540-37270-7.