Monads and Distributive Laws in Substructural Contexts

Soichiro Fujii
National Institute of Informatics, Tokyo, Japan

Yun Chen Tsai
National Institute of Informatics, Tokyo, Japan
SOKENDAI (The Graduate University for Advanced Studies), Kanagawa, Japan

Yoàv Montacute
National Institute of Informatics, Tokyo, Japan

Ichiro Hasuo
National Institute of Informatics, Tokyo, Japan
SOKENDAI (The Graduate University for Advanced Studies), Kanagawa, Japan
Imiron Co., Ltd., Tokyo, Japan

Abstract

We present a categorical theory of monads and distributive laws *in substructural contexts*. In the study of distributive laws, the roles of (the absence of) structural rules for variable contexts have been recognized; our theory formalizes these substructural situations using Tronin's *verbal categories* $\mathbf{W}$, in a uniform and presentation-independent manner. We introduce the classes of $\mathbf{W}$-*operadic monads* (those *defined* via the structural rules in $\mathbf{W}$) and of $\mathbf{W}$-*commutative monads* (those *invariant* under the structural rules in $\mathbf{W}$). We give a canonical construction of a distributive law $ST \to TS$ of monads on **Set**; it is applicable when S is $\mathbf{W}$-operadic and T is $\mathbf{W}$-commutative (under mild conditions). This accounts for many known and new distributive laws. Even when S fails to be $\mathbf{W}$-operadic, we can *refine* S and force $\mathbf{W}$-operadicity; this captures Varacca and Winskel's construction of indexed valuations.

2012 ACM Subject Classification Theory of computation → Categorical semantics; Theory of computation → Probabilistic computation; Theory of computation → Program semantics

Keywords and phrases Monad, distributive law, operad, category theory, effect

Digital Object Identifier 10.4230/LIPIcs.LICS.2026.45

Related Version *Extended Version*: `https://arxiv.org/abs/2605.13533v1` [14]

Funding All authors are supported by ASPIRE Grant No. JPMJAP2301, JST, Japan.
Yoàv Montacute: Supported by ACT-X, Grant No. JPMJAX24CR, JST, Japan.

Acknowledgements We thank the anonymous referees for their helpful comments.

1 Introduction

Monads. In program semantics and, more generally, in system behavior modeling, the categorical notion of *monad* has been successfully used as a uniform interface to various computational *effects*. Examples of effects include non-functional *side effects* in functional programming [43], additional behavioral layers for reactive systems such as (possibilistic or probabilistic) nondeterminism [2, 21], and so on. The mathematical theory of monads – together with a closely related theory of *universal algebra* – not only offers unified foundations for effects, but also enables novel programming principles such as *effect handlers* [48].

Distributive laws. In this context, *distributive laws* between monads have attracted attention. Originally a mathematical construct [3], distributive laws enable us to *combine effects*. Specifically, given monads S and T on the category $\mathbf{Set}$ of sets (in this paper, we focus on monads on $\mathbf{Set}$), a distributive law $\delta\colon ST \to TS$ of S over T equips the composite TS with a monad structure, with the consequence that unified monadic frameworks for programming, static analysis, verification, etc. accommodate the coexistence of two different layers of effects modeled by S and T.

The question of obtaining a distributive law has intrigued many theoreticians, too. This was initiated by Plotkin's observation that there is no distributive law $DP \to PD$ of the (finitely supported) probability distribution monad D over the powerset monad P (see [62]). This observation sparked much work: both on the positive side (obtaining distributive laws [11, 45, 17]) and the negative side (proving "no-go" theorems [27, 66]).

Two approaches: weak distributive laws and refinement. On the positive side, towards obtaining a distributive law, there have mainly been two technical approaches. The first "weak distributive law" approach, when a distributive law $ST \to TS$ does not exist, finds it sufficient to have a *weak distributive law* $\delta\colon ST \to TS$. Here "weak" means δ's compatibility with the unit η^S of S is dropped. A weak distributive law δ induces a combination of S and T, but this time it is not given simply by the composite TS. Roughly speaking, this approach modifies T: concretely, for $S = D$ and $T = P$, the combination essentially replaces $T = P$ with the convex powerset monad over the Eilenberg–Moore category $\mathbf{Set}^D$ [15, 17, 4, 1, 56].

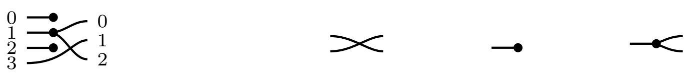
The second "refinement" approach modifies S instead. It finds a reason for the failure of a distributive law $\delta\colon ST \to TS$ in *equations* in (an equational presentation of) S, and thus proposes to use a *refinement* $\widehat{S}$ of S – obtained by dropping the problematic equations – in place of S. This approach is initiated in [62] for specific monads (variations of D and P – see § 6.3). Generalization of this approach is sought in [11, 45]. We follow this line of work.

Categorical theory of substructural contexts for monads and distributive laws. In this paper, we introduce a general framework for the aforementioned "refinement" approach to distributive laws, centered around the idea of *substructural contexts*. Our categorical framework uses *verbal categories* [57] for modeling substructural contexts.

The relevance of structural rules (exchange, weakening, and contraction) to algebraic theories (and thus monads) is widely acknowledged, explicitly or implicitly [10, 20, 33]. In particular, [11, 45, 38] study the roles of structural rules in distributive laws. Our framework here offers a unifying categorical picture parametrized by a verbal category $\mathbf{W}$.

We sketch our theory in the coming paragraphs.

Substructural contexts and verbal categories (§ 3). Restricting structural rules – *exchange* (E), *weakening* (W), and *contraction* (C) – in variable contexts is a common tool in logic. In the categorical study of algebraic theories, operads, and binding syntax, it is common to model the application of structural rules with 1) morphisms of $\mathbf{F}$, the category of finite ordinals and functions between them [13], and 2) their string diagram presentation [10, 33]. For example, the context transformation from $x_0, x_1, x_2 \vdash p(x_0, x_1, x_2)$ to $x_0, x_1, x_2, x_3 \vdash p(x_1, x_3, x_1)$ – obtained with E, W, and C for a term p – is represented by the function $\alpha\colon 3 \to 4$ with $\alpha(0) = \alpha(2) = 1$ and $\alpha(1) = 3$; the corresponding string diagram (in which we have $\alpha(i) = j$ iff i on the right is connected to j on the left) is depicted on the left of (1) below.



(1)

The structural rules E, W, and C correspond to the three "parts" of the string diagram as on the right of (1), respectively (see also [10, § 2] for a related discussion).

Conveniently, considering a *substructural context* – i.e. prohibiting the use of certain structural rules – corresponds to singling out a wide subcategory of $\mathbf{F}$. Concretely, allowing the rules designated below on the left corresponds to the subcategories of $\mathbf{F}$ on the right.

$$\begin{tikzcd} \emptyset \arrow[r] \arrow[d] & \{\mathrm{E}\} \arrow[r] \arrow[d] & \{\mathrm{E},\mathrm{C}\} \arrow[d] \\ \{\mathrm{W}\} \arrow[r] & \{\mathrm{E},\mathrm{W}\} \arrow[r] & \{\mathrm{E},\mathrm{W},\mathrm{C}\} \end{tikzcd} \qquad \begin{tikzcd} \mathbf{F}_{\mathrm{id}} \arrow[r] \arrow[d] & \mathbf{F}_{\mathrm{bij}} \arrow[r] \arrow[d] & \mathbf{F}_{\mathrm{surj}} \arrow[d] \\ \mathbf{F}_{\mathrm{m,inj}} \arrow[r] & \mathbf{F}_{\mathrm{inj}} \arrow[r] & \mathbf{F} \end{tikzcd} \tag{2}$$

Here, $\mathbf{F}_{\mathrm{id}}$ has only identity functions as morphisms; $\mathbf{F}_{\mathrm{m,inj}}$ has monotone injections (with respect to the total order $0 < 1 < \cdots < n-1$); $\mathbf{F}_{\mathrm{bij}}$ has bijections; $\mathbf{F}_{\mathrm{inj}}$ has injections; $\mathbf{F}_{\mathrm{surj}}$ has surjections. (The settings $\{\mathrm{C}\}$ and $\{\mathrm{W},\mathrm{C}\}$ are ill-behaved and thus omitted; see Ex. 10.)

The notion of *verbal category* [57] axiomatizes (wide) subcategories of $\mathbf{F}$ suited for modeling substructural contexts. Examples include the six categories in (2) (see also Ex. 9).

$\mathbf{W}$-operads and their monads (§ 3). The notion of $\mathbf{W}$-*operad* in [57] embodies the idea of "algebraic theories in the $\mathbf{W}$-substructural context". Its definition follows the same format – with id and subst – as for (symmetric or non-symmetric) operads.

For our theory, we introduce the notion of $\mathbf{W}$-*operadic monad*, as those which are induced by a $\mathbf{W}$-operad via a left Kan extension. This formalizes the idea of "a monad in the $\mathbf{W}$-substructural context". (For example, $\mathbf{F}$-operadic monads coincide with finitary monads.)

$\mathbf{W}$-commutative monads (§ 4). Our construction of a distributive law $ST \to TS$ (Thm. 42) requires, very roughly speaking, that

- S is a monad defined using the structural rules in $\mathbf{W}$, and
- T's effect is invariant under the structural rules in $\mathbf{W}$.

We just formalized the former as $\mathbf{W}$-operadic monads; for the latter, we introduce the notion of $\mathbf{W}$-*commutative monad*. The definition of $\mathbf{W}$-commutativity is by a suitable diagram, one similar to those used in [45, 11, 46]. It unifies some known notions such as (ordinary) *commutativity* [28], *affinity* [29], and *relevance* [20].

Canonical distributive laws and application (§§ 5–7). One of our main results in this paper is Thm. 42 where, as announced above, we present a general construction of a distributive law $\delta\colon ST \to TS$, assuming that

- S is a $\mathbf{W}$-operadic monad, and
- T is a $\mathbf{W}$-commutative monad.

The construction is categorical and canonical, along the coend description of (the left Kan extension used in) a $\mathbf{W}$-operadic monad.

The construction of this canonical δ is uniform for any verbal category $\mathbf{W}$. An additional condition is required for δ to be compatible with the multiplication μ^T of T: it is that $\mathbf{W}$ accommodates the exchange rule (E) – unless S is so lean that it is generated by operations of arity ≤ 1. See Thm. 42 for a precise statement.

Given the last condition (of accommodating (E)), it is natural for us to focus on the following four verbal categories.

$$\begin{tikzcd} & \mathbf{F} & \\ \mathbf{F}_{\mathrm{inj}} \arrow[ur] & & \mathbf{F}_{\mathrm{surj}} \arrow[ul] \\ & \mathbf{F}_{\mathrm{bij}} \arrow[ul] \arrow[ur] & \end{tikzcd} \tag{3}$$

We also show the following *monotonicity* results:

- $\mathbf{W}$-operadicity is up-closed (i.e. if S is $\mathbf{W}$-operadic and $\mathbf{W} \subseteq \mathbf{W}'$, then S is $\mathbf{W}'$-operadic, Prop. 19(2)); and
- $\mathbf{W}$-commutativity is down-closed (Rem. 33).

Therefore, for the application of the general construction in Thm. 42, we can look for $\mathbf{W}$ in the intersection of the upper set of S's $\mathbf{W}$-operadicity and the lower set of T's $\mathbf{W}$-commutativity.

This "find an intersection" strategy has two usages. The first usage is negative: given a pair (S, T), the empty intersection of the above two sets shows that the canonical distributive law in Thm. 42 does not work. While this does not prove a no-go result (other distributive laws may exist), it is often easy to check and thus a useful red flag. Of course, the intersection is empty for the combinations with known no-go results.

The second use is a positive one: when the intersection is empty, we can modify S or T for a nonempty intersection, so that the canonical distributive law works. While modifying T is possible – techniques for forcing $\mathbf{W}$-commutativity on a monad are studied for some $\mathbf{W}$, see e.g. [20, 6, 36] – we are interested in modifying S in this paper. This modification of S must extend the upper set of $\mathbf{W}$-operadicity downwards, so that a new monad relies on fewer structural rules.

To do so, we introduce a categorical construction, namely the $\mathbf{W}$*-operadic refinement* $\mathrm{Rf}_{\mathbf{W}}(S)$ of a monad S (§ 6). It is the monad induced, in the Kan extension-way described in § 3, by the $\mathbf{W}$-operad given by *restricting* S. The refinement $\mathrm{Rf}_{\mathbf{W}}(S)$ is $\mathbf{W}$-operadic by construction. We show that the *indexed valuation monad* IV in [62] – the first example of the "refinement" approach for distributive laws – is indeed an instance of $\mathbf{W}$-operadic refinement.

Contributions and organization. We introduce a theory of substructural contexts for monads and distributive laws. It is categorical and uniform, unifying previous observations for specific substructural contexts [45, 11, 46, 28, 29, 20] and the axiomatization by verbal categories $\mathbf{W}$ [57]. The theory has the following main components.

- After reviewing verbal categories [57], we introduce the classes of $\mathbf{W}$*-operadic* and $\mathbf{W}$*-commutative* monads (§§ 3 and 4).
- We present a construction of a *canonical distributive law* $\delta\colon ST \to TS$ for $\mathbf{W}$-operadic S and $\mathbf{W}$-commutative T (Thm. 42).
- We present the notion of $\mathbf{W}$*-operadic refinement* of a monad (§ 6). It can be used to force $\mathbf{W}$-operadicity required by Thm. 42.
- We discuss many examples of general constructs. These include various "multiset monads" (Exs 4, 5, and 23–25) and new distributive laws (e.g. one that resolves a problem in [61], see Ex. 50) obtained by forcing $\mathbf{W}$-operadicity (§ 7).

Most proofs are deferred to [14, Appendix D].

Related work. Many related works have been discussed in the context of our motivation, or will be discussed later in a suitable technical context. In particular, the relationship to closely related works [38, 45, 11] is discussed further in § 5.3.

Here are some other related works. A big motivation for distributive laws comes from the combination of (possibilistic) nondeterminism and probabilistic branching. This combination is observed e.g. in Markov decision processes (MDPs), a fundamental model for probabilistic model checking and reinforcement learning. Categorical studies of this specific combination include [22, 32, 44, 42, 41].

Once program/system semantics is defined by monads, reasoning about it with *program logic* is the next step. Recent works in this direction include [64, 35, 7, 8, 39, 51].

2 Preliminaries

The set of all natural numbers (or finite ordinals) is denoted by $\mathbb{N} = \{0, 1, 2, \dots\}$. Each $n \in \mathbb{N}$ is identified with the n-element set $\{0, 1, \dots, n-1\}$. The category of sets and functions between them is denoted by **Set**. It has a symmetric monoidal structure $(1, \times)$ via products. For any set X, we denote by $!_X \colon X \to 1$ the unique function from X to 1.

Since our focus in this paper is on monads on **Set**, by a *monad* we shall always mean a *monad on* **Set**. Specifically, a *monad* $\mathbb{T}$ (on **Set**) is a triple $\mathbb{T} = (T, \eta, \mu)$ where $T \colon \mathbf{Set} \to \mathbf{Set}$ is a functor (called the *functor part* of $\mathbb{T}$) and $\eta \colon \mathrm{id}_{\mathbf{Set}} \to T$ and $\mu \colon TT \to T$ are natural transformations (called the *unit* and *multiplication* of $\mathbb{T}$, respectively), such that $\mu \circ T\eta = \mathrm{id}_T = \mu \circ \eta T$ and $\mu \circ T\mu = \mu \circ \mu T$ hold. Here are some examples of monads.

▶ **Example 1** (C-exception monad $\mathbb{E}^C$)**.** Any set C has a unique monoid structure $(C, e_C \colon 0 \to C, \nabla_C \colon C + C \to C)$ with respect to coproducts. This induces a monad $\mathbb{E}^C$ called the *C-exception monad*. Its functor part is $C + (-) \colon \mathbf{Set} \to \mathbf{Set}$, and its unit and multiplication are $\eta_X = e_C + X \colon X \to C + X$ and $\mu_X = \nabla_C + X \colon C + C + X \to C + X$ for each set X. ⌟

▶ **Example 2** (C-reader monad $\mathbb{R}^C$)**.** Any set C also admits a unique comonoid structure $(C, !_C \colon C \to 1, \Delta_C \colon C \to C \times C)$ with respect to products. It induces the *C-reader monad* $\mathbb{R}^C$ in the same way as in Ex. 1: its functor part is $(-)^C \colon \mathbf{Set} \to \mathbf{Set}$, and its unit and multiplication are $\eta_X = X^{!_C} \colon X \to X^C$ and $\mu_X = X^{\Delta_C} \colon (X^C)^C \cong X^{C \times C} \to X^C$. ⌟

▶ **Example 3** (M-writer monad $\mathbb{W}^{\mathsf{M}}$)**.** Let $\mathsf{M} = (M, 1 \in M, \cdot \colon M \times M \to M)$ be a monoid (with respect to products). It induces the monad $\mathbb{W}^{\mathsf{M}}$ called the M-*writer monad*, whose functor part is $M \times (-) \colon \mathbf{Set} \to \mathbf{Set}$, and whose unit and multiplication are defined by $\eta_X(x) = (1, x)$ and $\mu_X(m, n, x) = (mn, x)$ for each set X, $x \in X$, and $m, n \in M$. ⌟

Recall that a *semiring* is a tuple $\mathsf{S} = (S, 0 \in S, + \colon S \times S \to S, 1 \in S, \cdot \colon S \times S \to S)$ where S is a set, $(S, 0, +)$ is a commutative monoid, $(S, 1, \cdot)$ is a monoid, and such that $\cdot \colon (S, 0, +) \times (S, 0, +) \to (S, 0, +)$ is a bihomomorphism of monoids, i.e., for all $x, y, z \in S$, $x \cdot 0 = 0$, $x \cdot (y + z) = x \cdot y + x \cdot z$, $0 \cdot z = 0$, and $(x + y) \cdot z = x \cdot z + y \cdot z$ hold. The semiring S is *commutative* if $(S, 1, \cdot)$ is a commutative monoid.

▶ **Example 4** (S-multiset monad $\mathbb{M}^{\mathsf{S}}$)**.** Let $\mathsf{S} = (S, 0, +, 1, \cdot)$ be a semiring. We define the monad $\mathbb{M}^{\mathsf{S}}$ called the S-*multiset monad*. The functor part M^{S} of $\mathbb{M}^{\mathsf{S}}$ maps each set X to the set $M^{\mathsf{S}}X$ of all functions $p \colon X \to S$ such that the set $\{\, x \in X \mid p(x) \neq 0 \,\}$ (called the *support* $\operatorname{supp} p$ of p) is finite, and maps each function $f \colon X \to Y$ to the function $M^{\mathsf{S}}f \colon M^{\mathsf{S}}X \to M^{\mathsf{S}}Y$ defined for each $p \in M^{\mathsf{S}}X$ by

$$\big((M^{\mathsf{S}}f)(p)\big)(y) = \textstyle\sum_{x \in f^{-1}(y) \cap \operatorname{supp} p} p(x).$$

We can also regard $p \in M^{\mathsf{S}}X$ as a formal linear combination

$$\textstyle\sum_{x \in \operatorname{supp} p} p(x) \cdot x \tag{4}$$

of elements of X with coefficients from S; from this viewpoint, the function $M^{\mathsf{S}}f$ maps (4) to $\sum_{x \in \operatorname{supp} p} p(x) \cdot f(x)$.

The unit η_X and multiplication μ_X of $\mathbb{M}^{\mathsf{S}}$ at $X \in \mathbf{Set}$ are given as follows. For each $x \in X$, $\eta_X(x) \in M^{\mathsf{S}}X$ is defined by $\big(\eta_X(x)\big)(x') = 1$ if $x' = x$, and $= 0$ otherwise. That is, $\eta_X(x)$ is x regarded as a trivial linear combination. For each $\varphi \in M^{\mathsf{S}}M^{\mathsf{S}}X$, $\mu_X(\varphi) \in M^{\mathsf{S}}X$ is defined by $\big(\mu_X(\varphi)\big)(x) = \sum_{p \in \operatorname{supp} \varphi} \varphi(p) \cdot p(x)$. This flattens a "linear combination of linear combinations" φ into a linear combination.

Here are some instances of $\mathbb{M}^{\mathsf{S}}$.

(1) When S is the semiring $\mathsf{N} = (\mathbb{N}, 0, +, 1, \cdot)$ of natural numbers, $\mathbb{M}^{\mathsf{N}}$ is called the *(finite) multiset monad* $\mathbb{M} = (M, \eta, \mu)$. For a set X, MX is the set of all finite multisets over X.

(2) When S is the semiring $2 = (\{0,1\}, 0, \vee, 1, \wedge)$, $\mathbb{M}^2$ is called the *finite powerset monad* $\mathbb{P}_{\mathrm{fin}} = (P_{\mathrm{fin}}, \eta, \mu)$. For each set X, $P_{\mathrm{fin}}X$ is the set of all finite subsets of X.

(3) When S is the semiring $\mathsf{R}_{\geq 0} = (\mathbb{R}_{\geq 0}, 0, +, 1, \cdot)$ of nonnegative real numbers, $\mathbb{M}^{\mathsf{R}_{\geq 0}}$ is called the *(finite) valuation monad* (see e.g. [62]) and is written as $\mathbb{V} = (V, \eta, \mu)$.

Note that the semirings N, 2, and $\mathsf{R}_{\geq 0}$ are all commutative. ⌟

▶ **Example 5** (affine S-multiset monad $\mathbb{A}^{\mathsf{S}}$)**.** Let $\mathsf{S} = (S, 0, +, 1, \cdot)$ be a semiring. We define the monad $\mathbb{A}^{\mathsf{S}} = (A^{\mathsf{S}}, \eta, \mu)$ called the *affine* S*-multiset monad.* This is the submonad of the S-multiset monad $\mathbb{M}^{\mathsf{S}}$ such that, for each set X, $A^{\mathsf{S}}X$ consists of all formal *affine* combinations of elements of X, i.e., $p \in M^{\mathsf{S}}X$ with $\sum_{x \in \mathrm{supp}\, p} p(x) = 1$ (cf. (4)). ($\mathbb{A}^{\mathsf{S}}$ is the *affine part* of $\mathbb{M}^{\mathsf{S}}$ in the sense of [36].) Here are the instances with $\mathsf{S} = 2$ (Ex. 4(2)) and $\mathsf{R}_{\geq 0}$ (Ex. 4(3)).

(1) $\mathbb{A}^2$ is the *nonempty finite powerset monad* $\mathbb{P}^+_{\mathrm{fin}}$.

(2) $\mathbb{A}^{\mathsf{R}_{\geq 0}}$ is the *(finitely supported) probability distribution monad* $\mathbb{D}$ (cf. [31, § 16]). ⌟

3 W-Operadic Monads

In this section, we will recall $\mathbf{W}$*-operads* [57, Def. 2], where $\mathbf{W}$ is any *verbal category* in the sense of [57, Def. 1]. We then define monads induced by $\mathbf{W}$-operads, and the "change of base" constructions for $\mathbf{W}$-operads.

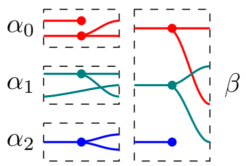
3.1 Verbal Categories

Towards a formal definition of verbal categories (Def. 8), we start with an informal introduction. Recall from § 1 that $\mathbf{F}$ is the category of finite ordinals and all functions between them, and that string diagrams represent morphisms in $\mathbf{F}$ (see (1)).

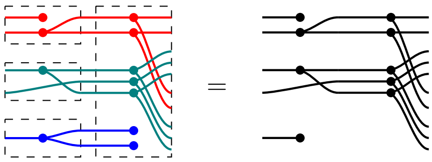
A verbal category $\mathbf{W}$ is defined to be a wide subcategory of $\mathbf{F}$ – meaning that $\mathrm{ob}\,\mathbf{W} = \mathrm{ob}\,\mathbf{F}$ – that is closed under what we call the *verbal substitution* operation $\star$ on morphisms of $\mathbf{F}$. We sketch the verbal substitution operation $\star$. Its type is as follows.

$$\frac{\beta\colon n' \to n \qquad \alpha_0\colon m'_0 \to m_0 \qquad \cdots \qquad \alpha_{n-1}\colon m'_{n-1} \to m_{n-1}}{\beta \star \vec{\alpha}\colon \sum_{i' \in n'} m'_{\beta(i')} \to \sum_{i \in n} m_i}$$

Here, $\vec{\alpha} = (\alpha_i)_{i \in n}$. Its concrete definition (Def. 7) is illustrated as follows. Let $\beta\colon 4 \to 3$, $\alpha_0\colon 2 \to 2$, $\alpha_1\colon 3 \to 2$, and $\alpha_2\colon 2 \to 1$, as depicted below.



The morphism $\beta \star \vec{\alpha}$ is then obtained by 1) finding the correspondence between α_i and the i-th end (on the left) of β, for each $i \in 3$ (we used colors to highlight this), 2) "duplicating" wires in β so that the numbers of wires match between β and $\vec{\alpha}$, and 3) connecting them.

$\beta \star \vec{\alpha}$ = 

In order to define $\star$ formally, we fix specific coproduct diagrams in the category $\mathbf{F}$.

▶ **Definition 6** (chosen coproducts in $\mathbf{F}$). *Let $n \in \mathbb{N}$, $\vec{m} = (m_i \in \mathbb{N})_{i \in n}$, and $m = \sum_{i \in n} m_i$. For each $i \in n$, we define the function*

$$\iota_{\vec{m},i} \colon m_i \to m$$

by $\iota_{\vec{m},i}(j) = \left(\sum_{k \in i} m_k\right) + j$ for each $j \in m_i$. The family $(\iota_{\vec{m},i} \colon m_i \to m)_{i \in n}$ exhibits m as a coproduct of $\vec{m}$ in $\mathbf{F}$. ⌟

▶ **Definition 7** ($\beta \star \vec{\alpha}$; cf. [53, § 2.2] and Prop. 21). *Given morphisms $\left(\beta \colon n' \to n, \vec{\alpha} = (\alpha_i \colon m'_i \to m_i)_{i \in n}\right)$ in $\mathbf{F}$, define $m' = \sum_{i' \in n'} m'_{\beta(i')}$ and $m = \sum_{i \in n} m_i$. We define*

$$\beta \star \vec{\alpha} \colon m' \to m$$

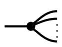
as the unique morphism in $\mathbf{F}$ making the following square commute for each $i' \in n'$.

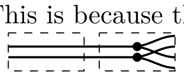
$$\begin{tikzcd}
m'_{\beta(i')} \arrow[r, "\alpha_{\beta(i')}"] \arrow[d, "\iota_{\vec{m}'\beta,i'}"'] & m_{\beta(i')} \arrow[d, "\iota_{\vec{m},\beta(i')}"] \\
m' \arrow[r, "\beta \star \vec{\alpha}"] & m
\end{tikzcd}$$

Here, $\vec{m}'\beta = (m'_{\beta(i')})_{i' \in n'}$ and $\vec{m} = (m_i)_{i \in n}$. ⌟

▶ **Definition 8** (verbal category). *A wide subcategory $\mathbf{W}$ of $\mathbf{F}$ is called a* verbal category *[57, Def. 1] if it is closed under the verbal substitution operation $\beta \star \vec{\alpha}$ of Def. 7. We say that a verbal category is* symmetric *if it contains $\mathbf{F}_{\mathrm{bij}}$ (see (2)).* ⌟

To see the equivalence of the above definition of verbal categories to the original [57, Def. 1], consider the cases of $\beta \star \vec{\alpha}$ where β or each of α_i is an identity map. We note that verbal categories are called *faithful cartesian clubs* in [50, Def. 2.6.3], and that symmetric verbal categories are called *contextual subcategories of* $\mathbf{F}$ in [34].

▶ **Example 9.** As already mentioned, the six categories in (2) are verbal categories [57, p. 748]. For every positive integer n, the verbal category $\mathbf{F}^{(n)}_{\mathrm{surj}}$ generated by $\mathrm{swap} \colon 2 \to 2$ and $!_n \colon n \to 1$ (i.e.,  and  $\}n$) consists of all surjections $\alpha \colon \ell \to m$ such that each fiber $\alpha^{-1}(i)$ has cardinality $k(n-1)+1$ for some $k \in \mathbb{N}$; this yields an infinite family of verbal categories, with $\mathbf{F}^{(1)}_{\mathrm{surj}} = \mathbf{F}_{\mathrm{bij}}$ and $\mathbf{F}^{(2)}_{\mathrm{surj}} = \mathbf{F}_{\mathrm{surj}}$. Intuitively, $\mathbf{F}^{(n)}_{\mathrm{surj}}$ models contexts with exchange and "contraction with arity n". ⌟

▶ **Example 10.** As non-examples, we note that the wide subcategory of $\mathbf{F}$ consisting of all monotone (resp. surjective monotone) maps (corresponding to the set $\{\mathrm{W}, \mathrm{C}\}$ (resp. $\{\mathrm{C}\}$) of structural rules; cf. (2)) is not a verbal category.[1] This is because the function $!_2 \colon 2 \to 1$ is monotone but $!_2 \star (\mathrm{id}_2) \colon 4 \to 2$ is not, as the picture  shows (cf. [10, § 2]). ⌟

[1] This shows that the third example in [57, p. 748] is incorrect.

3.2 W-Operads

A formal definition (Def. 11) of **W**-operads may look complicated. Illustration by string diagram, which will follow, should be more intuitive.

▶ **Definition 11** (**W**-operad [57, Def. 2]). *Let* **W** *be a verbal category. A* **W**-operad $\mathbb{O} = (O, \mathsf{id}, \mathsf{subst})$ *consists of the following.*

- *A functor* $O\colon \mathbf{W} \to \mathbf{Set}$. *For each object* $n \in \mathbf{W}$*, the set* On *is denoted by* O_n*. For each morphism* $\alpha\colon m \to n$ *in* **W** *and* $p \in O_m$*, the element* $(O\alpha)p \in O_n$ *is also denoted by* αp.
- *An element* $\mathsf{id} \in O_1$.
- *For each* $n \in \mathbb{N}$ *and* $\vec{m} = (m_i \in \mathbb{N})_{i\in n}$*, a function*

$$\mathsf{subst}_{n,\vec{m}}\colon O_n \times \textstyle\prod_{i\in n} O_{m_i} \to O_m, \tag{5}$$

where $m = \sum_{i\in n} m_i$. *Its value at* $(q,\vec{p}) \in O_n \times \prod_{i\in n} O_{m_i}$ *is denoted by* $\mathsf{subst}_{n,\vec{m}}(q,\vec{p})$.

These data are subject to the following axioms.

- *The compatibility of the* subst *functions* (5) *with the action by morphisms in* **W***: for any* $\beta\colon n' \to n$ *and* $\vec{\alpha} = (\alpha_i\colon m'_i \to m_i)_{i\in n}$ *in* **W***, the following diagram commutes.*

$$\begin{tikzcd}
O_{n'} \times \prod_{i\in n} O_{m'_i} \arrow[r, "1 \times \prod_{i'\in n'} \pi_{\beta(i')}"] \arrow[d, "O\beta \times \prod_{i\in n} O\alpha_i"'] & O_{n'} \times \prod_{i'\in n'} O_{m'_{\beta(i')}} \arrow[r, "\mathsf{subst}_{n',\vec{m}\beta}"] & O_{m'} \arrow[d, "O(\beta \star \vec{\alpha})"] \\
O_n \times \prod_{i\in n} O_{m_i} \arrow[rr, "\mathsf{subst}_{n,\vec{m}}"] & & O_m
\end{tikzcd} \tag{6}$$

Here, $\pi_i\colon \prod_{i\in n} O_{m'_i} \to O_{m'_i}$ *is the* i*-th projection (and hence* $\pi_{\beta(i')}$ *is the* $\beta(i')$*-th projection),* $\vec{m}\beta = (m_{\beta(i')})_{i'\in n'}$*, and* $\beta \star \vec{\alpha}$ *is from Def. 7.*

- *The unit laws: for each natural number* m *and each* $p \in O_m$*, we have*

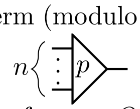
$$\mathsf{subst}_{1,(m)}(\mathsf{id},(p)) = p = \mathsf{subst}_{m,(1)_{i\in m}}(p,(\mathsf{id})_{i\in m}).$$

- *The associativity law:*

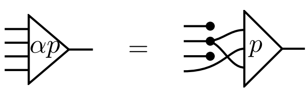
$$\mathsf{subst}_{n,\vec{m}}\big(r,\ \big(\mathsf{subst}_{m_j,\vec{\ell}^{(j)}}(q_j,\vec{p}^{(j)})\big)_{j\in n}\big) = \mathsf{subst}_{m,\vec{\ell}}\big(\mathsf{subst}_{n,\vec{m}}(r,\vec{q}),\ \vec{p}\big),$$

for all $n \in \mathbb{N}$, $\vec{m} = (m_j \in \mathbb{N})_{j\in n}$, $\vec{\ell} = (\vec{\ell}^{(j)})_{j\in n} = \big((\ell_i^{(j)} \in \mathbb{N})_{i\in m_j}\big)_{j\in n}$, $r \in O_n$, $\vec{q} = (q_j \in O_{m_j})_{j\in n}$*, and* $\vec{p} = (\vec{p}^{(j)})_{j\in n} = \big((p_i^{(j)} \in O_{\ell_i^{(j)}})_{i\in m_j}\big)_{j\in n}$*, where* $m = \sum_{j\in n} m_j$. ⌟

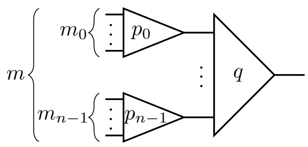
Note that the unit and associativity laws in Def. 11 are identical to those for (symmetric or non-symmetric) operads and multicategories (see e.g. [33, Def. 2.1.1]).

The idea is that a **W**-operad $\mathbb{O}$ encodes an algebraic theory in the "substructural context" corresponding to **W**. More specifically, an element p of O_n represents an n-ary term (modulo provable equality) in the algebraic theory, and can be depicted as the "gadget" 

with n input wires and one output wire. The action of morphisms of **W** on the functor O amounts to rearranging the input wires. For example, if $\mathbf{W} = \mathbf{F}$ and $\alpha\colon 3 \to 4$ is the function depicted on the left of (1), then $\alpha p \in O_4$ can be obtained from $p \in O_3$ as follows.



Syntactically, the 4-ary term αp is defined from the 3-ary term p as $(\alpha p)(x_0, x_1, x_2, x_3) = p(x_1, x_3, x_1)$. We depict $\mathsf{id} \in O_1$ as a wire (i.e. id = ——) and $\mathsf{subst}_{n,\vec{m}}(q, \vec{p})$ as follows.



(7)

Syntactically, id corresponds to the variable x_0 regarded as a unary term, and the term $\mathsf{subst}_{n,\vec{m}}(q, \vec{p})$ is obtained by suitably *substituting* $\vec{p}$ for the variables in q. For example,

$$\mathsf{subst}_{2,(2,1)}\big(\, q(x_0, x_1),\ \big(p_0(x_0, x_1),\ p_1(x_0)\big)\,\big) = q\big(\, p_0(x_0, x_1),\ p_1(x_2)\,\big);$$

note that the variable x_0 in the term p_1 is renamed to x_2. The unit and associativity laws for $\mathbf{W}$-operads will be trivial from these pictorial or syntactic viewpoints; the compatibility axiom (6) can be understood in an analogous manner, too. See [33, Chap. 2] for an extended discussion with many pictures in the case where $\mathbf{W} = \mathbf{F}_{\mathrm{id}}$ or $\mathbf{F}_{\mathrm{bij}}$. We will discuss examples of $\mathbf{W}$-operads in § 3.5.

The general notion of $\mathbf{W}$-operad specializes to the following:

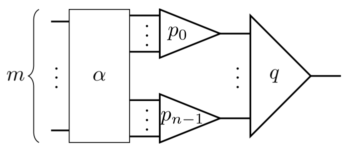
- $\mathbf{F}_{\mathrm{id}}$-operads coincide with *non-symmetric operads* [33],
- $\mathbf{F}_{\mathrm{bij}}$-operads coincide with *symmetric operads* [25],
- $\mathbf{F}_{\mathrm{surj}}$-operads coincide with *regular operads* [53], and
- $\mathbf{F}$-operads coincide with *(abstract) clones* [57, Thm. on p. 751].

The classes of $\mathbf{W}$-operads for various $\mathbf{W}$ can also be regarded as a generalization of the *Boom hierarchy*, whose relevance to distributive laws has been noted; see [26, § 7] and [65, Chap. 6].

▶ Remark 12. For any verbal category $\mathbf{W}$, the functor category $[\mathbf{W}, \mathbf{Set}]$ is known to have a monoidal structure $(J, \bullet)$ called the *substitution monoidal structure*, with the inclusion functor $J \colon \mathbf{W} \to \mathbf{Set}$ as the unit object; see [34, 54, 55] or [10, § 4]. Given $P, Q \in [\mathbf{W}, \mathbf{Set}]$ and $m \in \mathbf{W}$, an element of $(Q \bullet P)_m$ is a tuple

$$\big(\, q \in Q_n,\ (p_i \in P_{m_i})_{i \in n},\ \alpha \in \mathbf{W}(\textstyle\sum_{i \in n} m_i, m)\,\big) \tag{8}$$

modulo a suitable equivalence relation (see [10] for a coend formula). The data (8) can be depicted as follows.



Using the substitution monoidal structure $(J, \bullet)$, $\mathbf{W}$-operads can be identified with monoid objects in $([\mathbf{W}, \mathbf{Set}], J, \bullet)$.

While the substitution monoidal structure provides a nice conceptual viewpoint (cf. Rem. 18), its definition requires some work and is not necessary for the purposes of this paper. Hence, we give an exposition that does not rely on the substitution monoidal structure. ⌟

3.3 The Monad Induced by a W-Operad

In this section, we show that every **W**-operad $\mathbb{O}$ induces a monad $\mathrm{Mnd}(\mathbb{O})$ on **Set**. Intuitively, $\mathrm{Mnd}(\mathbb{O})$ is the (finitary) monad corresponding to the algebraic theory expressed by $\mathbb{O}$. One way of making this precise is to proceed as follows. Define the notion of $\mathbb{O}$-*algebra* in **Set** (as in [58, Def. 1.1]), show that the forgetful functor $\mathbf{Alg}(\mathbb{O}) \to \mathbf{Set}$ from the category $\mathbf{Alg}(\mathbb{O})$ of $\mathbb{O}$-algebras is monadic, and then let $\mathrm{Mnd}(\mathbb{O})$ be the induced monad on **Set**.

Instead, here we use left Kan extensions to define $\mathrm{Mnd}(\mathbb{O})$. This is useful for describing the canonical distributive laws later.

▶ **Definition 13** ($\mathrm{Mnd}(\mathbb{O})$)**.** *Let* $\mathbf{W}$ *be a verbal category and* $\mathbb{O} = (O, \mathsf{id}, \mathsf{subst})$ *be a* $\mathbf{W}$*-operad. The monad* $\mathrm{Mnd}(\mathbb{O})$ induced by $\mathbb{O}$ *is*

$$\mathrm{Mnd}(\mathbb{O}) = (\mathrm{Lan}_J O, \eta, \mu), \tag{9}$$

using the unit and multiplication described below. ⌟

Here, $\mathrm{Lan}_J O$ in (9) is the left Kan extension of O along the inclusion functor J:

$$\begin{tikzcd} \mathbf{W} \arrow[r, "O"] \arrow[d, "J"'] & \mathbf{Set} \\ \mathbf{Set} \arrow[ur, "\mathrm{Lan}_J O"'] \arrow[u, phantom, shift left=4, "\Downarrow"] & \end{tikzcd} \tag{10}$$

Concretely, for each $X \in \mathbf{Set}$, the set $(\mathrm{Lan}_J O)X$ is the coend

$$\int^{n \in \mathbf{W}} O_n \times X^n \tag{11}$$

(cf. e.g. [19, Prop. 4.1]), which can be constructed as follows. First form the set

$$\sum_{n \in \mathbb{N}} O_n \times X^n, \tag{12}$$

whose element we write as $\big(p, (x_i)_{i \in n}\big)$ where $p \in O_n$ and $x_i \in X$ (with $n \in \mathbb{N}$), or sometimes as $(p, \vec{x})$. The set $(\mathrm{Lan}_J O)X$ is the quotient of the set (12) under the smallest equivalence relation $\sim$ on (12) such that

$$\big(p, (x_{\alpha(i)})_{i \in m}\big) \sim \big(\alpha p, (x_j)_{j \in n}\big) \tag{13}$$

for each morphism $\alpha \colon m \to n$ in $\mathbf{W}$, $p \in O_m$, and $(x_j)_{j \in n} \in X^n$. (Writing $(x_{\alpha(i)})_{i \in m}$ as $\vec{x}\alpha$, (13) can be expressed concisely as $(p, \vec{x}\alpha) \sim (\alpha p, \vec{x})$.) The $\sim$-equivalence class containing $(p, \vec{x})$ is denoted by $[p, \vec{x}]$. Given a function $f \colon X \to Y$ between sets, the function $(\mathrm{Lan}_J O)f \colon (\mathrm{Lan}_J O)X \to (\mathrm{Lan}_J O)Y$ maps each $\big[p, (x_i)_{i \in n}\big] \in (\mathrm{Lan}_J O)X$ to $\big[p, (f x_i)_{i \in n}\big]$.

The unit η of the monad $\mathrm{Mnd}(\mathbb{O})$ at $X \in \mathbf{Set}$ is the function $\eta_X \colon X \to (\mathrm{Lan}_J O)X$ mapping each $x \in X$ to $[\mathsf{id} \in O_1, x \in X^1] \in (\mathrm{Lan}_J O)X$. The naturality of η is clear.

The multiplication μ of $\mathrm{Mnd}(\mathbb{O})$ at X is $\mu_X \colon (\mathrm{Lan}_J O)(\mathrm{Lan}_J O)X \to (\mathrm{Lan}_J O)X$ given by

$$\begin{aligned} &\big[\, q \in O_n,\ \big([p_j \in O_{m_j},\ (x_i^{(j)})_{i \in m_j} \in X^{m_j}]\big)_{j \in n} \in \big((\mathrm{Lan}_J O)X\big)^n \,\big] \\ &\qquad\qquad \longmapsto \quad \big[\, \mathsf{subst}_{n, \vec{m}}(q, \vec{p}) \in O_m,\ \big((x_i^{(j)})_{i \in m_j}\big)_{j \in n} \in X^m \,\big], \end{aligned}$$

where $\vec{m} = (m_j)_{j \in n}$, $m = \sum_{j \in n} m_j$, and $\vec{p} = (p_j)_{j \in n}$. Well-definedness of μ_X follows from (6), while the naturality of μ is clear. This definition of μ crucially uses the subst operation of $\mathbb{O}$ – its depiction in (7) with a string diagram should illustrate how μ works.

The monad axioms for $\mathrm{Mnd}(\mathbb{O}) = (\mathrm{Lan}_J O, \eta, \mu)$ follow from the unit and associativity axioms for the $\mathbf{W}$-operad $\mathbb{O}$.

It is not hard to see that the Eilenberg–Moore category of the monad $\mathrm{Mnd}(\mathbb{O})$ coincides with the category $\mathbf{Alg}(\mathbb{O})$ of $\mathbb{O}$-algebras defined in [58, Def. 1.1]. This shows that $\mathrm{Mnd}(\mathbb{O})$ can indeed be obtained by the construction sketched in the beginning of § 3.3. One can also see the coincidence by endowing the set $(\mathrm{Lan}_J O)X$ with the free $\mathbb{O}$-algebra structure over X, as in [58, Thm. 1.4].

▶ **Definition 14** **($\mathbf{W}$-operadicity).** *A monad $\mathbb{S}$ (on* **Set***) is* $\mathbf{W}$-operadic *if there exists a* $\mathbf{W}$*-operad $\mathbb{O}$ with $\mathrm{Mnd}(\mathbb{O}) \cong \mathbb{S}$. When moreover such a $\mathbf{W}$-operad $\mathbb{O}$ is unique up to isomorphism, we say that $\mathbb{S}$ is* uniquely $\mathbf{W}$-operadic*.* ⌟

▶ Remark 15. Intrinsic characterization of $\mathbf{W}$-operadicity is known for some $\mathbf{W}$.

(1) See [63, Def. 3.1 and Thms 10.1 and 2] for a characterization of $\mathbf{F}_{\mathrm{id}}$-operadic monads (also known as *strongly analytic monads*).
(2) See [63, Def. 3.2 and Thms 10.10 and 11] for a characterization of $\mathbf{F}_{\mathrm{bij}}$-operadic monads (also known as *analytic monads*).
(3) See [53, § 2.4] for a characterization of $\mathbf{F}_{\mathrm{surj}}$-operadic monads (also known as *semi-analytic monads* or *finitary taut monads*).
(4) A monad is $\mathbf{F}$-operadic iff it is finitary (folklore). ⌟

We can simplify the coend formula (11) in some cases.

▶ **Proposition 16.**

(1) *When $\mathbf{W} = \mathbf{F}_{\mathrm{id}}$, (11) coincides with (12).*
(2) *When $\mathbf{W} = \mathbf{F}_{\mathrm{bij}}$, (11) is $\sum_{n\in\mathbb{N}} O_n \times_{\mathbf{F}_{\mathrm{bij}}(n,n)} X^n$, where $O_n \times_{\mathbf{F}_{\mathrm{bij}}(n,n)} X^n$ is the quotient of $O_n \times X^n$ with respect to the following equivalence relation [24]:*

$$\{\, ((p, \vec{x}\alpha), (\alpha p, \vec{x})) \mid p \in O_n,\ \vec{x} \in X^n,\ \alpha \in \mathbf{F}_{\mathrm{bij}}(n,n) \,\}.$$

(3) *When $\mathbf{W} = \mathbf{F}_{\mathrm{surj}}$, (11) is $\sum_{S\in P_{\mathrm{fin}}X} O_{|S|}$, where $|S|$ is the cardinality of S.* ⌟

Here is an observation with potential consequences on distributive laws (see Rem. 44).

▶ **Proposition 17.** *When $\mathbf{W}$ is $\mathbf{F}_{\mathrm{bij}}$, $\mathbf{F}_{\mathrm{surj}}$, or $\mathbf{F}$, any $\mathbf{W}$-operadic monad is uniquely $\mathbf{W}$-operadic. When $\mathbf{W}$ is $\mathbf{F}_{\mathrm{id}}$, $\mathbf{F}_{\mathrm{m,inj}}$, or $\mathbf{F}_{\mathrm{inj}}$, there exists a $\mathbf{W}$-operadic monad which is not uniquely $\mathbf{W}$-operadic.* ⌟

▶ Remark 18. Recall Rem. 12. The left Kan extension and restriction along the inclusion functor $J\colon \mathbf{W} \to \mathbf{Set}$ induces a *lax monoidal* adjunction

$$([\mathbf{Set},\mathbf{Set}], 1, \circ) \underset{[J,\mathbf{Set}]}{\overset{\mathrm{Lan}_J}{\underset{\longrightarrow}{\overset{\longleftarrow}{\perp}}}} ([\mathbf{W},\mathbf{Set}], J, \bullet), \tag{14}$$

where the monoidal structure $(1, \circ)$ on $[\mathbf{Set},\mathbf{Set}]$ is given by composition of endofunctors (cf. [10, § 5]). It then follows that the adjunction (14) lifts to the categories of monoid objects:

$$\mathbf{Mon}([\mathbf{Set},\mathbf{Set}], 1, \circ) \underset{\mathbf{Mon}([J,\mathbf{Set}])}{\overset{\mathbf{Mon}(\mathrm{Lan}_J)}{\underset{\longrightarrow}{\overset{\longleftarrow}{\perp}}}} \mathbf{Mon}([\mathbf{W},\mathbf{Set}], J, \bullet).$$

Note that $\mathbf{Mon}([\mathbf{Set},\mathbf{Set}], 1, \circ)$ is the category $\mathbf{Mnd}(\mathbf{Set})$ of monads on $\mathbf{Set}$, whereas $\mathbf{Mon}([\mathbf{W},\mathbf{Set}], J, \bullet)$ is the category $\mathbf{W}$**-Opd** of $\mathbf{W}$-operads. Our construction $\mathrm{Mnd}(-)$ is $\mathbf{Mon}(\mathrm{Lan}_J)$. We discuss the functor $\mathbf{Mon}([J,\mathbf{Set}])$ in § 6; see in particular Prop. 57. ⌟

3.4 Change of Base

When one verbal category $\mathbf{W}$ is a subcategory of another $\mathbf{W}'$, we obtain two-way constructions between $\mathbf{W}$-operads and $\mathbf{W}'$-operads – this is much like the one in Rem. 18 between $\mathbf{W}$-operads and monads. We describe these constructions, together with their consequence on the relationship with monads.

Let $\mathbf{W}$ and $\mathbf{W}'$ be verbal categories such that $\mathbf{W}$ is a subcategory of $\mathbf{W}'$; this fixes the inclusion functor $K\colon \mathbf{W} \to \mathbf{W}'$. Firstly, given a $\mathbf{W}'$-operad $\mathbb{O}'$, it is easy to see that the *restriction* of $\mathbb{O}'$ along K is a $\mathbf{W}$-operad. Its underlying functor is given by $\mathbf{W} \xrightarrow{K} \mathbf{W}' \xrightarrow{O'} \mathbf{Set}$.

Conversely, given a $\mathbf{W}$-operad $\mathbb{O}$, we define its *extension* $\mathbb{O}' = (O', \mathsf{id}^{\mathbb{O}'}, \mathsf{subst}^{\mathbb{O}'})$ along K which is a $\mathbf{W}'$-operad. Its underlying functor is given, much like in (10), by the left Kan extension $O' = \mathrm{Lan}_K O$. Concretely, for any $n' \in \mathbf{W}'$, we have

$$O'_{n'} = (\mathrm{Lan}_K O)_{n'} \cong \textstyle\int^{n\in\mathbf{W}} O_n \times \mathbf{W}'(n, n'),$$

whose element is denoted by $[\, p \in O_n,\ \alpha \in \mathbf{W}'(n, n')\,]$.

The unit $\mathsf{id}^{\mathbb{O}'}$ for the extension $\mathbb{O}'$ is $[\,\mathsf{id}^{\mathbb{O}} \in O_1,\ \mathrm{id}_1 \in \mathbf{W}'(1, 1)\,] \in O'_1$. For $n' \in \mathbb{N}$ and $\vec{m}' = (m'_{i'} \in \mathbb{N})_{i'\in n'}$, the substitution operation

$$\mathsf{subst}^{\mathbb{O}'}_{n',\vec{m}'}\colon O'_{n'} \times \textstyle\prod_{i'\in n'} O'_{m'_{i'}} \to O'_{m'}$$

for $\mathbb{O}'$ (with $m' = \sum_{i'\in n'} m'_{i'}$) is given by

$$\begin{aligned}\big(\,[q \in O_n,\ \beta \in \mathbf{W}'(n, n')],\ \big([p_{i'} \in O_{m_{i'}},\ \alpha_{i'} \in \mathbf{W}'(m_{i'}, m'_{i'})]\big)_{i'\in n'}\,\big)&\\ \longmapsto\quad \big[\,\mathsf{subst}^{\mathbb{O}}_{n,\vec{m}\beta}(q, \vec{p}\beta) \in O_m,\ \beta \star \vec{\alpha} \in \mathbf{W}'(m, m')\,\big],&\end{aligned}$$

where $\vec{m}\beta = (m_{\beta(i)})_{i\in n}$, $\vec{p}\beta = (p_{\beta(i)})_{i\in n}$, $m = \sum_{i\in n} m_{\beta(i)}$, and $\vec{\alpha} = (\alpha_{i'})_{i'\in n'}$. The use of $\star$, from Def. 7, is crucial here.

Regarding the relationship to § 3.3, we have the following.

▶ **Proposition 19.**

(1) *Let $K\colon \mathbf{W} \to \mathbf{W}'$ be an inclusion of verbal categories and $\mathbb{O}'$ be the extension of a $\mathbf{W}$-operad $\mathbb{O}$ along K. Then $\mathbb{O}'$ is a $\mathbf{W}'$-operad and* $\mathrm{Mnd}(\mathbb{O}') \cong \mathrm{Mnd}(\mathbb{O})$.

(2) *Consequently, any $\mathbf{W}$-operadic monad is $\mathbf{W}'$-operadic, too.* ⌟

From the viewpoint of algebraic theories, a $\mathbf{W}$-operad $\mathbb{O}$ and its *extension* $\mathbb{O}'$ express the same algebraic theory. In contrast, the *restriction* construction may modify the theory.

3.5 Examples

Here we look at examples of $\mathbf{W}$-operads and $\mathbf{W}$-operadic monads. In general, for any verbal category $\mathbf{W}$, the functor $\Delta^{\mathbf{W}}_1\colon \mathbf{W} \to \mathbf{Set}$ constant at the terminal object $1 \in \mathbf{Set}$ has a unique structure of a $\mathbf{W}$-operad; we call it the *terminal* $\mathbf{W}$*-operad* and write $\Delta^{\mathbf{W}}_1$ for it.

▶ **Example 20 ($\mathbf{W} = \mathbf{F}_{\mathrm{id}}$).** Using Prop. 16(1), we obtain the following.

- The *list monad* $(-)^*$ (corresponding to the theory of monoids) is induced by the terminal $\mathbf{F}_{\mathrm{id}}$-operad $\Delta^{\mathbf{F}_{\mathrm{id}}}_1$.
- The C-exception monad $\mathbb{E}^C$ (Ex. 1) is induced by an $\mathbf{F}_{\mathrm{id}}$-operad; specifically $\mathbb{O} = (O, \mathsf{id}, \mathsf{subst})$ with $O_0 = C$, $O_1 = \{\mathsf{id}\}$, and $O_2 = O_3 = \cdots = \emptyset$.
- The M-writer monad $\mathbb{W}^{\mathsf{M}}$ (Ex. 3) is induced by an $\mathbf{F}_{\mathrm{id}}$-operad $\mathbb{O} = (O, \mathsf{id}, \mathsf{subst})$ with $O_0 = \emptyset$, $O_1 = M$ (where id is the unit $1 \in M$ of the monoid M), and $O_2 = O_3 = \cdots = \emptyset$.

Hence these monads are $\mathbf{F}_{\mathrm{id}}$-operadic. See [33, Chap. 2] for more examples. It follows from Rem. 15(1) that every $\mathbf{F}_{\mathrm{id}}$-operadic monad is *cartesian* in the sense of [33, Def. 4.1.1]. ⌟

We revisit the verbal substitution operation $\star$ (Def. 7).

▶ **Proposition 21.** *Let* $\mathbf{W}$ *be a verbal category and* $\mathbb{O}^{\mathrm{mon}}$ *be the* $\mathbf{W}$*-operad obtained from the terminal* $\mathbf{F}_{\mathrm{id}}$*-operad* $\Delta_1^{\mathbf{F}_{\mathrm{id}}}$ *via extension (§ 3.4).*

(1) $\mathbb{O}^{\mathrm{mon}}$ *is the* $\mathbf{W}$-operad for monoids*, with the underlying functor* $O^{\mathrm{mon}}_n = \sum_{n' \in \mathbb{N}} \mathbf{W}(n', n)$*.*

(2) *The substitution operation of* $\mathbb{O}^{\mathrm{mon}}$ *coincides with the verbal substitution operation* $\star$ *in* $\mathbf{W}$*, i.e.,* $\mathsf{subst}^{\mathbb{O}^{\mathrm{mon}}}(\beta, \vec{\alpha}) = \beta \star \vec{\alpha}$*.* ⌟

The intuition for Prop. 21(1) is that $\alpha \colon n' \to n$ in $\mathbf{W}$ represents a word of length n' over the alphabet $n = \{0, 1, \dots, n-1\}$ in the $\mathbf{W}$-substructural context.

We return to examples of $\mathbf{W}$-operads.

▶ **Example 22** ($\mathbf{W} = \mathbf{F}_{\mathrm{m,inj}}$)**.** Whereas not much seems to be known about $\mathbf{F}_{\mathrm{m,inj}}$-operads, we note the following.

- The terminal monad constant at $1 \in \mathbf{Set}$ (corresponding to the algebraic theory with a constant symbol c and the axiom $x = c$) is induced by the terminal $\mathbf{F}_{\mathrm{m,inj}}$-operad $\Delta_1^{\mathbf{F}_{\mathrm{m,inj}}}$.
- The monad $1 + (-)^*$ for the algebraic theory of *monoids with zero* ($\mathsf{P} = (P, 1, \cdot, 0)$ such that $(P, 1, \cdot)$ is a monoid and $0 \in P$ satisfies $0 \cdot x = 0 = x \cdot 0$ for all $x \in P$) is $\mathbf{F}_{\mathrm{m,inj}}$-operadic.

We can also see that these $\mathbf{F}_{\mathrm{m,inj}}$-operadic monads are *not* $\mathbf{F}_{\mathrm{id}}$-operadic, via the fact that every $\mathbf{F}_{\mathrm{id}}$-operadic monad is cartesian (see Ex. 20): the terminal monad is not cartesian since its unit is not; the monad $1 + (-)^*$ is not cartesian since its multiplication is not. ⌟

▶ **Example 23** ($\mathbf{W} = \mathbf{F}_{\mathrm{bij}}$)**.** The following monads are $\mathbf{F}_{\mathrm{bij}}$-operadic.

- The multiset monad $\mathbb{M}$ (Ex. 4(1)). It is induced by the terminal $\mathbf{F}_{\mathrm{bij}}$-operad $\Delta_1^{\mathbf{F}_{\mathrm{bij}}}$.
- More generally, the $\mathsf{N[M]}$-multiset monad $\mathbb{M}^{\mathsf{N[M]}}$ (Ex. 4), where $\mathsf{N[M]}$ is the *monoid semiring* of a monoid M; see below.

The monoid semiring construction $\mathsf{N}[-]$ (with natural numbers as coefficients) is an analogue of the well-known *group ring* construction, and is the left adjoint in the following adjunction between the categories $\mathbf{SRing}$ of semirings and $\mathbf{Mon}$ of monoids (folklore):

$$\begin{tikzcd} \mathbf{SRing} \arrow[r, shift right, "(-)_{\mathrm{mult}}"'] \arrow[r, phantom, "\perp"] & \mathbf{Mon}. \arrow[l, shift right, "\mathsf{N}[-]"'] \end{tikzcd} \tag{15}$$

Here, the right adjoint $(-)_{\mathrm{mult}}$ sends each semiring $\mathsf{S} = (S, 0, +, 1, \cdot)$ to its multiplicative monoid $\mathsf{S}_{\mathrm{mult}} = (S, 1, \cdot)$.

Given a monoid M, the $\mathbf{F}_{\mathrm{bij}}$-operad $\mathbb{O}^{\mathsf{M}}$ inducing the $\mathsf{N[M]}$-multiset monad $\mathbb{M}^{\mathsf{N[M]}}$ can be constructed as follows [60, Ex. 3]. Define the functor $O^{\mathsf{M}} \colon \mathbf{F}_{\mathrm{bij}} \to \mathbf{Set}$ with $O^{\mathsf{M}}_n = M^n$; this assignment is *covariantly* functorial because every morphism in $\mathbf{F}_{\mathrm{bij}}$ is invertible. We set $\mathsf{id} = 1 \in M$, and subst by componentwise multiplication: e.g.,

$$\mathsf{subst}_{2,(1,2)}\big((u, v), ((x), (y, z))\big) = (ux, vy, vz).$$

In contrast, the Z-multiset monad $\mathbb{M}^{\mathsf{Z}}$ (corresponding to the theory of abelian groups) is not $\mathbf{F}_{\mathrm{bij}}$-operadic; note that the ring Z of integers is not of the form $\mathsf{N[M]}$. For a rigorous proof of this, one can use Rem. 15(2) and observe that the functor part of the monad $\mathbb{M}^{\mathsf{Z}}$ does not preserve the following (weak) pullback in $\mathbf{Set}$: $\begin{smallmatrix} 0 & \to & 0 \\ \downarrow & & \downarrow \\ 2 & \to & 1 \end{smallmatrix}$. (In fact, this shows that $\mathbb{M}^{\mathsf{Z}}$ is not even $\mathbf{F}_{\mathrm{surj}}$-operadic, by Rem. 15(3).) ⌟

▶ **Example 24** ($\mathbf{W} = \mathbf{F}_{\mathrm{inj}}$)**.** Thanks to Prop. 19(2), any monad that is either $\mathbf{F}_{\mathrm{m,inj}}$-operadic (Ex. 22) or $\mathbf{F}_{\mathrm{bij}}$-operadic (Ex. 23) is $\mathbf{F}_{\mathrm{inj}}$-operadic. Here is an additional class of $\mathbf{F}_{\mathrm{inj}}$-operadic monads, combining certain features of the $\mathbf{F}_{\mathrm{m,inj}}$-operadic monad $1 + (-)^*$ for monoids with zero and the $\mathbf{F}_{\mathrm{bij}}$-operadic monad $\mathbb{M}^{\mathsf{N[M]}}$:

- The $\mathsf{N}_0[\mathsf{P}]$-multiset monad $\mathbb{M}^{\mathsf{N}_0[\mathsf{P}]}$ is $\mathbf{F}_{\text{inj}}$-operadic for any monoid with zero P.

Here, $\mathsf{N}_0[\mathsf{P}]$ is the *contracted monoid semiring* (cf. [9, § 5.2]) of P, which we now explain. The underlying set of $\mathsf{N}_0[\mathsf{P}]$ is the set $M(P \setminus \{0\})$ of all finite multisets (see Ex. 4(1)) over the set $P \setminus \{0\}$ of *nonzero* elements of P. The addition in $\mathsf{N}_0[\mathsf{P}]$ is given by the usual addition of multisets (so the additive commutative monoid of $\mathsf{N}_0[\mathsf{P}]$ is the free commutative monoid over $P \setminus \{0\}$), whereas the multiplication in $\mathsf{N}_0[\mathsf{P}]$ is defined using the multiplication of P. The construction $\mathsf{N}_0[-]$ admits a characterization similar to that for $\mathsf{N}[-]$ in (15): it is the left adjoint of the adjunction $\mathsf{N}_0[-] \dashv (-)_{\text{mult},0} \colon \mathbf{SRing} \to \mathbf{Mon}_0$, where $\mathbf{Mon}_0$ is the category of monoids with zero and the right adjoint $(-)_{\text{mult},0}$ maps each semiring $\mathsf{S} = (S, 0, +, 1, \cdot) \in \mathbf{SRing}$ to $\mathsf{S}_{\text{mult},0} = (S, 1, \cdot, 0) \in \mathbf{Mon}_0$.

An $\mathbf{F}_{\text{inj}}$-operad $\mathbb{O}^{\mathsf{P}}$ inducing the $\mathsf{N}_0[\mathsf{P}]$-multiset monad $\mathbb{M}^{\mathsf{N}_0[\mathsf{P}]}$ can be constructed as follows. The underlying functor $O^{\mathsf{P}} \colon \mathbf{F}_{\text{inj}} \to \mathbf{Set}$ is given by $O^{\mathsf{P}}_n = P^n$. Its action on morphisms of $\mathbf{F}_{\text{inj}}$ is defined using $0 \in P$; for example, if $\alpha \colon 2 \to 3$ is given by $\alpha(0) = 2$ and $\alpha(1) = 0$, then $O^{\mathsf{P}}\alpha$ maps $(p_0, p_1) \in O^{\mathsf{P}}_2$ to $\alpha(p_0, p_1) = (p_1, 0, p_0) \in O^{\mathsf{P}}_3$.[2] The rest of the operad structure of $\mathbb{O}^{\mathsf{P}}$ is similar to that of $\mathbb{O}^{\mathsf{M}}$ in Ex. 23.

The *indexed valuation monad* $\mathbb{IV}$ of [62] turns out to be an instance of $\mathbb{M}^{\mathsf{N}_0[\mathsf{P}]}$, with $\mathsf{P} = (\mathbb{R}_{\geq 0}, 1, \cdot, 0)$. Although it happens to be $\mathbf{F}_{\text{bij}}$-operadic as well – note that $\mathsf{N}_0[(\mathbb{R}_{\geq 0}, 1, \cdot, 0)] \cong \mathsf{N}[(\mathbb{R}_{>0}, 1, \cdot)]$ as semirings – it can be naturally obtained from the valuation monad $\mathbb{V} = \mathbb{M}^{\mathsf{R}_{\geq 0}}$ (Ex. 4(3)) as its $\mathbf{F}_{\text{inj}}$-*operadic refinement*; see § 6.3 for details. ⌟

▶ **Example 25** ($\mathbf{W} = \mathbf{F}_{\text{surj}}$)**.** Using Prop. 16(3), we see that the following monads are $\mathbf{F}_{\text{surj}}$-operadic.

- The finite powerset monad $\mathbb{P}_{\text{fin}}$ (Ex. 4(2)), induced by the terminal $\mathbf{F}_{\text{surj}}$-operad $\Delta_1^{\mathbf{F}_{\text{surj}}}$.
- The nonempty finite powerset monad $\mathbb{P}^+_{\text{fin}}$ (Ex. 5(1)).
- The probability distribution monad $\mathbb{D}$ (Ex. 5(2)).

For example, the probability distribution monad $\mathbb{D}$ is induced by the $\mathbf{F}_{\text{surj}}$-operad $\mathbb{O}$ with $O_n = \{p \in Dn \mid \operatorname{supp} p = n\}$ (see Ex. 4 for supp). More generally, for any semiring $\mathsf{S} = (S, 0, +, 1, \cdot)$ such that $0 \neq 1$ and that $S \setminus \{0\}$ is closed under $+$ and $\cdot$, the S-multiset monad $\mathbb{M}^{\mathsf{S}}$ (Ex. 4) and the affine S-multiset monad $\mathbb{A}^{\mathsf{S}}$ (Ex. 5) are $\mathbf{F}_{\text{surj}}$-operadic. ⌟

▶ **Example 26** ($\mathbf{W} = \mathbf{F}$)**.** As mentioned in Rem. 15(4), every finitary monad is $\mathbf{F}$-operadic. Thus all monads corresponding to finitary algebraic theories (such as those for groups, rings, etc.) are $\mathbf{F}$-operadic. In particular, all monads mentioned in § 2 are $\mathbf{F}$-operadic, with the proviso that for the C-reader monad $\mathbb{R}^C$ (Ex. 2) to be $\mathbf{F}$-operadic, C must be a finite set. ⌟

4 W-Commutative Monads

In this section, we introduce the notion of $\mathbf{W}$-*commutative monad* (on $\mathbf{Set}$) for any verbal category $\mathbf{W}$. The definition (Def. 32) uses a lax monoidal structure (η_1, ψ) on the functor part T of a monad $\mathbb{T} = (T, \eta, \mu)$ – the structure exists even when $\mathbb{T}$ is not commutative. We recall it in § 4.1. For verbal categories $\mathbf{W}$ in (2), the $\mathbf{W}$-commutativity of monads can be characterized by a suitable combination of known conditions on monads, namely (ordinary) commutativity, affinity, and relevance; see Prop. 36.

[2] Whereas letting $\alpha(p_0, p_1) = (p_1, 1, p_0)$ also defines a functor $\mathbf{F}_{\text{inj}} \to \mathbf{Set}$, the axiom (6) fails. A similar argument shows that [58, Ex. 2.3] with $\mathbf{W} = \mathbf{F}$ is incorrect.

4.1 Strengths and Lax Monoidal Structures

In order to define **W**-commutative monads, we first recall standard facts about *strength* and *lax monoidal structure* on a monad on **Set**. See [14, Appendix A] for details.

▶ **Definition 27** (strength τ [28, 30, 40])**.** *Let $T\colon \mathbf{Set} \to \mathbf{Set}$ be a functor. A* left (tensorial) strength *on T is a natural transformation*

$$\tau = \bigl(\tau_{X,Y}\colon X \times TY \to T(X \times Y)\bigr)_{X,Y \in \mathbf{Set}}$$

making the following diagrams commute for all sets X, Y, and Z.

$$\begin{tikzcd} 1 \times TX \arrow[r, "\tau_{1,X}"] \arrow[dr, "\cong"'] & T(1 \times X) \arrow[d, "\cong"] \\ & TX \end{tikzcd} \qquad \begin{tikzcd} (X \times Y) \times TZ \arrow[rr, "\tau_{X \times Y, Z}"] \arrow[d, "\cong"'] & & T\bigl((X \times Y) \times Z\bigr) \arrow[d, "\cong"] \\ X \times (Y \times TZ) \arrow[r, "1 \times \tau_{Y,Z}"] & X \times T(Y \times Z) \arrow[r, "\tau_{X, Y \times Z}"] & T\bigl(X \times (Y \times Z)\bigr) \end{tikzcd}$$

Dually, a right (tensorial) strength *on T is a natural transformation $\tau' = \bigl(\tau'_{X,Y}\colon TX \times Y \to T(X \times Y)\bigr)_{X,Y \in \mathbf{Set}}$ satisfying the dual axioms.* ⌟

▶ Remark 28 (see e.g. [40]). Any functor $T\colon \mathbf{Set} \to \mathbf{Set}$ has a unique left (resp. right) strength τ (resp. τ'). For sets X and Y, $\tau_{X,Y}\colon X \times TY \to T(X \times Y)$ is defined as follows. For each $x \in X$, $\tau_{X,Y}(x, -)\colon TY \to T(X \times Y)$ is Ti_x, where $i_x\colon Y \to X \times Y$ is the function mapping each $y \in Y$ to $(x, y) \in X \times Y$. (See [14, Prop. 68] for details.) ⌟

▶ **Definition 29** (lax monoidal structure ψ [28])**.** *Let $\mathbb{T} = (T, \eta, \mu)$ be a monad, and τ and τ' be the unique left and right strengths on T as in Rem. 28. Define a family*

$$\psi = \bigl(\psi_{X,Y}\colon TX \times TY \to T(X \times Y)\bigr)_{X,Y \in \mathbf{Set}}$$

of morphisms as follows.

$$\psi_{X,Y}\colon TX \times TY \xrightarrow{\tau'_{X,TY}} T(X \times TY) \xrightarrow{T\tau_{X,Y}} TT(X \times Y) \xrightarrow{\mu_{X \times Y}} T(X \times Y)$$ ⌟

▶ **Proposition 30.** *Let $\mathbb{T} = (T, \eta, \mu)$ be a monad. Then the family ψ in Def. 29 together with $\eta_1\colon 1 \to T1$ makes the functor part T of $\mathbb{T}$ into a lax monoidal functor $(T, \eta_1, \psi)\colon (\mathbf{Set}, 1, \times) \to (\mathbf{Set}, 1, \times)$ [28, Thm. 2.1]. Moreover, the unit η and the left strength τ are monoidal natural transformations.* ⌟

▶ Remark 31. Note that in Prop. 30 we only claim that the *functor part* T of a monad $\mathbb{T}$ is lax monoidal. In order for the *monad* $\mathbb{T}$ to be lax monoidal (in the sense that it is a monad in the 2-category of monoidal categories, lax monoidal functors, and monoidal natural transformations), the multiplication μ has to be monoidal, which is equivalent to $\mathbb{T}$ being commutative in the ordinary sense. ⌟

4.2 W-Commutativity

Let $\mathbb{T}$ be a monad (on **Set**). Thanks to Prop. 30, for any $n \in \mathbb{N}$ and n-tuple $\vec{X} = (X_i)_{i \in n}$ of sets, we obtain a well-defined morphism

$$\psi_{\vec{X}}\colon \textstyle\prod_{i \in n} TX_i \to T\bigl(\prod_{i \in n} X_i\bigr) \tag{16}$$

by composing various components of ψ appropriately. (When $n = 0$, (16) is $\eta_1 : 1 \to T1$; when $n = 1$, (16) is the identity on TX_0.) In particular, this means that we construct the morphisms (16) so that the following holds. For any $n \in \mathbb{N}$, n-tuple $\vec{X} = (X_i)_{i \in n}$ of sets, and $0 \le k \le n$, let $\vec{X}' = (X_i)_{i \in k}$ and $\vec{X}'' = (X_{k+i})_{i \in n-k}$. Then

$$\begin{tikzcd}
\prod_{i\in k} TX_i \times \prod_{i\in n-k} TX_{k+i} \arrow[r, "\cong"] \arrow[d, "\psi_{\vec{X}'} \times \psi_{\vec{X}''}"'] & \prod_{i\in n} TX_i \arrow[dd, "\psi_{\vec{X}}"] \\
T\bigl(\prod_{i\in k} X_i\bigr) \times T\bigl(\prod_{i\in n-k} X_{k+i}\bigr) \arrow[d, "\psi_{X',X''}"'] & \\
T\bigl(\prod_{i\in k} X_i \times \prod_{i\in n-k} X_{k+i}\bigr) \arrow[r, "\cong"] & T\bigl(\prod_{i\in n} X_i\bigr)
\end{tikzcd} \tag{17}$$

commutes, where $X' = \prod_{i\in k} X_i$ and $X'' = \prod_{i\in n-k} X_{k+i}$. As in [11, § 3.2], we set

$$\psi_X^{(n)} : (TX)^n \to T(X^n) \tag{18}$$

as the special case of (16) with $X_0 = X_1 = \cdots = X_{n-1} = X$. Note in particular that $\psi_X^{(0)} = \eta_1 : 1 \to T1$ and $\psi_X^{(1)}$ is the identity on TX.

▶ **Definition 32** ($\mathbf{W}$**-commutative monad).** *Let* $\mathbf{W}$ *be a verbal category. A monad* $\mathbb{T} = (T, \eta, \mu)$ *is* $\mathbf{W}$-commutative *if, for each morphism* $\alpha : m \to n$ *in* $\mathbf{W}$ *and set* X*, the diagram*

$$\begin{tikzcd}
(TX)^n \arrow[r, "\psi_X^{(n)}"] \arrow[d, "(TX)^\alpha"'] & T(X^n) \arrow[d, "T(X^\alpha)"] \\
(TX)^m \arrow[r, "\psi_X^{(m)}"] & T(X^m)
\end{tikzcd} \tag{19}$$

commutes. ⌟

▶ Remark 33. It follows directly from Def. 32 that $\mathbf{W}$-commutativity of a monad $\mathbb{T}$ is down-closed with respect to $\mathbf{W}$: if $\mathbb{T}$ is $\mathbf{W}'$-commutative, then it is $\mathbf{W}$-commutative for any $\mathbf{W} \subseteq \mathbf{W}'$. We discussed this monotonicity in § 1 (cf. (3)). ⌟

▶ Remark 34. The functor part T of a monad $\mathbb{T}$ admits another canonical lax monoidal structure (η_1, ψ'), where ψ' is obtained by swapping the order of τ and τ' in the definition of ψ (Def. 29). We have $\psi = \psi'$ iff $\mathbb{T}$ is commutative in the ordinary sense. Using ψ' instead of ψ, we obtain the evident analogues of (16) and (18), and hence of Def. 32. The resulting class of $\mathbf{W}$-commutative monads coincides with the one defined here using ψ. ⌟

The following result establishes certain "robustness" of the notion of $\mathbf{W}$-commutative monad, by giving an alternative characterization.

▶ **Proposition 35.** *Let* $\mathbb{T} = (T, \eta, \mu)$ *be a monad and* $\alpha : m \to n$ *be a morphism in* $\mathbf{F}$*. Then the following are equivalent.*

(1) *For each* n*-tuple* $\vec{X} = (X_j)_{j\in n}$ *of sets, the following diagram commutes.*

$$\begin{tikzcd}
\prod_{j\in n} TX_j \arrow[r, "\psi_{\vec{X}}"] \arrow[d, "\langle \pi_{\alpha(i)} \rangle_{i\in m}"'] & T\bigl(\prod_{j\in n} X_j\bigr) \arrow[d, "T\langle \pi_{\alpha(i)} \rangle_{i\in m}"] \\
\prod_{i\in m} TX_{\alpha(i)} \arrow[r, "\psi_{\vec{X}\alpha}"] & T\bigl(\prod_{i\in m} X_{\alpha(i)}\bigr)
\end{tikzcd} \tag{20}$$

Here, $\vec{X}\alpha = (X_{\alpha(i)})_{i\in m}$ *and* $\pi_{\alpha(i)}$ *is the* $\alpha(i)$*-th projection from the product* $\prod_{j\in n} TX_j$ *or* $\prod_{j\in n} X_j$*.*

(2) *For each set* X*, the diagram* (19) *commutes.* ⌟

We note that (19) is closely related to *residual diagrams* of [11, 45]; see § 5.3. The diagram (20) also appears in [59, Def. 2].

As already mentioned, for some verbal categories $\mathbf{W}$, the $\mathbf{W}$-commutativity of a monad can be characterized by means of known conditions on monads, such as (ordinary) *commutativity* [28], *affinity* [29], *relevance* [20], n-*relevance* [46, 45], and *hyperaffinity* [23, 16] (see [14, Appendix B] for details about these conditions).

▶ **Proposition 36** (cf. [45]). **(1)** *Any monad is $\mathbf{F}_{\mathrm{id}}$-commutative.*
(2) *A monad is $\mathbf{F}_{\mathrm{m,inj}}$-commutative iff it is affine.*
(3) *A monad is $\mathbf{F}_{\mathrm{bij}}$-commutative iff it is commutative.*
(4) *A monad is $\mathbf{F}_{\mathrm{inj}}$-commutative iff it is commutative and affine.*
(5) *A monad is $\mathbf{F}_{\mathrm{surj}}$-commutative iff it is commutative and relevant.*
(6) *A monad is $\mathbf{F}$-commutative iff it is hyperaffine.*
(7) *For $n > 0$, a monad is $\mathbf{F}^{(n)}_{\mathrm{surj}}$-commutative (Ex. 9) iff it is commutative and n-relevant.* ⌟

4.3 Examples

Here are examples of $\mathbf{W}$-commutative monads. We describe them via the characterizations of Prop. 36, since examples for those characterizations have been studied well.

▶ **Example 37** (commutative monads).
- The M-writer monad $\mathbb{W}^{\mathsf{M}}$ (Ex. 3) is commutative iff the monoid M is commutative.
- The S-multiset monad $\mathbb{M}^{\mathsf{S}}$ (Ex. 4) is commutative iff the semiring S is commutative. In particular, the finite powerset monad $\mathbb{P}_{\mathrm{fin}}$ is commutative, but is not affine or relevant.
- The C-exception monad $\mathbb{E}^C$ (Ex. 1) is commutative iff $|C| \le 1$. ⌟

▶ **Example 38** (affine monads). The affine S-multiset monad $\mathbb{A}^{\mathsf{S}}$ (Ex. 5) is affine for any semiring S. In particular, the probability distribution monad $\mathbb{D}$ and the nonempty finite powerset monad $\mathbb{P}^+_{\mathrm{fin}}$ are affine (and commutative). $\mathbb{D}$ and $\mathbb{P}^+_{\mathrm{fin}}$ are not relevant. ⌟

▶ **Example 39** (relevant monads). The M-writer monad $\mathbb{W}^{\mathsf{M}}$ (Ex. 3) is relevant iff the monoid M is idempotent. The C-exception monad $\mathbb{E}^C$ (Ex. 1) is relevant for any set C. ⌟

▶ **Example 40** (hyperaffine monads). The C-reader monad $\mathbb{R}^C$ (Ex. 2) is hyperaffine. ⌟

5 Canonical Distributive Laws

Having introduced the notions of $\mathbf{W}$-operad (§ 3) and $\mathbf{W}$-commutative monad (§ 4) for each verbal category $\mathbf{W}$, we are now ready to present the construction of the canonical distributive laws. We first recall the general definition of distributive laws.

▶ **Definition 41** (distributive law [3]). *Let $\mathbb{S} = (S, \eta^{\mathbb{S}}, \mu^{\mathbb{S}})$ and $\mathbb{T} = (T, \eta^{\mathbb{T}}, \mu^{\mathbb{T}})$ be monads. A* distributive law of $\mathbb{S}$ over $\mathbb{T}$ *is a natural transformation $\delta\colon ST \to TS$ making the following four diagrams (stating the compatibility of δ with $\eta^{\mathbb{S}}$, $\mu^{\mathbb{S}}$, $\eta^{\mathbb{T}}$, and $\mu^{\mathbb{T}}$, respectively) commute.*

$$\begin{tikzcd} & T \arrow[dl, "\eta^{\mathbb{S}}T"'] \arrow[dr, "T\eta^{\mathbb{S}}"] & \\ ST \arrow[rr, "\delta"] & & TS \end{tikzcd} \quad \begin{tikzcd} SST \arrow[r, "S\delta"] \arrow[d, "\mu^{\mathbb{S}}T"'] & STS \arrow[r, "\delta S"] & TSS \arrow[d, "T\mu^{\mathbb{S}}"] \\ ST \arrow[rr, "\delta"] & & TS \end{tikzcd} \quad \begin{tikzcd} & S \arrow[dl, "S\eta^{\mathbb{T}}"'] \arrow[dr, "\eta^{\mathbb{T}}S"] & \\ ST \arrow[rr, "\delta"] & & TS \end{tikzcd} \quad \begin{tikzcd} STT \arrow[r, "\delta T"] \arrow[d, "S\mu^{\mathbb{T}}"'] & TST \arrow[r, "T\delta"] & TTS \arrow[d, "\mu^{\mathbb{T}}S"] \\ ST \arrow[rr, "\delta"] & & TS \end{tikzcd}$$

When δ satisfies the first two axioms above (compatibility of δ with $\eta^{\mathbb{S}}$ and $\mu^{\mathbb{S}}$) but not necessarily the last two, we say that (T, δ) is a monad functor from $\mathbb{S}$ to $\mathbb{S}$ *[52], or that δ is a* distributive law of the monad $\mathbb{S}$ over the endofunctor T *(as opposed to the* monad $\mathbb{T}$*).* ⌟

The following theorem presents our construction of distributive laws. Note the structure of the result: without additional assumptions, it yields only a distributive law δ of a monad over an *endofunctor* T. To obtain a genuine distributive law of a monad over a *monad* $\mathbb{T}$, one must furthermore impose either condition 1) or 2) in the second paragraph of Thm. 42.

▶ **Theorem 42.** *Let* $\mathbf{W}$ *be a verbal category,* $\mathbb{O} = (O, \mathsf{id}, \mathsf{subst})$ *be a* $\mathbf{W}$*-operad, and* $\mathbb{T} = (T, \eta^{\mathbb{T}}, \mu^{\mathbb{T}})$ *be a* $\mathbf{W}$*-commutative monad. Let* $\mathrm{Mnd}(\mathbb{O}) = (\mathrm{Lan}_J O, \eta^{\mathrm{Mnd}(\mathbb{O})}, \mu^{\mathrm{Mnd}(\mathbb{O})})$ *denote the* $\mathbf{W}$*-operadic monad induced by* $\mathbb{O}$ *(§ 3.3). Then the family*

$$\delta = \big(\delta_X \colon (\mathrm{Lan}_J O)TX \to T(\mathrm{Lan}_J O)X\big)_{X \in \mathbf{Set}}, \tag{21}$$

of functions defined below (in § 5.1) gives rise to a distributive law of the monad $\mathrm{Mnd}(\mathbb{O})$ *over the* endofunctor T.

Moreover, δ *is a distributive law of* $\mathrm{Mnd}(\mathbb{O})$ *over the* monad $\mathbb{T}$ *–* δ *is compatible with* $\mu^{\mathbb{T}}$*, in particular – whenever either of the following conditions is satisfied: 1)* $\mathbb{T}$ *is commutative, or 2)* $O_n = 0$ *for all* $n > 1$.

▶ Remark 43. When $\mathbf{W}$ is symmetric (see Def. 8), the $\mathbf{W}$-commutativity of a monad $\mathbb{T}$ implies the ordinary commutativity of $\mathbb{T}$ (Prop. 36(3)), and hence Cond. 1) of Thm. 42. In contrast, Cond. 2) of Thm. 42 can never be satisfied when the verbal category $\mathbf{W}$ contains $\mathbf{F}_{\mathrm{m,inj}}$. We use Cond. 2) of Thm. 42 only when $\mathbf{W} = \mathbf{F}_{\mathrm{id}}$ (see Exs 51 and 52).

To see the necessity of the additional conditions, consider the case where $\mathbf{W} = \mathbf{F}_{\mathrm{id}}$, $\mathbb{O}$ is the terminal $\mathbf{F}_{\mathrm{id}}$-operad (thus $\mathrm{Mnd}(\mathbb{O})$ is the list monad; see Ex. 20), and $\mathbb{T}$ is the list monad. By Prop. 36(1), $\mathbb{T}$ is $\mathbf{F}_{\mathrm{id}}$-commutative. However, it is known that there is no distributive law of the list monad over itself [66, Ex. 4.18]. ⌟

As announced in § 1, the additional condition 1) for δ's compatibility with $\mu^{\mathbb{T}}$ directs our main focus to verbal categories that admit exchange (E), in particular $\mathbf{W} = \mathbf{F}_{\mathrm{bij}}, \mathbf{F}_{\mathrm{inj}}, \mathbf{F}_{\mathrm{surj}}$, and $\mathbf{F}$ (see (3)). In fact, compatibility with $\eta^{\mathbb{T}}$ holds without such a condition; see below.

5.1 The Construction

Here we define δ in (21), under the situation of the first paragraph of Thm. 42. Fix a set X. Since the domain $(\mathrm{Lan}_J O)TX$ of δ_X is the coend $\int^{n \in \mathbf{W}} O_n \times (TX)^n$ (see (11)), to give the function δ_X is equivalent to giving a family

$$\big(\bar{\delta}_{X,n} \colon O_n \times (TX)^n \to T(\mathrm{Lan}_J O)X\big)_{n \in \mathbf{W}} \tag{22}$$

of functions extranatural in $n \in \mathbf{W}$ (see e.g. [37, § IX.6]). We define $\bar{\delta}_{X,n}$ as the composite of

$$O_n \times (TX)^n \xrightarrow{1 \times \psi_X^{(n)}} O_n \times T(X^n) \xrightarrow{\tau_{O_n, X^n}} T(O_n \times X^n) \xrightarrow{T\kappa_n} T(\mathrm{Lan}_J O)X,$$

where 1) $\psi_X^{(n)} \colon (TX)^n \to T(X^n)$ was defined in (18), 2) τ_{O_n, X^n} is the left strength (Def. 27) of T, and 3) $\kappa_n \colon O_n \times X^n \to (\mathrm{Lan}_J O)X$ is the n-th coprojection of the coend (11).

We have to show the extranaturality of (22), that is, that the following diagram commutes for any morphism $\alpha \colon m \to n$ in $\mathbf{W}$:

$$\begin{tikzcd} O_m \times (TX)^n \arrow[r, "O\alpha \times 1"] \arrow[d, "1 \times (TX)^\alpha"'] & O_n \times (TX)^n \arrow[d, "\bar{\delta}_{X,n}"] \\ O_m \times (TX)^m \arrow[r, "\bar{\delta}_{X,m}"] & T(\mathrm{Lan}_J O)X. \end{tikzcd} \tag{23}$$

This can be seen by considering the following diagram.

$$
\begin{tikzcd}
O_m \times (TX)^n \arrow[rrr, "O\alpha \times 1"] \arrow[dr, "1 \times \psi_X^{(n)}"] \arrow[ddd, "1 \times (TX)^\alpha"'] & & & O_n \times (TX)^n \arrow[d, "1 \times \psi_X^{(n)}"'] \\
\arrow[r, phantom, "(\mathbf{W}\text{-commutativity (19)})"] & O_m \times T(X^n) \arrow[rr, "O\alpha \times 1"] \arrow[dr, "\tau_{O_m, X^n}"] \arrow[dd, "1 \times T(X^\alpha)"'] \arrow[urr, phantom, "(\text{bifunctoriality of } \times)"] & \arrow[r, phantom, "(\text{naturality of } \tau)"] & O_n \times T(X^n) \arrow[d, "\tau_{O_n, X^n}"'] \\
& \arrow[r, phantom, "(\text{naturality of } \tau)"] & T(O_m \times X^n) \arrow[r, "T(O\alpha \times 1)"] \arrow[d, "T(1 \times X^\alpha)"'] \arrow[dr, phantom, "(\text{extranaturality of } \kappa)"] & T(O_n \times X^n) \arrow[d, "T\kappa_n"'] \\
O_m \times (TX)^m \arrow[r, "1 \times \psi_X^{(m)}"] & O_m \times T(X^m) \arrow[r, "\tau_{O_m, X^m}"] & T(O_m \times X^m) \arrow[r, "T\kappa_m"] & T(\mathrm{Lan}_J O)X
\end{tikzcd}
$$

The defining condition (19) of $\mathbf{W}$-commutativity is crucial here.

See [14, Appendix D.6] for the rest of the proof of Thm. 42. We note that the compatibility of δ with $\eta^{\mathrm{Mnd}(\mathbb{O})}$, $\mu^{\mathrm{Mnd}(\mathbb{O})}$, and $\eta^{\mathbb{T}}$ holds without Conditions 1) or 2). It is the compatibility of δ with $\mu^{\mathbb{T}}$ where the additional assumption is needed.

Some examples are in § 5.2. Other known distributive laws can be seen to be *not* instances of Thm. 42, see e.g. [47, § 9].

▶ Remark 44. Let $\mathbb{S}$ be a $\mathbf{W}$-operadic monad which is *not* uniquely $\mathbf{W}$-operadic (cf. Prop. 17). Given a $\mathbf{W}$-commutative monad $\mathbb{T}$, the construction of the natural transformation $\delta\colon ST \to TS$ presented here refers to the choice of a $\mathbf{W}$-operad $\mathbb{O}$ inducing $\mathbb{S}$. Hence it is *a priori* possible that δ varies with this choice. At present, we are not aware of such an example. ⌟

In contrast, we can show that the choice of $\mathbf{W}$ does not matter, as we now explain. Let $\mathbf{W}$ and $\mathbf{W}'$ be verbal categories such that $\mathbf{W} \subseteq \mathbf{W}'$, $\mathbb{O}$ be a $\mathbf{W}$-operad, and $\mathbb{T}$ be a $\mathbf{W}'$-commutative monad. Then $\mathbb{T}$ is $\mathbf{W}$-commutative as well (Rem. 33). On the other hand, we saw in Prop. 19(1) that $\mathbb{O}$ induces a $\mathbf{W}'$-operad $\mathbb{O}'$ by extension, and that $\mathbb{O}$ and $\mathbb{O}'$ induce isomorphic monads, which we identify and write as $\mathbb{S}$. Thus we can apply the construction at the beginning of § 5.1 to $(\mathbf{W}, \mathbb{O}, \mathbb{T})$ as well as to $(\mathbf{W}', \mathbb{O}', \mathbb{T})$, obtaining natural transformations δ and δ' respectively, both of type $ST \to TS$.

▶ **Proposition 45.** *In the above situation, we have $\delta = \delta'$.* ⌟

5.2 Examples

Here is a concrete description – in algebraic terms, see e.g. [47, 49] – of the general construction of $\delta\colon ST \to TS$ in § 5.1. We give it by an example. Consider an $\mathbb{S}$-term over $\mathbb{T}$-terms over X

$$p\big(\, t_0(v, w),\; t_1(x, y, z) \,\big) \in STX, \tag{24}$$

where p is a binary $\mathbb{S}$-term, t_0 and t_1 are $\mathbb{T}$-terms, and $v, w, x, y, z \in X$. It is carried by δ_X to the following $\mathbb{T}$-term over $\mathbb{S}$-terms over X:

$$t_0\Big(\; t_1\big(\, p(v, x),\, p(v, y),\, p(v, z) \,\big),\;\; t_1\big(\, p(w, x),\, p(w, y),\, p(w, z)\big) \;\Big) \in TSX. \tag{25}$$

The term t_0 appears outside t_1 because, in the definition of the lax monoidal structure ψ (Def. 29), the left occurrence of T is pulled out first. When $\mathbb{T}$ is commutative – as in most cases of interest – swapping t_0 and t_1 in (25) (with the corresponding rearrangement of arguments) yields the same element of TSX.

▶ **Example 46.** Let $\mathbf{W} = \mathbf{F}_{\mathrm{bij}}$, $\mathbb{S}$ be the list monad $(-)^*$ (whose operations we write multiplicatively), and $\mathbb{T}$ be the abelian group monad $\mathbb{M}^{\mathbb{Z}}$ (written additively). Letting $p(x, y) = xy$, $t_0(v, w) = v + 2w$, and $t_1(x, y, z) = x + 3y + z$, (24) becomes $(v + 2w)(x + 3y + z)$

and (25) becomes $vx + 3vy + vz + 2wx + 6wy + 2wz$. Thus Thm. 42 gives rise to the "archetypal" distributive law δ expressing the "distribution of multiplication over addition". As is well-known, δ makes the composite TS into the ring monad [3, § 4(1)]. ⌟

▶ **Example 47.** Let $\mathbf{W} = \mathbf{F}_{\mathrm{inj}}$, $\mathbb{S}$ be the multiset monad $\mathbb{M}$ (Ex. 23), and $\mathbb{T}$ be the probability distribution monad $\mathbb{D}$ (Ex. 38). Thm. 42 induces the distributive law studied in [22, 32]. ⌟

▶ **Example 48.** Let $\mathbf{W} = \mathbf{F}_{\mathrm{surj}}$, $\mathbb{S}$ be either of the semigroup, commutative semigroup, or nonunital semilattice monad, and $\mathbb{T}$ be the maybe monad $1 + (-)$ (Exs 37 and 39). Thm. 42 induces the distributive law $S(1 + X) \to 1 + SX$ yielding the monoid, commutative monoid, or (unital) semilattice monad as the composite monad. ⌟

▶ **Example 49.** Let $\mathbf{W} = \mathbf{F}$, $\mathbb{S}$ be any finitary monad on **Set**, and $\mathbb{T}$ be the C-reader monad $\mathbb{R}^C$ (Ex. 40). Thm. 42 induces the canonical distributive law $S(X^C) \to (SX)^C$. A related distributive law appears in [12, Prop. 40]. In fact, this construction works for any (not necessarily finitary) monad $\mathbb{S}$ on **Set** [5, Prop. 2.2]. We note that our construction with $\mathbf{W} = \mathbf{F}$ coincides with [5, Cor. 4.3(c)], which is a special case of [5, Thm. 4.2]. ⌟

▶ **Example 50.** Let $\mathbf{W} = \mathbf{F}_{\mathrm{inj}}$, $\mathbb{S}$ be the *indexed valuation monad* $\mathbb{IV}$ of [62] (which we show is indeed $\mathbf{F}_{\mathrm{inj}}$-operadic in § 6.3), and $\mathbb{T}$ be the nonempty finite powerset monad $\mathbb{P}^{+}_{\mathrm{fin}}$. Thm. 42 induces the distributive law $IVP^{+}_{\mathrm{fin}} \to P^{+}_{\mathrm{fin}}IV$ constructed in [62] to model combination of probabilistic and nondeterministic choice.

Letting $\mathbb{T}$ be the probability distribution monad $\mathbb{D}$ instead, we obtain another distributive law $\delta\colon IVD \to DIV$. This δ turns out to be closely related to a family $\lambda = (\lambda_X : DDX \to DDX)_{X\in\mathbf{Set}}$ of morphisms introduced in [61] for a new semantics of MDPs. The family λ does not form a distributive law of $\mathbb{D}$ over itself since it is not natural; indeed, no such distributive law exists [66].

In summary, we observed that the general construction of Thm. 42 yields a distributive law $\delta\colon IVD \to DIV$, which resolves the categorical ill-behavedness of $\lambda\colon DD \to DD$ used in [61]. Its consequence in MDP semantics is an exciting direction of future work. The benefit of such "indexing" in program semantics is also observed in [35]. ⌟

Exs 51 and 52 below exploit Cond. 2 of Thm. 42.

▶ **Example 51.** Let $\mathbf{W} = \mathbf{F}_{\mathrm{id}}$, $\mathbb{S}$ be the C-exception monad $\mathbb{E}^C$ (Ex. 20), and $\mathbb{T}$ be an arbitrary monad on **Set**. Thm. 42 yields the distributive law $C + TX \to T(C + X)$ induced by the unit of $\mathbb{T}$ [3, § 4(2)]. ⌟

▶ **Example 52.** Let $\mathbf{W} = \mathbf{F}_{\mathrm{id}}$, $\mathbb{S}$ be the M-writer monad $\mathbb{W}^{\mathsf{M}}$ (Ex. 20), and $\mathbb{T}$ be an arbitrary monad on **Set**. Thm. 42 yields the distributive law $M \times TX \to T(M \times X)$ given by the left strength τ of T [3, § 4(3)]. ⌟

The distributive laws of Exs 51 and 52 (or more precisely, their generalizations to base categories other than **Set**) are used in [18, Cor. 3 and Thm. 12] to model combination of computational effects with exception and output.

5.3 Eilenberg–Moore Liftings and Kleisli Extensions

Let $\mathbb{S}$ and $\mathbb{T}$ be monads. A distributive law $\delta\colon ST \to TS$ is known to correspond to each of the following data (see e.g. [15, Prop. 6]).

Eilenberg–Moore lifting: A *lifting* $\overline{\mathbb{T}}$ of the monad $\mathbb{T}$ to the Eilenberg–Moore category $\mathbf{Set}^{\mathbb{S}}$ of $\mathbb{S}$, as on the left of (26).

Kleisli extension: An *extension* $\underline{\mathbb{S}}$ of the monad $\mathbb{S}$ to the Kleisli category $\mathbf{Set}_{\mathbb{T}}$ of $\mathbb{T}$, as on the right of (26).

$$\begin{array}{ccc} \mathbf{Set}^{\mathbb{S}} & \xrightarrow{\overline{\mathbb{T}}} & \mathbf{Set}^{\mathbb{S}} \\ {\scriptstyle U^{\mathbb{S}}}\downarrow & & \downarrow{\scriptstyle U^{\mathbb{S}}} \\ \mathbf{Set} & \xrightarrow{\mathbb{T}} & \mathbf{Set} \end{array} \qquad\qquad \begin{array}{ccc} \mathbf{Set} & \xrightarrow{\mathbb{S}} & \mathbf{Set} \\ {\scriptstyle F_{\mathbb{T}}}\downarrow & & \downarrow{\scriptstyle F_{\mathbb{T}}} \\ \mathbf{Set}_{\mathbb{T}} & \xrightarrow{\underline{\mathbb{S}}} & \mathbf{Set}_{\mathbb{T}} \end{array} \tag{26}$$

Here we comment on these aspects.

There are a few works [38, 45, 11] that pursue goals close to ours using the above Eilenberg–Moore characterization. In the latter, since the functor part T of any monad $\mathbb{T} = (T, \eta, \mu)$ on $\mathbf{Set}$ has the lax monoidal structure (η_1, ψ) (see Prop. 30), any n-ary operation $f\colon X^n \to X$ on a set X induces the n-ary operation

$$(TX)^n \xrightarrow{\psi_X^{(n)}} T(X^n) \xrightarrow{Tf} TX$$

on TX. Thus the key question towards the existence of $\overline{\mathbb{T}}$ is the following: does T preserve *equations* used in an equational presentation of the algebraic theory corresponding to $\mathbb{S}$?

In this context, it has been noted that any commutative monad preserves *linear* equations (expressible by $\mathbf{F}_{\mathrm{bij}}$-operads) [38, Lem. 4.3.3], any commutative and affine monad preserves *non-dup* equations (expressible by $\mathbf{F}_{\mathrm{inj}}$-operads) [11, Thm. 9], and any commutative and relevant monad preserves *non-drop* equations (expressible by $\mathbf{F}_{\mathrm{surj}}$-operads) [11, Thm. 8].

These results in [11] use what they call *residual diagrams* ([11, § 4.1]), which look close to our (19). The notion of $\mathbf{W}$-commutative monad (Def. 32) unifies various residual diagrams in our systematic, operadic, and presentation-independent framework parametrized in a verbal category $\mathbf{W}$. Residual diagrams are used in [45] for concrete applications of a general and abstract lifting result [45, Thm. 3.14] (which formalizes our algebraic arguments above).

An interpretation of our results in the second, Kleisli extension characterization, is as follows. From now on, let $\mathbf{W}$ be a symmetric verbal category, $\mathbb{O}$ be a $\mathbf{W}$-operad, and $\mathbb{T}$ be a $\mathbf{W}$-commutative monad. In this situation, the canonical symmetric monoidal structure of the Kleisli category $\mathbf{Set}_{\mathbb{T}}$ has enough structure to accommodate the notion of $\mathbb{O}$-*algebra in* $\mathbf{Set}_{\mathbb{T}}$. These $\mathbb{O}$-algebras then correspond to Eilenberg–Moore algebras of the extended monad $\underline{\mathrm{Mnd}(\mathbb{O})}$ on $\mathbf{Set}_{\mathbb{T}}$, corresponding to our δ in Thm. 42.

We sketch a definition of $\mathbb{O}$-algebras in $\mathbf{Set}_{\mathbb{T}}$, independently from the extended monad $\underline{\mathrm{Mnd}(\mathbb{O})}$. First we need:

▶ **Proposition 53.** *Every set X induces a $\mathbf{W}$-operad $\mathbb{E}\mathbf{nd}_{\mathbb{T}}(X) = (\mathrm{End}_{\mathbb{T}}(X), \mathsf{id}, \mathsf{subst})$ with*
- $\mathrm{End}_{\mathbb{T}}(X)_n = \mathbf{Set}_{\mathbb{T}}(X^n, X) = \mathbf{Set}(X^n, TX)$*;*
- $\mathsf{id} = \eta_X \in \mathbf{Set}(X, TX)$*; and*
- *given $g \in \mathbf{Set}(X^n, TX)$ and $\vec{f} = \big(f_i \in \mathbf{Set}(X^{m_i}, TX)\big)_{i\in n}$, $\mathsf{subst}(g, \vec{f}) \in \mathbf{Set}(X^m, TX)$ (with $m = \sum_{i\in n} m_i$) is the following composite:*

$$X^m \cong \textstyle\prod_{i\in n} X^{m_i} \xrightarrow{\prod_{i\in n} f_i} (TX)^n \xrightarrow{\psi_X^{(n)}} T(X^n) \xrightarrow{Tg} TTX \xrightarrow{\mu_X} TX. \qquad ⌟$$

We note that the commutativity of $\mathbb{T}$ is used to show the associativity axiom of $\mathbb{E}\mathbf{nd}_{\mathbb{T}}(X)$, whereas the $\mathbf{W}$-commutativity of $\mathbb{T}$ is used to show the compatibility axiom (6) for subst. As usual, we define the notion of $\mathbb{O}$-algebra via operad homomorphisms to endomorphism operads (see e.g. [33, § 6.4]).

▶ **Definition 54.** *An* $\mathbb{O}$-algebra in $\mathbf{Set}_{\mathbb{T}}$ *is a pair* (X, ξ)*, where* X *is a set and* ξ *is a homomorphism of* $\mathbf{W}$*-operads from* $\mathbb{O}$ *to* $\mathbb{E}\mathbf{nd}_{\mathbb{T}}(X)$ *(meaning a natural transformation* $O \to \mathrm{End}_{\mathbb{T}}(X)$ *compatible with* id *and* subst*). Similarly, define the notion of* homomorphism of $\mathbb{O}$-algebras $(X, \xi) \to (Y, \upsilon)$ *as a morphism* $X \to Y$ *in* $\mathbf{Set}_{\mathbb{T}}$ *suitably compatible with the algebra structures* ξ *and* υ*.* ⌟

▶ **Proposition 55.** *The category of* $\mathbb{O}$*-algebras in* $\mathbf{Set}_{\mathbb{T}}$ *and their homomorphisms is isomorphic to the Eilenberg–Moore category of the monad* $\underline{\mathrm{Mnd}(\mathbb{O})}$ *extended to the Kleisli category* $\mathbf{Set}_{\mathbb{T}}$ *via the distributive law* δ *of Thm. 42.* ⌟

6 W-Operadic Refinement of a Monad

Let $\mathbf{W}$ be a verbal category. In § 3.3, we constructed a monad $\mathrm{Mnd}(\mathbb{O})$ from any $\mathbf{W}$-operad $\mathbb{O}$. In this section, we first construct a $\mathbf{W}$-operad $\mathbf{W}\text{-}\mathrm{Opd}(\mathbb{S})$ from any monad $\mathbb{S}$ (§ 6.1; cf. Rem. 18). This provides us with a universal way of turning any monad $\mathbb{S}$ into a $\mathbf{W}$-operadic one, namely $\mathrm{Mnd}(\mathbf{W}\text{-}\mathrm{Opd}(\mathbb{S}))$, which we call the $\mathbf{W}$*-operadic refinement* of $\mathbb{S}$. As explained in § 1, this construction is useful when an obstacle to the existence of a distributive law $\delta \colon ST \to TS$ of a monad $\mathbb{S}$ over a $\mathbf{W}$-commutative monad $\mathbb{T}$ is the lack of $\mathbf{W}$-operadicity of $\mathbb{S}$. We illustrate its use in §§ 6.3 and 7.

6.1 From Monads to W-Operads

Let $\mathbb{S} = (S, \eta, \mu)$ be a monad (on $\mathbf{Set}$). The $\mathbf{W}$-operad $\mathbf{W}\text{-}\mathrm{Opd}(\mathbb{S})$ induced by $\mathbb{S}$ is defined as follows. The underlying functor of $\mathbf{W}\text{-}\mathrm{Opd}(\mathbb{S})$ is the composite $\mathbf{W} \xrightarrow{J} \mathbf{Set} \xrightarrow{S} \mathbf{Set}$. The element $\mathsf{id} \in SJ1 = S1$ is the one corresponding to $\eta_1 \colon 1 \to S1$. Given $n \in \mathbb{N}$ and $\vec{m} = (m_i \in \mathbb{N})_{i \in n}$, the function

$$\mathsf{subst}_{n,\vec{m}} \colon Sn \times \textstyle\prod_{i \in n} Sm_i \to Sm,$$

where $m = \sum_{i \in n} m_i$, is defined as follows. For each $q \in Sn$ and $\vec{p} = (p_i \in Sm_i)_{i \in n}$, we set $\mathsf{subst}_{n,\vec{m}}(q, \vec{p}) \in Sm$ as the element corresponding to the composite

$$1 \xrightarrow{q} Sn \xrightarrow{S\bar{p}} SSm \xrightarrow{\mu_m} Sm,$$

where $\bar{p} \colon n \to Sm$ maps each $i \in n$ to $(S\iota_{\vec{m},i})(p_i)$ ($\iota_{\vec{m},i} \colon m_i \to m$ is from Def. 6). It is straightforward to check that this defines a $\mathbf{W}$-operad $\mathbf{W}\text{-}\mathrm{Opd}(\mathbb{S})$.

▶ **Example 56.** Let S be a semiring. For the S-multiset monad $\mathbb{M}^{\mathsf{S}}$ (Ex. 4), we have:

- $\mathbf{F}_{\mathrm{bij}}\text{-}\mathrm{Opd}(\mathbb{M}^{\mathsf{S}})$ is the $\mathbf{F}_{\mathrm{bij}}$-operad $\mathbb{O}^{\mathsf{S}_{\mathrm{mult}}}$ (see Ex. 23) induced by the multiplicative monoid $\mathsf{S}_{\mathrm{mult}}$ of S, and
- $\mathbf{F}_{\mathrm{inj}}\text{-}\mathrm{Opd}(\mathbb{M}^{\mathsf{S}})$ is the $\mathbf{F}_{\mathrm{inj}}$-operad $\mathbb{O}^{\mathsf{S}_{\mathrm{mult},0}}$ (see Ex. 24) induced by the multiplicative monoid with zero $\mathsf{S}_{\mathrm{mult},0}$ of S. ⌟

▶ **Proposition 57** (cf. Rem. 18)**.** *Let* $\mathbf{W}$ *be a verbal category. The constructions* $\mathrm{Mnd}(-)$ *(from § 3.3) and* $\mathbf{W}\text{-}\mathrm{Opd}(-)$ *form an adjunction*

$$\begin{tikzcd} \mathbf{Mnd}(\mathbf{Set}) \arrow[r, shift right, "\mathbf{W}\text{-}\mathrm{Opd}(-)"'] \arrow[r, phantom, "\perp"] & \mathbf{W}\text{-}\mathbf{Opd} \arrow[l, shift right, "\mathrm{Mnd}(-)"'] \end{tikzcd} \tag{27}$$

between the categories $\mathbf{Mnd}(\mathbf{Set})$ *of monads on* $\mathbf{Set}$ *and* $\mathbf{W}\text{-}\mathbf{Opd}$ *of* $\mathbf{W}$*-operads.* ⌟

6.2 Turning Any Monad Into a $\mathbb{W}$-Operadic Monad

▶ **Definition 58** ($\mathbf{W}$-operadic refinement $\mathrm{Rf}_{\mathbf{W}}(\mathbb{S})$)**.** *Let* $\mathbf{W}$ *be a verbal category and* $\mathbb{S}$ *be a monad. The (*$\mathbf{W}$*-operadic) monad* $\mathrm{Mnd}(\mathbf{W}\text{-}\mathrm{Opd}(\mathbb{S}))$ *is called the* $\mathbf{W}$-operadic refinement *of* $\mathbb{S}$ *and is denoted by* $\mathrm{Rf}_{\mathbf{W}}(\mathbb{S})$. ⌟

In view of Prop. 57, $\mathrm{Rf}_{\mathbf{W}}(-)$ is the comonad on $\mathbf{Mnd}(\mathbf{Set})$ generated by the adjunction (27). Hence for each $\mathbb{S} \in \mathbf{Mnd}(\mathbf{Set})$, there exists a canonical monad morphism $\varepsilon_{\mathbb{S}} \colon \mathrm{Rf}_{\mathbf{W}}(\mathbb{S}) \to \mathbb{S}$ (the counit of (27) at $\mathbb{S}$) with the suitable universal property. In this sense, $\mathrm{Rf}_{\mathbf{W}}(\mathbb{S})$ is the *universal* $\mathbf{W}$-operadic monad induced by $\mathbb{S}$.

The following is an easy consequence of Exs 23, 24, and 56.

▶ **Proposition 59.** *Let* S *be a semiring.*

(1) *The* $\mathbf{F}_{\mathrm{bij}}$*-operadic refinement* $\mathrm{Rf}_{\mathbf{F}_{\mathrm{bij}}}(\mathbb{M}^{\mathsf{S}})$ *of the* S*-multiset monad* $\mathbb{M}^{\mathsf{S}}$ *(Ex. 4) is the* $\mathsf{N}[\mathsf{S}_{\mathrm{mult}}]$*-multiset monad* $\mathbb{M}^{\mathsf{N}[\mathsf{S}_{\mathrm{mult}}]}$.

(2) *The* $\mathbf{F}_{\mathrm{inj}}$*-operadic refinement* $\mathrm{Rf}_{\mathbf{F}_{\mathrm{inj}}}(\mathbb{M}^{\mathsf{S}})$ *of the* S*-multiset monad* $\mathbb{M}^{\mathsf{S}}$ *is the* $\mathsf{N}_0[\mathsf{S}_{\mathrm{mult},0}]$*-multiset monad* $\mathbb{M}^{\mathsf{N}_0[\mathsf{S}_{\mathrm{mult},0}]}$*. In particular, the* $\mathbf{F}_{\mathrm{inj}}$*-operadic refinement* $\mathrm{Rf}_{\mathbf{F}_{\mathrm{inj}}}(\mathbb{P}_{\mathrm{fin}})$ *of the finite powerset monad* $\mathbb{P}_{\mathrm{fin}}$ *is the multiset monad* $\mathbb{M}$. ⌟

For $\mathbf{W} = \mathbf{F}_{\mathrm{surj}}$ and $\mathbf{F}_{\mathrm{bij}}$, $\mathbf{W}$-operadic refinements have the following concrete presentations.

▶ **Proposition 60** (cf. Prop. 16(2) and (3))**.** *Let* $\mathbb{S}$ *be a monad.*

(1) *The functor part of its* $\mathbf{F}_{\mathrm{surj}}$*-operadic refinement* $\mathrm{Rf}_{\mathbf{F}_{\mathrm{surj}}}(\mathbb{S})$ *maps* X *to* $\sum_{A \in P_{\mathrm{fin}} X} SA$.

(2) *The functor part of its* $\mathbf{F}_{\mathrm{bij}}$*-operadic refinement* $\mathrm{Rf}_{\mathbf{F}_{\mathrm{bij}}}(\mathbb{S})$ *maps* X *to* $\left\{ (n, [p, \vec{x}]_{\sim_n}) \,\middle|\, n \in \mathbb{N}, p \in Sn, \vec{x} \in X^n \right\}$*, where* $\sim_n$ *is given by the natural action of bijections* $\alpha \colon n \xrightarrow{\cong} n$. ⌟

The presentation in Prop. 60(2) (for $\mathbf{W} = \mathbf{F}_{\mathrm{bij}}$) is almost the same as the general presentation

$$\left\{ (n, p, \vec{x}) \,\middle|\, n \in \mathbb{N}, p \in Sn, \vec{x} \in X^n \right\}/\!\sim, \qquad \text{where } \sim \text{ is induced by arrows in } \mathbf{W}, \tag{28}$$

of $\mathrm{Rf}_{\mathbf{W}}(\mathbb{S}) = \mathrm{Mnd}(\mathbf{W}\text{-}\mathrm{Opd}(\mathbb{S})) = \int^{n \in \mathbf{W}} Sn \times X^n$ for general $\mathbf{W}$ (cf. § 3.3 & Def. 58). The difference is that the equivalence $\sim_n$ in Prop. 60(2) is specific to each n, while $\sim$ in (28) can relate $(n, p, \vec{x}) \sim (n', p', \vec{x'})$ with $n \neq n'$. An operational intuition is discussed later in Ex. 66.

6.3 The Indexed Valuation Monad Revisited

As mentioned in § 1, in order to model the combination of probability and nondeterminism, Varacca and Winskel [62] consider the existence of a distributive law $VP^+_{\mathrm{fin}} \to P^+_{\mathrm{fin}}V$ of the valuation monad $\mathbb{V}$ (Ex. 4(3)) over the nonempty finite powerset monad $\mathbb{P}^+_{\mathrm{fin}}$ (Ex. 5(1)). Such a distributive law does not exist, but they obtain a distributive law $IVP^+_{\mathrm{fin}} \to P^+_{\mathrm{fin}}IV$ of the *indexed valuation monad* $\mathbb{IV}$ over $\mathbb{P}^+_{\mathrm{fin}}$. They construct $\mathbb{IV}$ by removing “problematic” equational axioms from an equational presentation of $\mathbb{V}$.

Since $\mathbb{P}^+_{\mathrm{fin}}$ is $\mathbf{F}_{\mathrm{inj}}$-commutative (Ex. 38), from our point of view, a natural approach is to replace $\mathbb{V}$ with its $\mathbf{F}_{\mathrm{inj}}$-refinement $\mathrm{Rf}_{\mathbf{F}_{\mathrm{inj}}}(\mathbb{V})$.

In this case, these two approaches lead to the same result:

▶ **Proposition 61.**

(1) *The* $\mathbf{F}_{\mathrm{inj}}$*-operadic refinement* $\mathrm{Rf}_{\mathbf{F}_{\mathrm{inj}}}(\mathbb{V})$ *of the valuation monad* $\mathbb{V}$ *is the indexed valuation monad* $\mathbb{IV}$.

(2) *The distributive law of* $\mathbb{IV}$ *over* $\mathbb{P}^+_{\mathrm{fin}}$ *obtained in [62, § 4.3] coincides with the one obtained from Thm. 42.* ⌟

The key to the proof of Prop. 61(1) is to observe that the set $IV(X)$ of finite indexed valuations [62, Def. 4.3] can be written as the coend $IV(X) = \int^{n\in\mathbf{F}_{\mathrm{inj}}} M^{\mathsf{R}_{\geq 0}} n \times X^n$. Props 61(1) and 59(2) yield the following novel characterization of $\mathbb{IV}$.

▶ **Corollary 62.** *The indexed valuation monad $\mathbb{IV}$ is isomorphic to the $\mathsf{N}_0[(\mathsf{R}_{\geq 0})_{\mathrm{mult},0}]$-multiset monad $\mathbb{M}^{\mathsf{N}_0[(\mathsf{R}_{\geq 0})_{\mathrm{mult},0}]}$.* ⌟

▶ Remark 63. In general, simply removing "problematic" axioms from an equational presentation of an algebraic theory does not give the same result as the **W**-operadic refinement. (An easy way to see this is to notice that the **W**-operadic refinement construction is not idempotent in general.) In fact, the former construction is presentation-dependent, whereas our **W**-operadic refinement construction is a universal construction independent of presentations (see Prop. 57). We leave it for future work to see whether the similar modification of the theory done in [11, § 5] can be captured by a suitable **W**-operadic refinement. ⌟

7 Application: Distributive Laws via Operadic Refinements

We have discussed examples of **W**-operadic monads (§ 3.5), **W**-commutative monads (§ 4.3), and canonical distributive laws between them (§ 5.2). Those distributive laws, in particular, can be seen as application of our theory. In this section, we discuss application of our categorical framework as a whole.

We follow the storyline sketched in § 1. Recall that

- we are mainly interested in four verbal categories $\mathbf{W} = \mathbf{F}_{\mathrm{bij}}, \mathbf{F}_{\mathrm{inj}}, \mathbf{F}_{\mathrm{surj}}$, and $\mathbf{F}$, with the inclusion order shown in (3);
- and that, to get the canonical distributive law (Thm. 42) to work, we need a nonempty intersection between 1) the upper set of $\mathbb{S}$'s **W**-operadicity and 2) the lower set of $\mathbb{T}$'s **W**-commutativity.

The empty intersection of the two sets is a sign that a distributive law may not exist (although a non-canonical distributive law may exist, see e.g. [47, § 9]).

▶ **Example 64.**

(1) Let $\mathbb{S} = \mathbb{D}$ (the probability distribution monad) and $\mathbb{T} = \mathbb{P}_{\mathrm{fin}}$ (the finite powerset monad). Then $\mathbb{S} = \mathbb{D}$ is not $\mathbf{F}_{\mathrm{bij}}$-operadic (by Rem. 15(2)). In contrast, the lower set of $\mathbb{T}$'s commutativity is seen to be $\{\mathbf{F}_{\mathrm{bij}}\}$ by Prop. 36. Thus this is an "empty intersection" situation.
Indeed, this is the combination for which Plotkin's counterexample shows there is no distributive law [62].

(2) Similarly, one can show that $(\mathbb{S}, \mathbb{T}) = (\mathbb{P}_{\mathrm{fin}}, \mathbb{P}_{\mathrm{fin}})$ is an "empty intersection" situation. ⌟

In such situations with the empty intersection, as discussed in § 1, we consider using operadic refinement in § 6 to modify $\mathbb{S}$.

▶ **Example 65.** For the above pairs $(\mathbb{S}, \mathbb{T}) = (\mathbb{D}, \mathbb{P}_{\mathrm{fin}})$, $(\mathbb{P}_{\mathrm{fin}}, \mathbb{P}_{\mathrm{fin}})$, we replace $\mathbb{S}$ with the $\mathbf{F}_{\mathrm{bij}}$-refinement $\mathrm{Rf}_{\mathbf{F}_{\mathrm{bij}}}(\mathbb{S})$; the refinements are intensional and indexed versions of $\mathbb{S}$ (cf. Prop. 60(2)). We obtain a canonical distributive law $\delta\colon \big(\mathrm{Rf}_{\mathbf{F}_{\mathrm{bij}}}(\mathbb{S})\big)T \to T\big(\mathrm{Rf}_{\mathbf{F}_{\mathrm{bij}}}(\mathbb{S})\big)$; Thm. 42 guarantees that δ is well-defined and satisfies the required axioms. ⌟

▶ **Example 66.** Let $\mathbb{S} = \mathbb{P}_{\mathrm{fin}}$ and $\mathbb{T} = \mathbb{W}^{\mathsf{M}}$ (the writer monad) with commutative M. Then $\mathbb{S}$ is not $\mathbf{F}_{\mathrm{bij}}$-operadic while $\mathbb{T}$ is only $\mathbf{F}_{\mathrm{bij}}$-commutative in general (§ 4.3). Therefore we replace $\mathbb{S}$ with $\mathrm{Rf}_{\mathbf{F}_{\mathrm{bij}}}(\mathbb{S})$, and obtain the following canonical distributive law.

$$\begin{array}{ccc} \mathrm{Rf}_{\mathbf{F}_{\mathrm{bij}}}(\mathbb{P}_{\mathrm{fin}})(W^{\mathsf{M}}X) & \xrightarrow{\delta_X} & W^{\mathsf{M}}(\mathrm{Rf}_{\mathbf{F}_{\mathrm{bij}}}(\mathbb{P}_{\mathrm{fin}})X) \\ \big(n,\ [p, (m_i, x_i)_{i\in n}]_{\sim_n}\big) & \longmapsto & \big(\prod_{i\in n} m_i,\ \big(n,\ [p, (x_i)_{i\in n}]_{\sim_n}\big)\big) \end{array} \tag{29}$$

Here we used the presentation of $\mathrm{Rf}_{\mathbf{F}_{\mathrm{bij}}}(\mathbb{P}_{\mathrm{fin}})$ in Prop. 60(2), with $n \in \mathbb{N}$, $p \in P_{\mathrm{fin}}n$, $m_i \in M$ and $x_i \in X$. The operational intuition here is that 1) n processes are spawned, 2) Process i (for each $i \in n$) produces output x_i, writing m_i as side effect, and 3) this m_i is used, in the combined effect $\prod_{i \in n} m_i$, no matter if Process i's output x_i is used in the $\mathbb{P}_{\mathrm{fin}}$-effect context $p \in P_{\mathrm{fin}}n$. See [14, Appendix C] for further discussions. ⌟

8 Conclusions and Future Work

We developed a general theory of monads and distributive laws that is parametrized by a verbal category $\mathbf{W}$, a notion for substructural contexts by Tronin [58]. The theory is categorical, unifying previous observations e.g. in [45, 11] in a presentation-independent manner. We constructed a canonical distributive law $ST \to TS$ that stands in the balance between $\mathbf{W}$-operadicity of $\mathbb{S}$ and $\mathbf{W}$-commutativity of $\mathbb{T}$.

Extension to base categories other than **Set** is an obvious future direction. Accommodating "modify $\mathbb{T}$" approaches [20, 6, 36] is another. Application to program logics [35, 39] and MDP model checking [61] is a practically important direction.

References

1. Quentin Aristote. Monotone weak distributive laws over the lifted powerset monad in categories of algebras. In *42nd International Symposium on Theoretical Aspects of Computer Science (STACS 2025)*, volume 327 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2025. `doi:10.4230/LIPIcs.STACS.2025.10`.
2. Falk Bartels. *On Generalised Coinduction and Probabilistic Specification Formats: Distributive laws in coalgebraic modelling*. PhD thesis, Vrije Universiteit Amsterdam, 2004. URL: `https://hdl.handle.net/1871.1/56f40c24-8cb4-4359-beb3-c28c3263838e`.
3. Jon Beck. Distributive laws. In *Seminar on Triples and Categorical Homology Theory*, volume 80 of *Lecture Notes in Mathematics*, pages 119–140. Springer Berlin Heidelberg, 1969. `doi:10.1007/BFb0083084`.
4. Filippo Bonchi and Alessio Santamaria. Convexity via weak distributive laws. *Logical Methods in Computer Science*, 18(4):8:1–8:55, 2022. `doi:10.46298/LMCS-18(4:8)2022`.
5. Marta Bunge. On the relationship between composite and tensor product triples. *Journal of Pure and Applied Algebra*, 13(2):139–156, 1978. `doi:10.1016/0022-4049(78)90004-X`.
6. Titouan Carette, Louis Lemonnier, and Vladimir Zamdzhiev. Central submonads and notions of computation: soundness, completeness and internal languages. In *Proceedings of the 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2023)*, pages 1–13. IEEE, 2023. `doi:10.1109/LICS56636.2023.10175687`.
7. Chris Chen, Annabelle McIver, and Carroll Morgan. Probabilistic datatypes. In *Theoretical Aspects of Computing – ICTAC 2024*, volume 15373 of *Lecture Notes in Computer Science*, pages 3–16. Springer Nature Switzerland, 2025. `doi:10.1007/978-3-031-77019-7_1`.
8. Chris Chen, Annabelle McIver, and Carroll Morgan. Forward and backward simulations for partially observable probability. In *Theoretical Aspects of Computing – ICTAC 2025*, volume 16237 of *Lecture Notes in Computer Science*, pages 279–297. Springer Nature Switzerland, 2026. `doi:10.1007/978-3-032-11176-0_17`.
9. Alfred Hoblitzelle Clifford and Gordon Bamford Preston. *The Algebraic Theory of Semigroups, Volume I*, volume 7 of *Mathematical Surveys*. American Mathematical Society, 1961. `doi:10.1090/surv/007.1/01`.
10. Pierre-Louis Curien. Operads, clones, and distributive laws. In *Operads and universal algebra. Proceedings of the summer school and international conference, Tianjin, China, July 5–9, 2010*, volume 9 of *Nankai Series in Pure, Applied Mathematics and Theoretical Physics*, pages 25–49. World Scientific, 2012. `doi:10.1142/9789814365123_0002`.

11 Fredrik Dahlqvist, Louis Parlant, and Alexandra Silva. Layer by layer – combining monads. In *Theoretical Aspects of Computing – ICTAC 2018*, volume 11187 of *Lecture Notes in Computer Science*, pages 153–172. Springer International Publishing, 2018. `doi:10.1007/978-3-030-02508-3_9`.

12 Valeria de Paiva. The dialectica categories. Technical Report UCAM-CL-TR-213, University of Cambridge, Computer Laboratory, 1991. `doi:10.48456/tr-213`.

13 Marcelo Fiore, Gordon Plotkin, and Daniele Turi. Abstract syntax and variable binding. In *Proceedings of the Fourteenth Annual IEEE Symposium on Logic in Computer Science (LICS 1999)*, pages 193–202. IEEE Computer Society, 1999. `doi:10.1109/LICS.1999.782615`.

14 Soichiro Fujii, Yun Chen Tsai, Yoàv Montacute, and Ichiro Hasuo. Monads and distributive laws in substructural contexts (extended version), 2026. `arXiv:2605.13533v1`.

15 Richard Garner. The Vietoris monad and weak distributive laws. *Applied Categorical Structures*, 28(2):339–354, 2020. `doi:10.1007/s10485-019-09582-w`.

16 Richard Garner. Cartesian closed varieties I: the classification theorem. *Algebra Universalis*, 85(4):38, 2024. `doi:10.1007/s00012-024-00869-1`.

17 Alexandre Goy and Daniela Petrişan. Combining probabilistic and non-deterministic choice via weak distributive laws. In *Proceedings of the 35th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2020)*, pages 454–464. ACM, 2020. `doi:10.1145/3373718.3394795`.

18 Martin Hyland, Gordon Plotkin, and John Power. Combining effects: sum and tensor. *Theoretical Computer Science*, 357(1):70–99, 2006. `doi:10.1016/j.tcs.2006.03.013`.

19 Martin Hyland and John Power. The category theoretic understanding of universal algebra: Lawvere theories and monads. In *Computation, Meaning, and Logic: Articles dedicated to Gordon Plotkin*, volume 172 of *Electronic Notes in Theoretical Computer Science*, pages 437–458. Elsevier, 2007. `doi:10.1016/j.entcs.2007.02.019`.

20 Bart Jacobs. Semantics of weakening and contraction. *Annals of Pure and Applied Logic*, 69(1):73–106, 1994. `doi:10.1016/0168-0072(94)90020-5`.

21 Bart Jacobs. Trace semantics for coalgebras. In *Proceedings of the Seventh Workshop on Coalgebraic Methods in Computer Science (CMCS 2004)*, volume 106 of *Electronic Notes in Theoretical Computer Science*, pages 167–184. Elsevier, 2004. `doi:10.1016/j.entcs.2004.02.031`.

22 Bart Jacobs. From multisets over distributions to distributions over multisets. In *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2021)*, pages 1–13. IEEE, 2021. `doi:10.1109/LICS52264.2021.9470678`.

23 Peter Johnstone. Collapsed toposes and cartesian closed varieties. *Journal of Algebra*, 129(2):446–480, 1990. `doi:10.1016/0021-8693(90)90230-L`.

24 André Joyal. Foncteurs analytiques et espèces de structures. In *Combinatoire énumérative*, volume 1234 of *Lecture Notes in Mathematics*, pages 126–159. Springer Berlin Heidelberg, 1986. `doi:10.1007/BFb0072514`.

25 Gregory Maxwell Kelly. On the operads of J.P. May. *Reprints in Theory and Applications of Categories*, 13:1–13, 2005. URL: `http://www.tac.mta.ca/tac/reprints/articles/13/tr13abs.html`.

26 David King and Philip Wadler. Combining monads. In *Functional Programming, Glasgow 1992*, pages 134–143. Springer London, 1993. `doi:10.1007/978-1-4471-3215-8_12`.

27 Bartek Klin and Julian Salamanca. Iterated covariant powerset is not a monad. In *Proceedings of the Thirty-Fourth Conference on the Mathematical Foundations of Programming Semantics (MFPS XXXIV)*, volume 341 of *Electronic Notes in Theoretical Computer Science*, pages 261–276. Elsevier, 2018. `doi:10.1016/j.entcs.2018.11.013`.

28 Anders Kock. Monads on symmetric monoidal closed categories. *Archiv der Mathematik*, 21:1–10, 1970. `doi:10.1007/BF01220868`.

29 Anders Kock. Bilinearity and Cartesian closed monads. *Mathematica Scandinavica*, 29:161–174, 1972. `doi:10.7146/math.scand.a-11042`.

30 Anders Kock. Strong functors and monoidal monads. *Archiv der Mathematik*, 23:113–120, 1972. doi:10.1007/BF01304852.

31 Anders Kock. Commutative monads as a theory of distributions. *Theory and Applications of Categories*, 26:97–131, 2012. URL: http://www.tac.mta.ca/tac/volumes/26/4/26-04abs.html.

32 Dexter Kozen and Alexandra Silva. Multisets and distributions. In *Logics and Type Systems in Theory and Practice – Essays Dedicated to Herman Geuvers on The Occasion of His 60th Birthday*, volume 14560 of *Lecture Notes in Computer Science*, pages 168–187. Springer, 2024. doi:10.1007/978-3-031-61716-4_11.

33 Tom Leinster. *Higher Operads, Higher Categories*, volume 298 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 2004. doi:10.1017/CBO9780511525896.

34 Fabian Lenke, Stefan Milius, and Henning Urbat. A unified treatment of substitution for presheaves, nominal sets, renaming sets, and so on, 2026. arXiv:2602.11907v1.

35 Jack Liell-Cock and Sam Staton. Compositional imprecise probability: A solution from graded monads and Markov categories. *Proceedings of the ACM on Programming Languages*, 9(POPL):54:1–54:31, 2025. doi:10.1145/3704890.

36 Harald Lindner. Affine parts of monads. *Archiv der Mathematik*, 33:437–443, 1980. doi:10.1007/BF01222782.

37 Saunders Mac Lane. *Categories for the Working Mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 1998. doi:10.1007/978-1-4757-4721-8.

38 Ernie Manes and Philip Mulry. Monad compositions I: General constructions and recursive distributive laws. *Theory and Applications of Categories*, 18:172–208, 2007. URL: http://www.tac.mta.ca/tac/volumes/18/7/18-07abs.html.

39 Cristina Matache and Sam Staton. A sound and complete logic for algebraic effects. In *Foundations of Software Science and Computation Structures (FOSSACS 2019)*, volume 11425 of *Lecture Notes in Computer Science*, pages 382–399. Springer, 2019. doi:10.1007/978-3-030-17127-8_22.

40 Dylan McDermott and Tarmo Uustalu. What makes a strong monad? In *Proceedings of the Ninth Workshop on Mathematically Structured Functional Programming (MSFP 2022)*, volume 360 of *Electronic Proceedings in Theoretical Computer Science*, pages 113–133. Open Publishing Association, 2022. doi:10.4204/EPTCS.360.6.

41 Matteo Mio, Ralph Sarkis, and Valeria Vignudelli. Combining nondeterminism, probability, and termination: Equational and metric reasoning. In *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2021)*, pages 1–14. IEEE, 2021. doi:10.1109/LICS52264.2021.9470717.

42 Matteo Mio and Valeria Vignudelli. Monads and quantitative equational theories for nondeterminism and probability. In *31st International Conference on Concurrency Theory (CONCUR 2020)*, volume 171 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 28:1–28:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CONCUR.2020.28.

43 Eugenio Moggi. Notions of computation and monads. *Information and Computation*, 93(1):55–92, 1991. doi:10.1016/0890-5401(91)90052-4.

44 Shawn Ong, Stephanie Ma, and Dexter Kozen. Probabilistic Kleene algebra with angelic nondeterminism. *Proceedings of the ACM on Programming Languages*, 9(PLDI):897–919, 2025. doi:10.1145/3729286.

45 Louis Parlant. *Monad Composition via Preservation of Algebras*. PhD thesis, University College London, 2020. URL: https://discovery.ucl.ac.uk/id/eprint/10112228.

46 Louis Parlant, Jurriaan Rot, Alexandra Silva, and Bas Westerbaan. Preservation of equations by monoidal monads. In *45th International Symposium on Mathematical Foundations of Computer Science (MFCS 2020)*, volume 170 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 77:1–77:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.MFCS.2020.77.

47 Maciej Piróg and Sam Staton. Backtracking with cut via a distributive law and left-zero monoids. *Journal of Functional Programming*, 27:e17, 2017. doi:10.1017/S0956796817000077.

48 Gordon Plotkin and Matija Pretnar. Handling algebraic effects. *Logical Methods in Computer Science*, 9(4):23, 2013. doi:10.2168/LMCS-9(4:23)2013.

49 Aloïs Rosset, Maaike Zwart, Helle Hvid Hansen, and Jörg Endrullis. Correspondence between composite theories and distributive laws. In *Coalgebraic Methods in Computer Science (CMCS 2024)*, volume 14617 of *Lecture Notes in Computer Science*, pages 194–215. Springer, 2024. doi:10.1007/978-3-031-66438-0_10.

50 Michael Shulman. Categorical logic from a categorical point of view. Draft, version of July 28, 2016. URL: https://mikeshulman.github.io/catlog/catlog.pdf.

51 Alex Simpson and Niels Voorneveld. Behavioural equivalence via modalities for algebraic effects. *ACM Transactions on Programming Languages and Systems*, 42(1):4:1–4:45, 2020. doi:10.1145/3363518.

52 Ross Street. The formal theory of monads. *Journal of Pure and Applied Algebra*, 2(2):149–168, 1972. doi:10.1016/0022-4049(72)90019-9.

53 Stanisław Szawiel and Marek Zawadowski. Monads of regular theories. *Applied Categorical Structures*, 23(3):215–262, 2015. doi:10.1007/s10485-013-9331-x.

54 Miki Tanaka and John Power. Pseudo-distributive laws and axiomatics for variable binding. *Higher-Order and Symbolic Computation*, 19(2-3):305–337, 2006. doi:10.1007/s10990-006-8750-x.

55 Miki Tanaka and John Power. A unified category-theoretic semantics for binding signatures in substructural logics. *Journal of Logic and Computation*, 16(1):5–25, 2006. doi:10.1093/logcom/exi070.

56 Regina Tix, Klaus Keimel, and Gordon Plotkin. Semantic domains for combining probability and non-determinism. *Electronic Notes in Theoretical Computer Science*, 222:3–99, 2009. doi:10.1016/j.entcs.2009.01.002.

57 Serge Tronin. Abstract clones and operads. *Siberian Mathematical Journal*, 43(4):746–755, 2002. doi:10.1023/A:1016340806503.

58 Serge Tronin. Operads and varieties of algebras defined by polylinear identities. *Siberian Mathematical Journal*, 47(3):670–694, 2006. doi:10.1007/s11202-006-0067-9.

59 Serge Tronin. On algebras over multicategories. *Russian Mathematics*, 60(2):52–61, 2016. doi:10.3103/S1066369X16020092.

60 Serge Tronin and Oleg Kopp. Matrix linear operads. *Russian Mathematics*, 44(6):50–59, 2000.

61 Yun Chen Tsai, Kittiphon Phalakarn, Sundararaman Akshay, and Ichiro Hasuo. Chance and mass interpretations of probabilities in Markov decision processes. In *36th International Conference on Concurrency Theory (CONCUR 2025)*, volume 348 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 33:1–33:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2025. doi:10.4230/LIPIcs.CONCUR.2025.33.

62 Daniele Varacca and Glynn Winskel. Distributing probability over non-determinism. *Mathematical Structures in Computer Science*, 16(1):87–113, 2006. doi:10.1017/S0960129505005074.

63 Mark Weber. Generic morphisms, parametric representations and weakly cartesian monads. *Theory and Applications of Categories*, 13:191–234, 2004. URL: http://www.tac.mta.ca/tac/volumes/13/14/13-14abs.html.

64 Noam Zilberstein, Dexter Kozen, Alexandra Silva, and Joseph Tassarotti. A demonic outcome logic for randomized nondeterminism. *Proceedings of the ACM on Programming Languages*, 9(POPL):539–568, 2025. doi:10.1145/3704855.

65 Maaike Zwart. *On the Non-Compositionality of Monads via Distributive Laws*. PhD thesis, Department of Computer Science, University of Oxford, September 2020. URL: https://www.cs.ox.ac.uk/publications/publication14285-abstract.html.

66 Maaike Zwart and Dan Marsden. No-go theorems for distributive laws. *Logical Methods in Computer Science*, 18(1):13:1–13:61, 2022. doi:10.46298/lmcs-18(1:13)2022.