No-Go Theorems for Distributive Laws

Maaike Zwart
University of Oxford
maaike.zwart@cs.ox.ac.uk

Dan Marsden
University of Oxford
daniel.marsden@cs.ox.ac.uk

***Abstract*—Monads are commonplace in computer science, and can be composed using Beck's distributive laws. Unfortunately, finding distributive laws can be extremely difficult and error-prone. The literature contains some principles for constructing distributive laws. However, until now there have been no general techniques for establishing when no such law exists.**

We present two families of theorems for showing when there can be no distributive law for two monads. The first widely generalizes a counterexample attributed to Plotkin. It covers all the previous known no-go results for specific pairs of monads, and includes many new results. The second family is entirely novel, encompassing various new practical situations. For example, it negatively resolves the open question of whether the list monad distributes over itself, and also reveals a previously unobserved error in the literature.

I. Introduction

Monads have become a key tool in computer science. They are, amongst other things, used to provide semantics for computational effects such as state, exceptions, and IO [1]. They are also used to structure functional programs [2], [3], and even appear explicitly in the standard library of the Haskell programming language [4].

Monads are a categorical concept. A monad on a category $\mathcal{C}$ consists of an endofunctor T and two natural transformations $1 \Rightarrow T$ and $\mu : T \circ T \Rightarrow T$ satisfying axioms described in Definition II.1. Given two monads with underlying functors S and T, it is natural to ask if $T \circ S$ always carries the structure of a monad. This would, for example, provide a way to combine simple monads together to model more complex computational effects. Unfortunately, monads cannot generally be combined in this way. Beck has shown that the existence of a *distributive law* provides sufficient conditions for $T \circ S$ to form a monad [5]. A distributive law is a natural transformation of type $S \circ T \Rightarrow T \circ S$, satisfying four equations described in Definition II.8. This important idea has since been generalized to notions of distributive law for combining monads with comonads, monads with pointed endofunctors, endofunctors with endofunctors and various other combinations, see for example [6].

General-purpose techniques have been developed for constructing distributive laws [7], [8], [9], [10]. These methods are highly valuable, for as stated in [9]: *"It can be rather difficult to prove the defining axioms of a distributive law."* In fact, it can be so difficult that on occasion a distributive law has been published which later turned out to be incorrect; see [11] for an overview of such cases involving the powerset monad.

978-1-7281-3608-0/19/$31.00

The literature has tended to focus on positive results, either demonstrating specific distributive laws, or developing general-purpose techniques for constructing them. By comparison, there is a relative paucity of negative results, showing when no distributive law can exist. The most well-known result of this type appears in [12], where it is shown that there is no distributive law combining the powerset and probability distribution monads, via a proof credited to Plotkin. This result was strengthened to show that the composite functor carries no monad structure at all in [13]. Recently, the same proof technique was used to show that composing the covariant powerset functor with itself yields an endofunctor that does not carry any monad structure [11], correcting an earlier error in the literature [7]. To the best our knowledge, these are currently the only published impossibility results.

In this paper we present several theorems for proving the absence of distributive laws for large classes of monads on the category of sets and functions. These theorems significantly extend the current understanding of distributive laws. Our results can roughly be divided into two classes:

- Firstly, we generalize Plotkin's method to a general-purpose theorem covering all the previously published no-go results about distributive laws, and yielding new results as well.
- Secondly, we identify an entirely novel approach, distinct from Plotkin's method. This technique leads to new theorems, covering combinations of monads that were previously out of reach.

As one application of our second method, we show that the list monad cannot distribute over itself, resolving an open question [7], [8] and previous error [14] in the literature. In addition, the no-go theorems produced by the second method reveal yet another faulty distributive law in the literature, involving the list and exception monads.

Monads have deep connections with universal algebra. We exploit this algebraic viewpoint on monads by using an explicit description of the algebraic nature of distributive laws from [15], which was inspired by the work in [16]. Formulating our results in algebraic form is a key contribution of our work, simplifying and clarifying the essentials of our proofs, which can be obscured by more direct calculations.

In physics, theorems proving the impossibility of certain things are called *no-go theorems*, because they clearly identify theoretical directions that cannot succeed. We follow this example, and hope that by sharing our results, we prevent others from wasting time on forlorn searches for distributive laws that do not exist.

A. Contributions

We briefly outline our contribution. By taking an algebraic perspective, we demonstrate the non-existence of distributive laws for large classes of monads:

- In Section III we widely generalize the essentials of a counterexample due do Plotkin [12]:
  - We establish in Theorem III.6 purely algebraic conditions under which a no-go result holds. This theorem recovers all the known negative results we are aware of, and many useful new results. The key ingredients for this theorem are binary terms that are idempotent and commutative.
  - In Theorem III.13 we generalize further, showing there is nothing essential about binary terms.
  - Finally, Theorem III.15 eliminates commutativity assumptions, yielding more useful insights.
- In Section IV we present three entirely new no-go theorems:
  - Theorem IV.4 states conditions under which algebraic theories with more than one constant do not combine well with other theories. The usefulness of this theorem is demonstrated by Example IV.5, which identifies a previously unnoticed error in a theorem in the literature.
  - Theorem IV.10 is a general no-go theorem for monads that do not satisfy the so-called *abides* equation [17]. One of its applications is that it negatively answers the open question of whether the list monad distributes over itself.
  - Theorem IV.15 is a general no-go theorem focussing on the combination of idempotence and units. From this theorem it follows that there is no distributive law for the powerset monad over the multiset monad: $PM \Rightarrow MP$.

Also in Section IV, Theorem IV.7 gives conditions under which at most one distributive law can exist. For example, the well-known distributive law for the multiset monad overself is unique.

Throughout this paper we restrict our attention to monads on the category of sets and functions, as this is already an incredibly rich setting. Our results are general purpose in that they are phrased in terms of abstract properties of the algebraic theories corresponding to both monads.

Remark I.1. *To compose monads $\langle S, \eta, \mu\rangle$ and $\langle T, \eta, \mu\rangle$ one may ask if all we really want is some arbitrary monad structure on the functor TS? Generally, a monad structure arising from a distributive law is vastly preferable to an arbitrary one, as it has many desirable properties. For example, there are canonical monad morphisms $S \Rightarrow TS$ and $T \Rightarrow TS$ inducing functors between both the Eilenberg-Moore and Kleisli categories of the corresponding monads. Furthermore, T lifts to a monad on the Eilenberg-Moore category of S, and S lifts to a monad on the Kleisli category of T. More succinctly, there is a relationship between the whole and the parts.*

II. Preliminaries

A. Monads and Distributive Laws

We introduce monads, distributive laws, and various examples that will recur in later sections, primarily to fix notation. The material is standard, and may be skipped by the expert reader.

Definition II.1 (Monad)**.** *For any category $\mathcal{C}$, a **monad** $\langle T, \eta, \mu\rangle$ on $\mathcal{C}$ consists of an endofunctor $T : \mathcal{C} \to \mathcal{C}$, and natural transformations $\eta : 1 \Rightarrow T$ and $\mu : TT \Rightarrow T$ referred to as the **unit** and **multiplication**, satisfying the following axioms:*

$$\mu \cdot T\eta = id \qquad \text{(unit 1)}$$

$$\mu \cdot \eta T = id \qquad \text{(unit 2)}$$

$$\mu \cdot T\mu = \mu \cdot \mu T \qquad \text{(associativity)}$$

Example II.2. *For any set E, the **exception monad** $(- + E)$ is given by:*

- *$(- + E)$ maps a set X to the disjoint union $X + E$.*
- *η^E_X is the left inclusion morphism.*
- *μ^E_X is the identity on X, and collapses the two copies E down to a single copy.*

*When E is a singleton set, this monad is also known as the **maybe monad**, written as $(-)_\perp$.*

Example II.3. *The **list monad** L is given by:*

- *$L(X)$ is the set of all finite lists of elements of X.*
- *$\eta^L_X(x)$ is the singleton list $[x]$.*
- *μ^L_X concatenates a list of lists.*

Example II.4. *The **multiset monad** M is given by:*

- *$M(X)$ is the set of all finite multisets of elements of X.*
- *$\eta^M_X(x)$ is the singleton multiset $\{x : 1\}$.*
- *μ^M_X takes a union of multisets, adding multiplicities.*

Example II.5. *The **finite powerset monad** P is given by:*

- *$P(X)$ is the set of all finite subsets of X.*
- *$\eta^P_X(x)$ is the singleton set $\{x\}$.*
- *μ^P_X takes a union of sets.*

Example II.6. *The **probability distribution monad** D is given by:*

- *$D(X)$ is the set of all finitely supported probability distributions over X.*
- *$\eta^D_X(x)$ is the point distribution at x.*
- *$\mu^D(e)(x)$ is the weighted average $\sum_{d \in \mathrm{supp}(e)} e(d)d(x)$.*

Example II.7. *For any set of states R, the **reader monad** $(-)^R$ is given by:*

- *X^R is the set of functions from R to X.*
- *$\eta^R_X(x)$ is constantly x.*
- *$\mu^R_X(f)(r) = f(r)(r)$.*

Given a pair of monads, sufficient conditions for the composite functor to form a monad are given by Beck's distributive laws [5]:

Definition II.8 (Distributive Law)**.** *Given monads S and T, a **distributive law** for monad composition TS is a natural transformation $\lambda : ST \Rightarrow TS$ satisfying the following axioms:*

$$\lambda \cdot \eta^S T = T\eta^S \tag{unit1}$$

$$\lambda \cdot S\eta^T = \eta^T S \tag{unit2}$$

$$\lambda \cdot \mu^S T = T\mu^S \cdot \lambda S \cdot S\lambda \tag{multiplication1}$$

$$\lambda \cdot S\mu^T = \mu^T S \cdot T\lambda \cdot \lambda T \tag{multiplication2}$$

Remark II.9. *For a pair of monads S, T the expression "S distributes over T" is often used. This phrasing is somewhat ambiguous and prone to errors. We will therefore explicitly state the type of the natural transformation, for example "there is a distributive law of type $ST \Rightarrow TS$".*

Theorem II.10 (Beck [5])**.** *Let $\mathcal{C}$ be a category, and $\langle S, \eta^S, \mu^S \rangle$ and $\langle T, \eta^T, \mu^T \rangle$ two monads on $\mathcal{C}$. If $\lambda : ST \Rightarrow TS$ is a distributive law, then TS carries a monad structure with unit $\eta^T\eta^S$ and multiplication $\mu^T\mu^S \cdot T\lambda S$.*

Example II.11 (The Ring Monad [5])**.** *The motivating example of a distributive law involves the list monad and the Abelian group monad, $\lambda : LA \Rightarrow AL$. It captures exactly the distributivity of multiplication over addition:*

$$\lambda\,(a \cdot (b + c)) = (a \cdot b) + (a \cdot c) \tag{1}$$

The term 'distributive law' is motivated by this example, and many other distributive laws exploit similar algebraic properties. As we will see in Section IV, caution is needed though: the validity of an equation such as (1) *does not automatically imply the existence of a distributive law.*

Example II.12 (Multiset Monad Composed with Itself)**.** *The multiset monad distributes over itself in a manner analogous to distributing multiplication over addition.*

B. Algebraic Theories and Composite Theories

We now outline the connections between algebras, monads, and distributive laws that we will require in later sections.

Definition II.13 (Algebraic Theories)**.** *An **algebraic signature** is a set of operation symbols Σ, each with an associated natural number referred to as its **arity**. The set of Σ**-terms** over a set X contains X as **variables** and is inductively closed under forming terms $\sigma(t_1, ..., t_n)$ for an n-ary operation symbol σ and terms $t_1, ..., t_n$.*

*An **algebraic theory** $\mathbb{T}$ consists of a signature $\Sigma^{\mathbb{T}}$, and a set $E^{\mathbb{T}}$ of pairs of Σ-terms referred to as **equations** or **axioms**. We will often write a pair $(s, t) \in E^{\mathbb{T}}$ as $s =_{\mathbb{T}} t$ or simply $s = t$ when convenient. For a subset $Y \subseteq X$ and a term t we write $Y \vdash_{\mathbb{T}} t$ or $Y \vdash t$ to indicate that the variables appearing in t are contained in the **variable context** Y. The precise set of variables appearing t will be denoted $\mathrm{var}(t)$.*

The following is well known [18], [19], [20]:

Proposition II.14. *Given a theory $(\Sigma^{\mathbb{T}}, E^{\mathbb{T}})$, the **free model monad** over that signature maps X to the set of $\Sigma^{\mathbb{T}}$-terms over X, quotiented by provable equality in equational logic from the axioms $E^{\mathbb{T}}$. The unit maps a variable to its corresponding equivalence class, and the multiplication flattens a term-of-terms to a term in the obvious way. If a monad is isomorphic to a free model monad, it is said to be **presented** by the corresponding theory.*

Example II.15 (Monoids)**.** *The algebraic **theory of monoids** has a signature containing a constant and a binary operation, satisfying left and right unitality and associativity. The **theory of commutative monoids** extends this theory with the commutativity equation. The **theory of join semilattices** further extends the theory of commutative monoids with an additional idempotence axiom.*

The corresponding free model monads are the list, multiset and finite powerset monads respectively.

As shown in [21], many monads used to describe computational effects have natural algebraic presentations.

Example II.16. *An algebraic presentation of the reader monad of Example II.7, with state space $\{0, 1\}$, has a signature containing a single binary operation. Intuitively, $x * y$ is a process that proceeds as x if the state is 0, and y otherwise. This operation should satisfy:*

$$x * x = x \quad \text{and} \quad (w * x) * (y * z) = w * z$$

These axioms generalize naturally to larger state spaces.

Example II.17. *The distribution monad Example II.6 can be presented by a family of binary operations $+^p$, for $p \in (0, 1)$, satisfying the following axioms [22], [23]:*

$$x +^p x = x$$

$$x +^p y = y +^{1-p} x$$

$$x +^p (y +^r z) = (x +^{\frac{p}{p+(1-p)r}} y) +^{p+(1-p)r} z$$

We require the key notion of composite theories, as introduced in [15].

Definition II.18 (Composite Theories)**.** *Let $\mathbb{U}$ be an algebraic theory that contains two theories $\mathbb{S}$ and $\mathbb{T}$.*

1) *A term in $\mathbb{U}$ is **separated** if it is of the form $t[s_x/x]$, where $X \vdash_{\mathbb{T}} t$ and s_x is a family of $\mathbb{S}$-terms indexed by $x \in X$.*
2) *If there are terms $X \vdash_{\mathbb{T}} t, X' \vdash_{\mathbb{T}} t'$ and families of $\mathbb{S}$-terms $s_x, s'_{x'}$, we say that $t[s_x/x]$ and $t'[s'_{x'}/x']$ are **equal modulo** $(\mathbb{T}, \mathbb{S})$ if there are functions $h : X \to Y, h' : X' \to Y$ and terms $\bar{s}_y$, such that:*
   a) *$t[h(x)/x] =_{\mathbb{T}} t'[h'(x')/x']$*
   b) *$s_x =_{\mathbb{S}} \bar{s}_{h(x)} (\forall x \in X)$*
   c) *$s'_{x'} =_{\mathbb{S}} \bar{s}_{h'(x')} (\forall x' \in X')$*

*The theory $\mathbb{U}$ is said to be a **composite** of $\mathbb{T}$ after $\mathbb{S}$, if every term u in $\mathbb{U}$ is equal to a separated term, and moreover this term is **essentially unique** in the sense that if v, v' are separated and $v =_{\mathbb{U}} u =_{\mathbb{U}} v'$ then v and v' are equal modulo $(\mathbb{T}, \mathbb{S})$. Note that this is an oriented notion, a composite of $\mathbb{T}$ after $\mathbb{S}$ is not equivalent to a composite of $\mathbb{S}$ after $\mathbb{T}$.*

In this paper, we shall use an equivalent definition of equality modulo $(\mathbb{T}, \mathbb{S})$.

Theorem II.19. *Let $\mathbb{S}$ and $\mathbb{T}$ be two algebraic theories, and let $\mathbb{U}$ be an algebraic theory that contains both $\mathbb{S}$ and $\mathbb{T}$. For terms $X \vdash_{\mathbb{T}} t$, $X' \vdash_{\mathbb{T}} t'$ and families of $\mathbb{S}$-terms $s_x, s'_{x'}$, the following are equivalent:*

1) *The terms $t[s_x/x]$ and $t'[s'_{x'}/x']$ are equal modulo $(\mathbb{T}, \mathbb{S})$.*
2) *There are functions $f : X \to Z$, $f' : X' \to Z$ satisfying:*
   a) $t[f(x)/x] =_{\mathbb{T}} t'[f'(x')/x']$.
   b) $f(x_1) = f(x_2) \Leftrightarrow s_{x_1} =_{\mathbb{S}} s_{x_2}$.
   c) $f'(x'_1) = f'(x'_2) \Leftrightarrow s'_{x'_1} =_{\mathbb{S}} s'_{x'_2}$.
   d) $f(x) = f'(x') \Leftrightarrow s_x =_{\mathbb{S}} s'_{x'}$.

Proof. That condition 2 implies condition 1 is straightforward. Taking Y to be the union of the ranges of f and f', requirements 2b-2d ensure that we can choose $\bar{s}_y$ such that:

$$\bar{s}_y = \begin{cases} s_x, & \text{if } f(x) = y \\ s_{x'}, & \text{if } f'(x') = y. \end{cases}$$

To show that condition 1 implies condition 2, notice that by transitivity of equality, h and h' from Definition II.18 already have the properties:

1) $t[h(x)/x] =_{\mathbb{T}} t'[h'(x')/x']$.
2) $h(x_1) = h(x_2) \Rightarrow s_{x_1} =_{\mathbb{S}} s_{x_2}$.
3) $h'(x'_1) = h'(x'_2) \Rightarrow s'_{x'_1} =_{\mathbb{S}} s'_{x'_2}$.
4) $h(x) = h'(x') \Rightarrow s_x =_{\mathbb{S}} s'_{x'}$.

So all we need to show are the reverse implications of the latter three points. To this end, pick a function $g : Y \to Z$ such that:

- If $s_{x_1} =_{\mathbb{S}} s_{x_2}$, then $g(h(x_1)) = g(h(x_2))$
- If $s'_{x'_1} =_{\mathbb{S}} s'_{x'_2}$, then $g(h'(x'_1)) = g(h'(x'_2))$
- If $s_x =_{\mathbb{S}} s'_{x'}$, then $g(h(x)) = g(h'(x'))$

It is clear that such a g exists. The compositions $f = g \circ h$ and $f' = g \circ h'$ then preserve properties 1-4. By definition, they also satisfy the reverse implications, and so they satisfy condition 2. ☐

Composite theories and monads induced by distributive laws are intimately related.

Theorem II.20 (Piróg & Staton [15])**.** *Let S and T be* **Set**-*monads presented by theories $\mathbb{S}$ and $\mathbb{T}$. There is a distributive law of type $ST \Rightarrow TS$ if and only if there is a composite theory of $\mathbb{T}$ after $\mathbb{S}$.*

We will frequently exploit Theorem II.20 by showing that no composite theory exists, and therefore no distributive law. In addition, it will sometimes be useful to know the actual action of the distributive law promised by Theorem II.20, if we know the composite theory.

Proposition II.21. *Let $\mathbb{S}$ and $\mathbb{T}$ be algebraic theories presenting monads S and T, and let $\mathbb{U}$ be a composite theory of $\mathbb{T}$ after $\mathbb{S}$. Then the free model monad U is isomorphic to the composition of $T \circ S$ via the distributive law mapping the equivalence class of representative $s[t_x/x]$ to the suitable equivalence class of a separated term in $\mathbb{U}$ equal to $s[t_x/x]$.*

Proof. From [5] we know that we can express the distributive law λ in terms of the multiplication μ^U of U:

$$\lambda = \mu^U \cdot \eta^T ST \eta^S$$

We know the action of μ^U since it is the multiplication of the free model monad of theory $\mathbb{U}$. Combining this knowledge with the knowledge that $\mathbb{U}$ is a composite of $\mathbb{T}$ after $\mathbb{S}$ gives us the action of λ. The essential uniqueness property ensures that λ as stated above is well-defined. ☐

III. General Plotkin Theorems

In this section we develop algebraic generalizations of the counterexample attributed to Gordon Plotkin in [12]. This counterexample showed that there is no distributive law of type $DP \Rightarrow PD$, where D is the distribution monad of Example II.6 and P is the finite powerset monad of Example II.5. We present a slight rephrasing of this counterexample, augmented with commentary indicating the main proof ideas that will be used in the later generalizations. Thoughout this section we adopt the notational conventions of [12], [15] whenever possible to ease comparison with those papers.

Counterexample III.1 (Probability does not distribute over non-determinism)**.** *Assume, for contradiction, that there is a distributive law of type $\lambda : DP \Rightarrow PD$. Fix the set $X = \{a, b, c, d\}$, and consider the element $\Xi \in DP(X)$ defined by:*

$$\Xi = \{a, b\} +^{\frac{1}{2}} \{c, d\}$$

We define three functions $f_1, f_2, f_3 : X \to X$:

$$\begin{array}{lll} f_1(a) = a & f_2(a) = a & f_3(a) = a \\ f_1(b) = b & f_2(b) = b & f_3(b) = a \\ f_1(c) = a & f_2(c) = b & f_3(c) = c \\ f_1(d) = b & f_2(d) = a & f_3(d) = c \end{array}$$

The plan of the proof is to analyze how Ξ travels around the naturality square for λ, for each of the three functions. The element Ξ and the three functions have been carefully chosen so that the distributive law unit axioms can be applied during the proof.

$$\begin{array}{ccc} DP(X) & \xrightarrow{\lambda_X} & PD(X) \\ {\scriptstyle DP(f_i)}\downarrow & & \downarrow{\scriptstyle PD(f_i)} \\ DP(X) & \xrightarrow[\lambda_X]{} & PD(X) \end{array} \tag{2}$$

We proceed as follows:

- *Trace Ξ around the naturality square (2) for both f_1 and f_2. We note that as $\{-,-\}$ is commutative, and $+^{\frac{1}{2}}$ is idempotent:*

  $$DP(f_1)(\Xi) = \eta^D_{PX}\{a, b\} = DP(f_2)(\Xi)$$

  Commutativity, and particularly idempotence, will be important ideas our subsequent generalizations. For both f_1

and f_2 we can apply the first distributive law unit axiom to conclude that:

$$\lambda_X \circ DP(f_1)(\Xi) = \{\eta^D_X(a), \eta^D_X(b)\} = \lambda_X \circ DP(f_2)(\Xi)$$

Now a careful consideration of the actions $PD(f_1)$ and $PD(f_2)$ allows us to deduce that $\lambda_X(\Xi)$ must be a subset of:

$$\{\eta^D_X(a), \eta^D_X(b), \eta^D_X(c), \eta^D_X(d)\}$$

This part is less straightforward to generalize. In principle it involves inverse images of equivalence classes of terms in one algebraic theory, with variables labelled by equivalence classes of terms in a second algebraic theory. This motivates our move to an explicitly algebraic formulation. We can see this whole step as establishing an upper bound on the set of variables appearing in $\lambda_X(\Xi)$.

- *We then trace Ξ around the naturality square (2) for f_3. In this case, we exploit the idempotence of the operation $\{-,-\}$ to conclude:*

$$DP(f_3)(\Xi) = \{a\} +^{\frac{1}{2}} \{c\} \tag{3}$$

This indicates idempotence is actually an important aspect of both monads for this argument to work. Equation (3) allows us to apply the second unit axiom for λ to conclude:

$$\lambda_X \circ DP(f_3)(\Xi) = \{a +^{\frac{1}{2}} c\}$$

By considering the action of $PD(f_3)$ as before, we conclude that $\lambda_X(\Xi)$ must contain an element mapped onto $a +^{\frac{1}{2}} c$ by $PD(f_3)$, placing a lower bound on the set of variables that appear in $\lambda_X(\Xi)$.

- *The lower and upper bounds established in the previous two steps contradict each other, and so no distributive law of type $DP \Rightarrow PD$ can exist.*

In summary, the argument requires two components:

- Some operations satisfying certain algebraic equational properties such as idempotence and commutativity.
- Some slightly more mysterious properties of our monads, making the "inverse image" parts of the argument work correctly.

Remark III.2. *The original counterexample is actually shown for what is known as the free real cone or finite valuation monad, as this requires slightly weaker assumptions. We state it here for the distribution monad simply to avoid the distraction of introducing yet another monad. This is essentially a cosmetic decision, and our later results are equally applicable to the original counterexample for the free real cone monad.*

Our aim in this section is to extract general methods for showing no-go results for constructing distributive laws, derived from the essential steps in this counterexample. In order to do this, we isolate sufficient conditions on algebraic theories inducing two monads, such that there can be no distributive law between them. Earlier generalizations of this counterexample have appeared in [11], [13]. All the existing approaches involve direct calculations with the distributive law axioms, leading to somewhat opaque conditions. They also remain limited to the case where one of the two monads is the powerset monad, restricting their scope of application.

We introduce terminology for some special sets of terms in an algebraic theory. The theorems are stated in terms of these special sets, which in some cases can restrict the scope for which certain "global" conditions need to apply, broadening the range of applicability.

Definition III.3 (Universal Terms)**.** *For an algebraic theory, we say that a set of terms T is:*

- ***Universal*** *if every term is provably equal to a term in T.*
- ***Stable*** *if T is closed under substitution of variables for variables.*

Example III.4. *Some examples of universal and stable sets:*

[1] For any theory, the set of all terms is a stable universal set.

[2] For the theory of real vector spaces, every term is equal to a term in which scaling by the zero element does not appear. Terms that do not contain the scale by zero operation are clearly also stable under variable renaming. Therefore the terms not containing the scale by zero operation are a stable universal set.

[3] In the theory of groups, every term is equal to a term in which no subterm and its inverse are "adjacent". This set is therefore universal. It is not stable, as variable renaming may introduce a subterm adjacent to its inverse.

Remark III.5. *On first reading, it is probably easiest to take the universal stable sets required in subsequent theorems to be the set of all terms in a theory. This is by far the most common case.*

Throughout this section, the variable labels for any algebraic theory will range over the natural numbers. We will also write n for the set $\{1, ..., n\}$, so for example, $2 \vdash t$ means t is a term containing at most two variables.

We proceed in three steps. Theorem III.6 is an algebraic generalization of Plotkin's counterexample, capturing the algebraic properties required of both theories in order for a proof of this type to work. In Theorem III.13 we generalize further, removing the restriction to binary terms that was sufficient for the original application. This generalization complicates the proof slightly, and so to clarify the methods involved we present two separate theorems. Finally, Theorem III.15 addresses the need for commutativity-like axioms, and brings further combinations of monads into scope.

Theorem III.6. *Let $\mathbb{P}$ and $\mathbb{V}$ be two algebraic theories, $T_{\mathbb{P}}$ a stable universal set of $\mathbb{P}$-terms, and $T_{\mathbb{V}}$ a stable universal set of $\mathbb{V}$-terms. If there are terms:*

$$2 \vdash_{\mathbb{P}} p \qquad \textit{and} \qquad 2 \vdash_{\mathbb{V}} v$$

such that:

(P1) p is commutative:

$$2 \vdash p(1,2) = p(2,1)$$

(P2) p is idempotent:

$$1 \vdash p(1,1) = 1$$

(P3) For all $p' \in T_{\mathbb{P}}$:

$$\Gamma \vdash p(1,2) = p' \quad \Rightarrow \quad 2 \vdash p'$$

(V1) v is idempotent:

$$1 \vdash v(1,1) = 1$$

(V2) For all $v' \in T_{\mathbb{V}}$, and x a variable:

$$\Gamma \vdash x = v' \quad \Rightarrow \quad \{x\} \vdash v'$$

(V3) For all $v' \in T_{\mathbb{V}}$:

$$\Gamma \vdash v(1,2) = v' \quad \Rightarrow \quad \neg(\{1\} \vdash v' \vee \{2\} \vdash v')$$

Then there is no composite theory of $\mathbb{P}$ after $\mathbb{V}$.

Remark III.7. *Properties (P3), (V2), and (V3) are constraints on the variables appearing in certain terms, which are needed for the "inverse image" part of Counterexample III.1. Property (P3) states that any term equal to the special binary term p can have at most two free variables. Property (V2) states that any term equal to a variable can only contain that variable, and property (V3) states that any term equal to the special binary term v must have at least two free variables. Notice that the upper/lower bound principle from the original argument is reflected in these conditions.*

Proof. Assume for contradiction that a composite theory $\mathbb{U}$ of $\mathbb{P}$ after $\mathbb{V}$ exists. Then as $\mathbb{U}$ is composite, there exist $X \vdash p'$ and $\Gamma \vdash v'_x$ for each $x \in X$ such that:

$$\Gamma \vdash_{\mathbb{U}} v(p(1,2), p(3,4)) = p'[v'_x/x] \tag{4}$$

Without loss of generality, we may assume $p' \in T_{\mathbb{P}}$ and $v'_x \in T_{\mathbb{V}}$ by universality. Define partial function f_1 as follows:

$$f_1(1) = f_1(3) = 1$$
$$f_1(2) = f_1(4) = 2$$

Then, using this substitution of variables and assumption (V1):

$$\Gamma \vdash p(1,2) = p'[v'_x[f_1]/x]$$

As these are two separated terms, by essential uniqueness there are functions $g_1 : 2 \to Z$ and $g_2 : X \to Z$ such that:

$$p[g_1] = p'[g_2] \qquad \text{(Thm II.19, 2a)}$$
$$g_1(1) \neq g_1(2) \qquad \text{(Thm II.19, 2b)}$$
$$g_1(1) = g_2(x) \Leftrightarrow 1 = v'_x[f_1] \qquad \text{(Thm II.19, 2d)}$$
$$g_1(2) = g_2(x) \Leftrightarrow 2 = v'_x[f_1] \qquad \text{(Thm II.19, 2d)}$$

As $T_{\mathbb{P}}$ is stable, any variable renaming of p' is also in $T_{\mathbb{P}}$. And so by assumption (P3) we must have, for all $x \in \mathrm{var}(p')$: $g_2(x) = g_1(1)$ or $g_2(x) = g_1(2)$, which means for each v'_x:

$$\Gamma \vdash 1 = v'_x[f_1] \quad \vee \quad \Gamma \vdash 2 = v'_x[f_1]$$

Then using assumption (V2) and the preimage of f_1, for all v'_x:

$$\{1,3\} \vdash v'_x \quad \vee \quad \{2,4\} \vdash v'_x \tag{5}$$

Define a second partial function f_2 as follows:

$$f_2(1) = f_2(4) = 1$$
$$f_2(2) = f_2(3) = 2$$

Using this substitution and assumptions (P1) and (V1):

$$\Gamma \vdash p(1,2) = p'[v'_x[f_2]/x]$$

Again, by essential uniqueness and assumption (P3), for all v'_x:

$$\Gamma \vdash 1 = v'_x[f_2] \quad \vee \quad \Gamma \vdash 2 = v'_x[f_2]$$

and so, using assumption (V2), for all v'_x:

$$\{1,4\} \vdash v'_x \qquad \vee \qquad \{2,3\} \vdash v'_x \tag{6}$$

Combining conditions (5) and (6), for all v'_x:

$$\bigvee_{n \in 4} \{n\} \vdash v'_x \tag{7}$$

Finally, we define a third partial function f_3 as follows:

$$f_3(1) = f_3(2) = 1$$
$$f_3(3) = f_3(4) = 2$$

Using this final substitution and (P2):

$$\Gamma \vdash v(1,2) = p'[v'_x[f_3]/x]$$

By essential uniqueness (and consistency of $\mathbb{P}$), there exists x such that:

$$\Gamma \vdash v(1,2) = v'_x[f_3]$$

As $T_{\mathbb{V}}$ is stable, each $v'_x[f_3]$ is a element of $T_{\mathbb{V}}$, and so by (V3) this contradicts Equation (7). Therefore the assumed composite theory cannot exist. □

The following corollary reflects our real interest in monads:

Corollary III.8. *If monads P and V have presentations $\mathbb{P}$ and $\mathbb{V}$ such that the conditions of Theorem III.6 can be satisfied, then there is no distributive law of type $VP \Rightarrow PV$.*

The subsequent Theorems III.13 and III.15 have similar corollaries, we will not state these explicitly.

Example III.9 (Powerset and Distribution Monads)**.** *Consider the terms $1 \vee 2$ and $1 +^{\frac{1}{2}} 2$ in the theories representing the powerset and distribution monads of Examples II.15 and II.17. Since both of these terms are binary, commutative, and idempotent, and the remaining axioms are satisfied, Theorem III.6 captures the known results that there are no distributive laws of type $DP \Rightarrow PD$ [12], $PP \Rightarrow PP$ [11], or $PD \Rightarrow DP$ [24] (the latter is stated without proof). In addition, Theorem III.6 yields the new result that there is no distributive law of type $DD \Rightarrow DD$, completing the picture for these monads.*

Example III.10 (Powerset and Distribution Monads again)**.** *We can also consider the distribution monad to be presented by binary operations $+^p$ with p in the* closed *interval $[0,1]$, and in fact this is the more common formulation. In this case, Theorem III.6 can still be directly applied, without having to move to the more parsimonious presentation. We simply note* *that the terms not involving the operations $+^1$ and $+^0$ form a stable universal set satisfying the required axioms. The results discussed in the previous example can then be recovered using the conventional presentation of the distribution monad.*

Non-Example III.11 (Reader Monad)**.** *It is well known that the reader monad distributes over itself. Looking at the presentation of the reader monad given in Example II.16, we see that although it has idempotent terms, there is no commutative term and hence Theorem III.6 does not apply.*

A natural question to ask with regard to Theorem III.6 is whether the choice of *binary* terms for both p and v is necessary. We thank Prakash Panangaden for posing this question during an informal presentation of an earlier version of this work [25]. The answer is that we can generalize to terms with any arities strictly greater than one. We first introduce a lemma that is central to establishing the upper bound part of the argument.

Lemma III.12. *Let n, m be strictly positive natural numbers, and σ a fixed-point free permutation of $\{1, ..., m\}$. For distinct variables a_i^j, $1 \leq i \leq m$, $1 \leq j \leq n$, the sets:*

$$\begin{gathered}\{a^1_{i_1}, a^2_{i_1}, a^3_{i_1}, \dots, a^n_{i_1}\}\\ \{a^1_{i_2}, a^2_{\sigma(i_2)}, a^3_{i_2}, \dots, a^n_{i_2}\}\\ \vdots\\ \{a^1_{i_n}, a^2_{i_n}, a^3_{i_n}, \dots, a^n_{\sigma(i_n)}\}\end{gathered}$$

have at most one common element. Here, each i_k is an element of $\{i \mid 1 \leq i \leq m\}$, not necessarily unique.

We then get a more general variant of Theorem III.6.

Theorem III.13. *Let $\mathbb{P}$ and $\mathbb{V}$ be two algebraic theories, $T_{\mathbb{P}}$ a stable universal set of $\mathbb{P}$-terms, and $T_{\mathbb{V}}$ a stable universal set of $\mathbb{V}$-terms. If there are terms:*

$$m \vdash_{\mathbb{P}} p \qquad \text{and} \qquad n \vdash_{\mathbb{V}} v$$

such that:

(P4) p is stable under a fixed-point free permutation σ:

$$m \vdash p = p[\sigma]$$

(P5) p is idempotent:

$$1 \vdash p[1/i] = 1$$

(P6) For all $p' \in T_{\mathbb{P}}$:

$$\Gamma \vdash p = p' \quad \Rightarrow \quad m \vdash p'$$

(V4) v is idempotent:

$$1 \vdash v[1/i] = 1$$

(V5) For all $v' \in T_{\mathbb{V}}$, and x a variable:

$$\Gamma \vdash x = v' \quad \Rightarrow \quad \{x\} \vdash v'$$

(V6) For all $v' \in T_{\mathbb{V}}$:

$$\Gamma \vdash v = v' \quad \Rightarrow \quad \neg\left(\bigvee_{i\in\Gamma} \{i\} \vdash v'\right)$$

Then there is no composite theory of $\mathbb{P}$ after $\mathbb{V}$.

Remark III.14. *The required properties are generalizations of the binary conditions in Theorem III.6. Most are straightforward, but axiom (P4), the analogue of binary commutativity, is perhaps slightly surprising. Here we only require stability under a single fixed-point free permutation.*

Proof. Assume for contradiction that a composite theory $\mathbb{U}$ of $\mathbb{P}$ after $\mathbb{V}$ exists. Let a_i^j, $1 \leq i \leq m$, $1 \leq j \leq n$ denote distinct natural numbers. Then as $\mathbb{U}$ is composite, there exist $X \vdash p'$ and $\Gamma \vdash v'_x$ for each $x \in X$ such that:

$$\Gamma \vdash_{\mathbb{U}} v(p(a^1_1, \dots, a^1_m), \dots, p(a^n_1, \dots, a^n_m)) = p'[v'_x/x]$$

Without loss of generality, we may assume $p' \in T_{\mathbb{P}}$ and $v'_i \in T_{\mathbb{V}}$ by universality.

Define substitution f_1 as follows:

$$f_1(a_i^j) = a_i^1$$

We then have:

$$\Gamma \vdash_{\mathbb{U}} v(p(a^1_1, \dots, a^1_m), \dots, p(a^1_1, \dots, a^1_m)) = p'[v'_x[f_1]/x]$$

By assumption (V4)

$$\Gamma \vdash_{\mathbb{U}} p(a^1_1, \dots, a^1_m) = p'[v'_x[f_1]/x]$$

As $T_{\mathbb{P}}$ is stable, any variable renaming of p' is also in $T_{\mathbb{P}}$. Therefore by essential uniqueness and (P6):

$$\forall x\ \exists i:\ \Gamma \vdash a_i^1 = v'_x[f_1]$$

By assumption (V5):

$$\forall x\ \exists i:\ \{a_i^1\} \vdash v'_x[f_1]$$

And so:

$$\forall x\ \exists i:\ \{a_i^1, \dots, a_i^n\} \vdash v'_x \tag{8}$$

Now we define a family of substitutions for $2 \leq k \leq n$ as follows:

$$f_k(a_i^j) = \begin{cases} a^k_{\sigma(i)} & \text{if } j = k \\ a_i^k & \text{otherwise}\end{cases}$$

If we follow a similar argument as before, also exploiting assumption (P4), we conclude that:

$$\forall x, k\ \exists i_k : \{a^k_{i_k}\} \vdash v'_x[f_k]$$

And so:

$$\forall x, k\ \exists i_k:\ \{a^j_{\sigma^{-1}(i_k)} \mid j = k\} \cup \{a^j_{i_k} \mid j \neq k\} \vdash v'_x \tag{9}$$

Then we note that by Lemma III.12, conditions (8) and (9):

$$\forall x\ \exists i, j:\ \{a_i^j\} \vdash v'_x \tag{10}$$

Define another substitution:

$$f_{n+1}(a_i^j) = a_1^j$$

Applying this substitution:

$$\Gamma \vdash v(p(a^1_1, \dots, a^1_1), \dots, p(a^n_1, \dots, a^n_1)) = p'[v'_x[f_{n+1}]/x]$$

Using assumption (P5):

$$\Gamma \vdash v(a_1^1, \ldots, a_1^n) = p'[v'_x[f_{n+1}]/x]$$

By essential uniqueness and consistency:

$$\exists x : \; \Gamma \vdash v(a_1^1, \ldots, a_1^n) = v'_x[f_{n+1}]$$

As $T_{\mathbb{V}}$ is stable, this $v'_x[f_{n+1}]$ is an element of $T_{\mathbb{V}}$, and so by assumption (V6), v'_x must contain at least two variables, but this contradicts conclusion (10), and so the assumed composite theory cannot exist. □

It is clear that the simpler Theorem III.6 is a special case of Theorem III.13. Besides providing greater generality, the main point of Theorem III.13 is that it clearly demonstrates that there is nothing special about binary terms. This further clarifies our understanding of what abstract properties make the original Plotkin counterexample work. By moving to such a high level of abstraction it is also easier to see that our second method, described in Section IV, is not simply a further generalization of Plotkin's counterexample, as it makes fundamentally different assumptions upon the underlying algebraic theories.

In Theorem III.6 we require the special term p to be commutative in order to establish that no composite theory exists. In Theorem III.13 this commutativity was generalized to stability under the action of a fixed-point free permutation. It is natural to consider whether axioms of this form are essential to this type of proof. In fact, this is not the case, and a similar no-go theorem can be established under modified assumptions that make no use of commutativity.

Theorem III.15. *Let $\mathbb{P}$ and $\mathbb{V}$ be two algebraic theories, $T_{\mathbb{P}}$ a stable universal set of $\mathbb{P}$-terms, and $T_{\mathbb{V}}$ a stable universal set of $\mathbb{V}$-terms. If there are terms:*

$$2 \vdash_{\mathbb{P}} p \qquad \textit{and} \qquad 2 \vdash_{\mathbb{V}} v$$

such that axioms (P2), (P3), (V1), (V2) and (V3) hold, and:

(P7) For all $p' \in T_{\mathbb{P}}$, and x a variable:

$$\Gamma \vdash x = p' \quad \Rightarrow \quad \{x\} \vdash p'$$

(P8) For all $p' \in T_{\mathbb{P}}$:

$$\Gamma \vdash p(1,2) = p' \quad \Rightarrow \quad \neg(\emptyset \vdash p')$$

Then there is no composite theory of $\mathbb{P}$ after $\mathbb{V}$.

Proof. We only have room for a brief outline of the details. As in the proof of Theorem III.6, we start from Equation (4). In this case the proof is more delicate. We must establish bounds on the variables appearing in both the terms v'_x *and* the term p'. In order to do so, we must consider substitutions replacing variables with more complex terms, contrasting with the previous arguments. Pursuing this approach, a case analysis eventually lets us exploit axiom (P8) to derive a contradiction with regard to the variables that must appear in a substitution instance of p'. □

Remark III.16. *The proofs in this section require different types of substitutions, and this difference impacts their scope of application. The proofs of Theorems III.6 and III.13 only require variable-for-variable substitutions, and actually preclude the existence of distributive laws for pointed endofunctors, generalizing [11, Theorem 2.4]. The proof of Theorem III.15 requires more complex substitutions, implicitly assuming the multiplication axioms. Therefore, Theorem III.15 only applies to distributive laws between monads.*

Example III.17. *If we consider the algebraic theory of an idempotent binary operation, Theorem III.15 shows that the induced monad cannot distribute over itself. This remains true if we add either units or associativity, showing various non-commutative variants of non-determinism cannot be distributed over themselves.*

Similarly, if we denote any of these monads by T, there is no distributive law $DT \Rightarrow TD$, where D is the distribution monad.

Non-Example III.18 (Reader Monad)**.** *The presentation of the binary reader monad given in Example II.16 satisfies:*

$$x = (x * y) * (z * x)$$

Since an idempotent term must contain variables, this theory has no term p satisfying both axioms (P2) and (P3). So as expected, we cannot apply Theorem III.15 to the binary reader monad.

IV. No-Go Theorems Beyond Plotkin

So far, all our impossibility results involve monads with an idempotent term in their corresponding algebraic theories. But the absence of an idempotent term does not guarantee the existence of a distributive law. Consider the list monad for example. This monad is quite similar to the multiset monad, and we observed in Example II.12 that the "times over plus" law of Equation (1) induces a distributive law for the multiset monad over itself. If we assume this also yields a distributive law for the list monad over itself, then from one of the multiplication axioms:

$$\lambda([[a,b],[c,d]]) = [[a,c],[a,d],[b,c],[b,d]]$$

whilst from the other:

$$\lambda([[a,b],[c,d]]) = [[a,c],[b,c],[a,d],[b,d]]$$

Obviously not both can be true, and so list does not distribute over itself in this way.

Not all distributive laws resemble the distributivity of times over plus, so we cannot yet rule out the possibility of a distributive law for the list monad. In fact, Manes and Mulry found three further distributive laws for the *non-empty* list monad over itself [7, Example 5.1.9], [8, Example 4.10]. However, all of these fail to extend to distributive laws for the full list monad over itself. The theorems in this section will show that the search for a distributive law for the list monad over itself is a futile one; no such law exists.

We present several no-go theorems in this section. Whereas idempotence was the main property of interest in Section III, the focus now becomes unitality equations. In addition, the **abides** equation (above-besides, [17]) will be important:

$$(a * b) * (c * d) = (a * c) * (b * d) \tag{11}$$

We will be looking at theories in which this equation does *not* hold. This is made precise in property (T4)b below.

Throughout this section, we will consider two algebraic theories $\mathbb{S}$ and $\mathbb{T}$. For $\mathbb{S}$ we identify the following properties:

(S1) For any two terms s_1, s_2:

$$\emptyset \vdash s_1 \ \wedge\ \Gamma \vdash s_1 = s_2 \quad \Rightarrow \quad \emptyset \vdash s_2$$

(S2) For any term s and variable x:

$$\Gamma \vdash s = x \quad \Rightarrow \quad \{x\} \vdash s$$

(S3) For every n-ary operation ϕ ($n \geq 1$) in the signature $\Sigma^{\mathbb{S}}$, there is a constant e_ϕ, which acts as a unit for ϕ. If f is the substitution $x_i \mapsto e_\phi$ for all but one $i \leq n$, then:

$$\{x\} \vdash \phi[f] = x$$

(S4) $\mathbb{S}$ has a binary term $\mathbf{s}$ such that:

a) $e_{\mathbf{s}}$ is a unit for $\mathbf{s}$:

$$\{x\} \vdash \mathbf{s}(x, e_{\mathbf{s}}) = x = \mathbf{s}(e_{\mathbf{s}}, x)$$

b) $\mathbf{s}$ is idempotent:

$$\{x\} \vdash \mathbf{s}(x, x) = x$$

And for $\mathbb{T}$:

(T1) For any two terms t_1, t_2:

$$\emptyset \vdash t_1 \ \wedge\ \Gamma \vdash t_1 = t_2 \quad \Rightarrow \quad \emptyset \vdash t_2$$

(T2) For any term t and variable y:

$$\Gamma \vdash t = y \quad \Rightarrow \quad \{y\} \vdash t$$

(T3) $\mathbb{T}$ has a constant $e_{\mathbf{t}}$.

(T4) $\mathbb{T}$ has a binary term $\mathbf{t}$ such that:

a) $e_{\mathbf{t}}$ is a unit for $\mathbf{t}$:

$$\{y\} \vdash \mathbf{t}(y, e_{\mathbf{t}}) = y = \mathbf{t}(e_{\mathbf{t}}, y)$$

b) The abides equation does not hold in $\mathbb{T}$:

$$\Gamma \vdash \mathbf{t}(\mathbf{t}(y_1, y_2), \mathbf{t}(y_3, y_4)) = \mathbf{t}(\mathbf{t}(y_1, y_3), \mathbf{t}(y_2, y_4)) \\ \Rightarrow \ \#\Gamma \leq 3$$

Remark IV.1. *The properties (S1), (S2), (T1), (S2) are all constraints on the variables appearing in terms. The other properties are algebraic in nature, requiring constants to be units or terms to be idempotent.*

Example IV.2. *Algebraic properties of key monads:*

- *The list monad corresponds to the theory of monoids, which satisfies (S1) and (S2). The monoid unit satisfies (S3), and the monoid multiplication satisfies (S4)a, but the theory of monoids does not satisfy (S4)b. The equation $(x * y) * (z * w) = (x * z) * (y * w)$ holds in the theory of monoids if and only if $y = z$, and so it satisfies all of (T1), (T2), (T3), (T4)a, and (T4)b.*
- *The powerset monad is presented by the theory of join semilattices, which satisfies (S4)b in addition to (S1), (S2), (S3), and (S4)a. However, this theory does not have property (T4)b as the join is commutative and associative and so satisfies the abides equation* (11)*. Properties (T1), (T2), (T3), and (T4)a still hold.*
- *The exception monad corresponds to an algebraic theory with a signature containing constants for each exception, and no axioms. It satisfies (S1), (S2), and (S3). It does not satisfy (S4)a and (S4)b as there are no binary terms. Similarly, it satisfies (T1), (T2), and (T3), but not (T4)a or (T4)b.*

Given these properties, we will first consider when an $\mathbb{S}$-term of $\mathbb{T}$-terms is equal to the constant promised by property (T3), yielding Proposition IV.3. One application is that any putative distributive law of the list monad over itself must satisfy:

$$\lambda[L_1, \ldots, L_n] = [] \text{ if there is an } i \text{ s.t. } L_i = []$$

Furthermore, multiple constants immediately lead to a contradiction, and hence give us no-go Theorem IV.4.

Building on Proposition IV.3, we then derive conditions under which a distributive law has to behave like the distributivity of times over plus (1), resulting in Theorem IV.7. Again this theorem applies to the list monad, meaning there can be at most one distributive law for the list monad over itself.

Finally, in Theorems IV.10 and IV.15 we identify properties that together with Theorem IV.7 provide two more no-go theorems. The first property is a lack of the abides equation, which yields Theorem IV.10, from which we can conclude that there is no distributive law for the list monad over itself. The second requires an idempotent term in *one* of the algebraic theories, resulting in Theorem IV.15, which shows for example that there is no distributive law $PM \Rightarrow MP$ for the powerset monad P over the multiset monad M.

A. Multiplicative Zeroes

Our first focus is on properties (S3) and (T3). The goal is to prove that in a composite of theories $\mathbb{S}$ and $\mathbb{T}$, the constant $e_{\mathbf{t}}$ behaves like a multiplicative zero, consuming any $\mathbb{S}$-term it appears in.

Proposition IV.3. *Given algebraic theories $\mathbb{S}$ and $\mathbb{T}$ with properties (S3), and (T1) and (T3) respectively, let $\mathbb{U}$ be a composite of $\mathbb{T}$ after $\mathbb{S}$. Then for any $\mathbb{S}$-term $X \vdash s$ with at least one free variable, and any family of terms $u_x, x \in X$ for which there is an x_0 such that $x_0 \in \mathrm{var}(s)$ and $u_{x_0} = e_{\mathbf{t}}$:*

$$s[u_x/x] =_{\mathbb{U}} e_{\mathbf{t}}$$

Proof. We first prove the statement for any $\mathbb{S}$-term made of variables and a single operation symbol, where we substitute just one of the variables with $e_{\mathbf{t}}$. Then we generalize using substitutions and induction.

So let ϕ be any operation from the signature of $\mathbb{S}$ and let $x_0 \in \mathrm{var}(\phi)$. We will show that $\phi[e_{\mathbf{t}}/x_0] =_{\mathbb{U}} e_{\mathbf{t}}$.

The statement is trivial in the case that ϕ is a unary operation, because property (S3) implies that we must have $\phi(s) =_{\mathbb{S}} s$. We may therefore assume that ϕ has at least two free variables. From the fact that $\mathbb{U}$ is a composite of the theories $\mathbb{S}$ and $\mathbb{T}$, we know that every term in $\mathbb{U}$ is separated. And so, there is a $\mathbb{T}$-term $Y \vdash t$ and a family of $\mathbb{S}$-terms $s_y, y \in Y$ such that:

$$\phi[e_{\mathbf{t}}/x_0] =_{\mathbb{U}} t[s_y/y] \tag{12}$$

Since we assume property (S3), we know ϕ has unit e_ϕ. Consider the substitution $x \mapsto e_\phi$ for all $x \neq x_0$. Then:

$$\begin{aligned}
&\phi[e_{\mathbf{t}}/x_0] =_{\mathbb{U}} t[s_y/y] \\
\Rightarrow \quad & \{ \text{ Substitution } \} \\
&\phi[e_{\mathbf{t}}/x_0, e_\phi/x \neq x_0] =_{\mathbb{U}} t[s_y[e_\phi/x \neq x_0]/y] \\
\Leftrightarrow \quad & \{ \ e_\phi \text{is a unit for } \phi \ \} \\
&e_{\mathbf{t}} =_{\mathbb{U}} t[s_y[e_\phi/x \neq x_0]/y]
\end{aligned}$$

Because of essential uniqueness we conclude that there are functions $f : \emptyset \to Z, f' : Y \to Z$ such that:

$$e_{\mathbf{t}} =_{\mathbb{T}} t[f'(y)/y]$$

By assumption (T1) we may conclude that t does not have any variables and hence we have $e_{\mathbf{t}} =_{\mathbb{T}} t$. Going back to Equation (12), we see that:

$$\begin{aligned}
&\phi[e_{\mathbf{t}}/x_0] =_{\mathbb{U}} t[s_y/y] \\
\Leftrightarrow \quad & \{ \ \mathrm{var}(t) = \emptyset \ \} \\
&\phi[e_{\mathbf{t}}/x_0] =_{\mathbb{U}} t \\
\Leftrightarrow \quad & \{ \ t = e_{\mathbf{t}} \ \} \\
&\phi[e_{\mathbf{t}}/x_0] =_{\mathbb{U}} e_{\mathbf{t}}
\end{aligned}$$

The axiom of substitution in equational logic lets us generalize this result to the case where we substitute any terms u_x for every $x \neq x_0$:

$$\phi[e_{\mathbf{t}}/x_0, u_x/x \neq x_0] =_{\mathbb{U}} e_{\mathbf{t}}$$

Induction then proves the statement for any $\mathbb{S}$-term that has $e_{\mathbf{t}}$ substituted into one of its free variables. $\square$

If $\mathbb{T}$ has more than one constant, this could lead to inconsistencies. Our next no-go theorem makes this precise:

Theorem IV.4 (No-Go Theorem: Too Many Constants)**.** *Let $\mathbb{S}$ and $\mathbb{T}$ be algebraic theories with properties (S3) and (T1) respectively. Further assume that there is a term s in $\mathbb{S}$ with at least two free variables. Then if $\mathbb{T}$ has $n \geq 2$ constants, there exists no composite theory of $\mathbb{T}$ after $\mathbb{S}$.*

Proof. Suppose that $\mathbb{U}$ is a composite theory of $\mathbb{T}$ after $\mathbb{S}$ and let c_1 and c_2 be distinct constants in $\mathbb{T}$. Then by Proposition IV.3 we have:

$$c_1 =_{\mathbb{U}} s[c_1/x_1, c_2/x_2] =_{\mathbb{U}} c_2$$

By essential uniqueness, we may conclude that also $c_1 =_{\mathbb{T}} c_2$. Contradiction. So $\mathbb{U}$ cannot be a composite of $\mathbb{T}$ after $\mathbb{S}$. $\square$

By Theorem II.20, this also yields a no-go theorem for distributive laws. The usefulness of this result is illustrated by the following example.

Example IV.5 (An Error in the Literature)**.** *We have seen in Example IV.2 that the theory of monoids, corresponding to the list monad, satisfies (S3). It also has a term with two free variables, namely $x * y$, where $*$ is the binary operation in the signature of the theory. The exception monad satisfies (T1), so when the exception monad has more than one exception, Theorem IV.4 states that there is no distributive law $L \circ (- + E) \Rightarrow (- + E) \circ L$.*

However, [8, Example 4.12] claim to have a distributive law of this type for the case where $E = \{a, b\}$, given by:

$$\begin{aligned}
\lambda[] &= [] \\
\lambda[e] &= e \textit{ for any exception } e \in E \\
\lambda L &= L \textit{ if no element of L is in } E \\
\lambda L &= a \textit{ otherwise.}
\end{aligned}$$

We check more concretely that this cannot be a distributive law by showing that it fails the first multiplication axiom from Definition II.8:

$$\begin{tikzcd}
{[[b],[]]} \arrow[r, "L(\lambda_X)"] \arrow[d, "\mu^L_{EX}"] & {[b,[]]} \arrow[r, "\lambda_{LX}"] & a \arrow[d, "E(\mu^L_X)"] \\
{[b]} \arrow[rr, "\lambda_X"] & & b \neq a
\end{tikzcd}$$

The given distributive law seems to follow directly from [8, Theorem 4.6]. This would indicate that in its current form, the published theorem produces invalid distributive laws. We leave identifying the scope of the problem with this theorem to later work.

It is important to notice that Theorem IV.4 does not contradict the well-known result that the exception monad distributes over every set monad T; that result is for the other direction $(- + E) \circ T \Rightarrow T \circ (- + E)$.

Non-Example IV.6. *It is well known that the exception monad distributes over itself. Even though the corresponding theory satisfies properties (S3) and (T1), there are no terms with more than one free variable, and hence Theorem IV.4 does not apply.*

B. The One Distributive Law, If It Exists

Needing just the properties (S3), (T1), and (T3), Proposition IV.3 already greatly restricts the possibilities for a distributive law between monads S and T. We will now see that if both $\mathbb{S}$ and $\mathbb{T}$ have binary terms with units, then in a composite theory, the binary of $\mathbb{S}$ distributes over the binary of $\mathbb{T}$ like times over plus in Equation (1). For the monads corresponding to these theories, this means that there is only one candidate distributive law to consider.

Theorem IV.7 (Times over Plus Theorem)**.** *Let $\mathbb{S}$ and $\mathbb{T}$ be two algebraic theories satisfying (S1), (S2), (S3), and (S4)a and*

(T1), (T2), (T3), and (T4)a respectively. Let $\mathbb{U}$ be a composite theory of $\mathbb{T}$ after $\mathbb{S}$. Then $\mathbf{s}$ distributes over $\mathbf{t}$:

$$\mathbf{s}(\mathbf{t}(y_1, y_2), x_0) =_{\mathbb{U}} \mathbf{t}(\mathbf{s}(y_1, x_0), \mathbf{s}(y_2, x_0)) \tag{13}$$
$$\mathbf{s}(x_0, \mathbf{t}(y_1, y_2)) =_{\mathbb{U}} \mathbf{t}(\mathbf{s}(x_0, y_1), \mathbf{s}(x_0, y_2)) \tag{14}$$

Proof. We only have space for a rough sketch. We derive the distributional behaviour in three stages, relying heavily on the separation and essential uniqueness in a composite theory: every term can be written as a $\mathbb{T}$-term of $\mathbb{S}$-terms in an essentially unique way. If $\mathbf{s}$ is the binary in $\mathbb{S}$, and $\mathbf{t}$ the binary in $\mathbb{T}$, and $t_0[s_y/y]$ is the separated term equal to $\mathbf{s}(\mathbf{t}(y_1, y_2), x_0)$, then we derive the following about $t_0[s_y/y]$:

1) First we prove which variables appear in the terms s_y of the separated term: $\mathrm{var}(s_y) = \{y_1, x_0\}$ or $\mathrm{var}(s_y) = \{y_2, x_0\}$.
2) Then, we prove that each of the s_y is either equal to $\mathbf{s}(y_1, x_0)$ or to $\mathbf{s}(y_2, x_0)$.
3) Finally, we derive that the separated term $t_0[s_y/y]$ has to be equal to $\mathbf{t}(\mathbf{s}(y_1, x_0), \mathbf{s}(y_2, x_0))$.

□

In suitable cases, Theorem IV.7 reduces the search space for distributive laws to a single possibility. From Proposition II.21 we know that the action of distributive laws is determined by the separated terms in the composite theory. And so:

Corollary IV.8. *Let S and T be two monads presented by algebraic theories $\mathbb{S}$ and $\mathbb{T}$, having signatures with at least one constant and one binary operation. If for both theories the constant acts as a unit for the binary operation and the theories further satisfy (S1) and (S2), and (T1) and (T2) respectively, then any distributive law $ST \Rightarrow TS$ distributes the binary from $\mathbb{S}$ over the binary from $\mathbb{T}$ as in Equation* (1)*.*

Example IV.9 (Unique Distributive Laws)**.** *Let S be any of the monads tree*[1]*, list, or multiset. Then the corresponding algebraic theory $\mathbb{S}$ contains only linear equations. Let T be either the multiset or powerset monad. Since the multiset and powerset monads are commutative monads, it follows from [7, Theorem 4.3.4] that there is a distributive law $ST \Rightarrow TS$. Corollary IV.8 states that this distributive law is unique. In particular, the distributive law for the multiset monad over itself mentioned in Example II.12 is unique.*

C. Lacking the Abides Property: a No-Go Theorem

With Theorem IV.7 narrowing down the possible distributive laws for two monads, it is easier to find cases in which no distributive law can exist at all. We identify two properties that clash with Theorem IV.7, one for $\mathbb{T}$ and one for $\mathbb{S}$. In this section we show that not satisfying the abides equation, property (T4)b, in combination with Theorem IV.7 prevents the existence of a distributive law. In the next section we do the same for idempotence, property (S4)b. Both properties are sufficiently common to cover a broad class of monads.

[1]respresented by the theory of a single binary operation and a constant, satisfying only left and right unitality.

Theorem IV.10 (No-Go Theorem: Lacking Abides)**.** *If $\mathbb{S}$ and $\mathbb{T}$ are algebraic theories satisfying the conditions of Theorem IV.7 and $\mathbb{T}$ additionally satisfies (T4)b, then there does not exist a composite theory of $\mathbb{T}$ after $\mathbb{S}$.*

Proof. Suppose there exists a composite theory $\mathbb{U}$. Given Theorem IV.7, there are two ways to compute a separated term equal in $\mathbb{U}$ to $\mathbf{s}(\mathbf{t}(y_1, y_2), \mathbf{t}(y_3, y_4))$:

- The first is to view $\mathbf{s}(\mathbf{t}(y_1, y_2), \mathbf{t}(y_3, y_4))$ as $\mathbf{s}(\mathbf{t}(y_1, y_2), x_0)[\mathbf{t}(y_3, y_4)/x_0]$. Then apply Theorem IV.7 to $\mathbf{s}(\mathbf{t}(y_1, y_2), x_0)$, substitute $\mathbf{t}(y_3, y_4)$ back in and apply Theorem IV.7 again to get the final result. This yields:

$$\begin{aligned}&\mathbf{s}(\mathbf{t}(y_1, y_2), \mathbf{t}(y_3, y_4)) =_{\mathbb{U}} \\ &\mathbf{t}(\mathbf{t}(\mathbf{s}(y_1, y_3), \mathbf{s}(y_1, y_4)), \mathbf{t}(\mathbf{s}(y_2, y_3), \mathbf{s}(y_2, y_4)))\end{aligned} \tag{15}$$

- The second way is to view $\mathbf{s}(\mathbf{t}(y_1, y_2), \mathbf{t}(y_3, y_4))$ as $\mathbf{s}(x_0, \mathbf{t}(y_3, y_4))[\mathbf{t}(y_1, y_2)/x_0]$. Now apply Theorem IV.7 to $\mathbf{s}(x_0, \mathbf{t}(y_3, y_4))$, substitute $\mathbf{t}(y_1, y_2)$ back in and apply Theorem IV.7 again. This yields:

$$\begin{aligned}&\mathbf{s}(\mathbf{t}(y_1, y_2), \mathbf{t}(y_3, y_4)) =_{\mathbb{U}} \\ &\mathbf{t}(\mathbf{t}(\mathbf{s}(y_1, y_3), \mathbf{s}(y_2, y_3)), \mathbf{t}(\mathbf{s}(y_1, y_4), \mathbf{s}(y_2, y_4)))\end{aligned} \tag{16}$$

Of course, both computations are equally valid, so the terms in Equations (15) and (16) must be equal:

$$\begin{aligned}&\mathbf{t}(\mathbf{t}(\mathbf{s}(y_1, y_3), \mathbf{s}(y_1, y_4)), \mathbf{t}(\mathbf{s}(y_2, y_3), \mathbf{s}(y_2, y_4))) \\ =_{\mathbb{U}}\, &\mathbf{t}(\mathbf{t}(\mathbf{s}(y_1, y_3), \mathbf{s}(y_2, y_3)), \mathbf{t}(\mathbf{s}(y_1, y_4), \mathbf{s}(y_2, y_4)))\end{aligned}$$

Since these are two separated terms that are equal, we can apply the essential uniqueness property. This leads to the conclusion that there are distinct variables z_1, z_2, z_3, z_4 such that:

$$\mathbf{t}(\mathbf{t}(z_1, z_2), \mathbf{t}(z_3, z_4)) =_{\mathbb{T}} \mathbf{t}(\mathbf{t}(z_1, z_3), \mathbf{t}(z_2, z_4)) \tag{17}$$

But this contradicts assumption (T4)b, and hence no such composite theory exists. □

Corollary IV.11. *If monads S and T are presented by algebraic theories $\mathbb{S}$ and $\mathbb{T}$, satisfying the axioms of Theorem IV.10, then there does not exist a distributive law $ST \Rightarrow TS$.*

Example IV.12 (Resolving an Open Question)**.** *This finally settles the question of whether the list monad distributes over itself, posed by [7], [8]. The theory of monoids satisfies all the conditions required of both theories in Theorem IV.10, and hence there is no distributive law for the list monad over itself.*

Note that a distributive law for lists was claimed in [14], although it was subsequently shown to be incorrect in [26].

Remark IV.13. *Although there is no distributive law for the list monad over itself, the functor LL does still carry a monad structure. We are very grateful to Bartek Klin for pointing this out to us. The monad structure on LL can be described as follows:*

- *There is a distributive law for the list monad over the non-empty list monad $LL^+ \Rightarrow L^+L$ [7].*

- *There is a distributive law for the resulting monad over the maybe monad $(L^+L)(-)_\perp \Rightarrow (-)_\perp(L^+L)$, derived from general principles [7].*
- *The resulting functor $(-)_\perp(L^+L)$ is isomorphic to LL, and carries a monad structure. Hence LL carries a monad structure, but not one that can be derived from a distributive law $LL \Rightarrow LL$.*

Non-Example IV.14 (Multiset Monad)**.** *The multiset monad is closely related to the list monad, with an algebraic theory having just one extra equation compared to the list monad: commutativity. Because of this equation, the theory does not have property (T4)b. As we have seen in Example IV.9, there is a unique distributive law for the multiset monad over itself.*

D. Yet Another No-Go Theorem Caused by Idempotence

In Section III we learned that having idempotent terms in both algebraic theories could preclude the existence of a distributive law. We will now see that together with Theorem IV.7, having an idempotent term in $\mathbb{S}$ is enough.

Theorem IV.15 (No-Go Theorem: Idempotence and Units)**.** *Let $\mathbb{S}$ and $\mathbb{T}$ be algebraic theories satisfying (S1), (S2), (S3), (S4)a, and (S4)b and (T1), (T2), (T3), and (T4)a respectively. Then there exists no composite theory of $\mathbb{T}$ after $\mathbb{S}$.*

Proof. (Sketch.) Using idempotence of $\mathbf{s}$ and Theorem IV.7 we can equate:

$$\mathbf{t}(y_1, y_2) =_{\mathbb{U}} \mathbf{t}(\mathbf{t}(y_1, \mathbf{s}(y_1, y_2)), \mathbf{t}(\mathbf{s}(y_2, y_1), y_2))$$

We can then apply essential uniqueness to retrieve separate equalities of terms in $\mathbb{T}$ and $\mathbb{S}$. Further substitutions, using the constant $e_{\mathbf{t}}$ as a unit for $\mathbf{t}$ and exploiting property (T1) yield information about the variables appearing in the $\mathbb{T}$-terms. Via essential uniqueness again, this allows us to conclude:

$$\mathbf{s}(y_1, y_2) =_{\mathbb{S}} y_1 \quad \text{or} \quad \mathbf{s}(y_1, y_2) =_{\mathbb{S}} y_2$$

Both equations contradict (S2). Therefore, the composite theory $\mathbb{U}$ cannot exist. □

Corollary IV.16. *If monads S and T are presented by algebraic theories $\mathbb{S}$ and $\mathbb{T}$, satisfying the axioms of Theorem IV.15, then there does not exist a distributive law $ST \Rightarrow TS$.*

Example IV.17 (Powerset over Itself Again)**.** *The theory of join semilattices satisfies all the axioms required of both theories in theorem IV.15. Therefore, there is no distributive law for the powerset monad over itself. This was already shown by [11] using similar methods as in Section III. Theorem IV.15 gives a second, independent proof of this fact.*

Example IV.18 (Filling in the Gap: Multiset and Powerset)**.** *From [7] we know that there are distributive laws $MM \Rightarrow MM$ and $MP \Rightarrow PM$, where M is the multiset monad and P the powerset monad. So the only combination of multiset and powerset that is not covered by previous theorems is $PM \Rightarrow MP$. Theorem IV.15 fills this gap, saying there is no distributive law of that type.*

Non-Example IV.19 (The Sweet Spot)**.** *We come back to the multiset monad. In Counterexample IV.14 we saw that the algebraic theory presenting the multiset monad had one extra equation compared to the theory for the list monad: commutativity. Because of this equation, property (T4)b did not hold, and therefore Theorem IV.10 did not apply.*

There is a similar relation between the multiset monad and the powerset monad. Compared to the powerset monad, the theory corresponding to the multiset monad lacks *just one equation: idempotence, which is exactly what property (S4)b requires. The lack of this equation in the theory for the multiset monad therefore means that Theorem IV.15 does not apply to multiset either. So multiset holds a sort of 'sweet spot' in between the two no-go theorems, where a distributive law $MM \Rightarrow MM$ still can and does exist.*

V. Conclusion

We have shown there can be no distributive law between large classes of monads:

- Section III developed general theorems for demonstrating when distributive laws cannot exist, derived from a classical counterexample of Plotkin.
- Section IV went further, showing a new approach, yielding counterexamples beyond those possible in Section III.

Our results cover many naturally occurring combinations of monads, including all known negative results. They also identify issues in the existing literature, and resolve the open question of whether the list monad distributes over itself.

We strongly advocated the use of algebraic methods. These techniques were used for all of our proofs. Taking this approach, rather than direct calculations involving Beck's axioms, enabled us to single out the essentials of each proof, so that the resulting theorems could be stated in full generality.

Lastly, we would like to emphasize that the methods described in this paper are of broader application, beyond our specific theorems. For example, Julian Salamanca has shown that there is no distributive law of the group monad over the powerset monad $GP \Rightarrow PG$ [27]. Although elements of his proof are very similar to those used in section Section III, the proof itself uses further ideas outside the scope of our theorems.

Acknowledgements

We are very grateful to Jeremy Gibbons, Bartek Klin, Hector Miller-Bakewell, Julian Salamanca and our reviewers for their insightful feedback on earlier versions of this paper. We also would like to thank Prakash Panangaden, Fredrik Dahlqvist, Louis Parlant, Sam Staton, Ohad Kammar, Jules Hedges, and Ralf Hinze for inspiring discussions, which all had a positive influence on this paper.

This work was partially supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (No.2015-0-00565, Development of Vulnerability Discovery Technologies for IoT Software Security).

References

[1] E. Moggi, "Notions of computation and monads," *Information and Computation*, vol. 93, no. 1, pp. 55 – 92, 1991, selections from 1989 IEEE Symposium on Logic in Computer Science.

[2] P. Wadler, "Monads for functional programming," in *Advanced Functional Programming*, J. Jeuring and E. Meijer, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995, pp. 24–52.

[3] S. Peyton Jones, *Tackling the Awkward Squad: monadic input/output, concurrency, exceptions, and foreign-language calls in Haskell*. Amsterdam: IOS press, 2001, pp. 47–96.

[4] P. Hudak, S. Peyton Jones, P. Wadler, B. Boutel, J. Fairbairn, J. Fasel, M. M. Guzmán, K. Hammond, J. Hughes, T. Johnsson *et al.*, "Report on the programming language Haskell: a non-strict, purely functional language version 1.2," *ACM SigPlan notices*, vol. 27, no. 5, pp. 1–164, 1992.

[5] J. Beck, "Distributive laws," in *Seminar on Triples and Categorical Homology Theory*, B. Eckmann, Ed. Springer Berlin Heidelberg, 1969, pp. 119–140.

[6] M. Lenisa, J. Power, and H. Watanabe, "Distributivity for endofunctors, pointed and co-pointed endofunctors, monads and comonads," *Electronic Notes in Theoretical Computer Science*, vol. 33, pp. 230–260, 2000.

[7] E. Manes and P. Mulry, "Monad compositions I: general constructions and recursive distributive laws," *Theory and Applications of Categories*, vol. 18, pp. 172–208, 04 2007.

[8] ——, "Monad compositions II: Kleisli strength," *Mathematical Structures in Computer Science*, vol. 18, no. 3, p. 613643, 2008.

[9] M. M. Bonsangue, H. H. Hansen, A. Kurz, and J. Rot, "Presenting distributive laws," in *Algebra and Coalgebra in Computer Science*, R. Heckel and S. Milius, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 95–109.

[10] F. Dahlqvist, L. Parlant, and A. Silva, "Layer by layer - combining monads," in *Theoretical Aspects of Computing - ICTAC 2018 - 15th International Colloquium, Stellenbosch, South Africa, October 16-19, 2018, Proceedings*, ser. Lecture Notes in Computer Science, B. Fischer and T. Uustalu, Eds., vol. 11187. Springer, 2018, pp. 153–172. [Online]. Available: https://doi.org/10.1007/978-3-030-02508-3_9

[11] B. Klin and J. Salamanca, "Iterated covariant powerset is not a monad," in *Proceedings 34th Conference on the Mathematical Foundations of Programming Semantics, MFPS 2018*, 2018.

[12] D. Varacca and G. Winskel, "Distributing probability over non-determinism," *Mathematical Structures in Computer Science*, vol. 16, no. 1, pp. 87–113, 2006.

[13] F. Dahlqvist and R. Neves, "Compositional semantics for new paradigms: Probabilistic, hybrid and beyond," *CoRR*, vol. abs/1804.04145, 2018. [Online]. Available: http://arxiv.org/abs/1804.04145

[14] D. J. King and P. Wadler, "Combining monads," in *Functional Programming, Glasgow 1992*. Springer, 1993, pp. 134–143.

[15] M. Pirog and S. Staton, "Backtracking with cut via a distributive law and left-zero monoids," *Journal of Functional Programming*, vol. 27, 2017.

[16] E. Cheng, "Distributive laws for Lawvere theories," *CoRR*, vol. abs/1112.3076v2, 2011. [Online]. Available: https://arxiv.org/abs/1112.3076v2

[17] R. Bird, *Lectures on Constructive Functional Programming*. Oxford University Computing Laboratory Programming Research Group, 1988.

[18] F. E. Linton, "Some aspects of equational categories," in *Proceedings of the Conference on Categorical Algebra*. Springer, 1966, pp. 84–94.

[19] F. W. Lawvere, "Functorial semantics of algebraic theories," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 50, no. 5, pp. 869–872, 1963.

[20] E. G. Manes, *Algebraic theories*. Springer, 1976, vol. 26.

[21] G. Plotkin and J. Power, "Notions of computation determine monads," in *International Conference on Foundations of Software Science and Computation Structures*. Springer, 2002, pp. 342–356.

[22] B. Jacobs, "Convexity, duality and effects," in *IFIP International Conference on Theoretical Computer Science*. Springer, 2010, pp. 1–19.

[23] M. H. Stone, "Postulates for the barycentric calculus," *Annali di Matematica Pura ed Applicata*, vol. 29, no. 1, pp. 25–30, 1949.

[24] D. Varacca, "Probability, nondeterminism and concurrency: Two denotational models for probabilistic computation," Ph.D. dissertation, BRICS, Department of Computer Science, University of Aarhus, November 2003.

[25] P. Panangaden, "Private communication," 2018.

[26] M. P. Jones and L. Duponcheel, "Composing monads," Technical Report YALEU/DCS/RR-1004, Department of Computer Science. Yale , Tech. Rep., 1993.

[27] J. Salamanca, "Private communication," 2018.