Correspondence between Composite Theories and Distributive Laws

Aloïs Rosset[1], Maaike Zwart[2],
Helle Hvid Hansen[3], and Jörg Endrullis[1]

[1] Vrije Universiteit Amsterdam, Amsterdam, Netherlands
`{a.rosset, j.endrullis}@vu.nl`
[2] IT University of Copenhagen, Copenhagen, Denmark
`maaike.annebeth@gmail.com`
[3] University of Groningen, Groningen, Netherlands
`h.h.hansen@rug.nl`

Abstract. Composite theories are the algebraic equivalent of distributive laws. In this paper, we delve into the details of this correspondence and concretely show how to construct a composite theory from a distributive law and vice versa. Using term rewriting methods, we also describe when a minimal set of equations axiomatises the composite theory.

Keywords: monad · distributive law · algebraic theory · composite theory · term rewriting

1 Introduction

Monads are categorical structures [4, 20] with many applications in (co)algebraic approaches to program semantics, notably to model effects such as nondeterminism, probabilities and exceptions [24, 27, 6, 17]. Monads that occur in the specification of programs and are used in reasoning about programs are often finitary and Set-based, and hence can be presented as algebraic theories [7, 21, 1].

The algebraic view on monads has been especially useful when studying monad compositions [8, 14, 25, 26, 36]. Composing monads is a way to combine multiple computational effects, and is usually done categorically via a distributive law [5, 22]. However, the required distributive laws do not always exist, and the use of algebraic theories was instrumental in proving so-called no-go theorems, which tell us when two finitary monads cannot be composed via a distributive law [36].

Central to these results is the correspondence between composites of algebraic theories, and distributive laws between the corresponding monads. Briefly stated, a composite of two algebraic theories $\mathbb{S}$ and $\mathbb{T}$ is a theory $\mathbb{U}$ that contains all the function symbols and equations of $\mathbb{S}$ and $\mathbb{T}$ as well as a set of distribution axioms that specify how equality of mixed terms can be reduced to equality in $\mathbb{S}$ and $\mathbb{T}$. Composite theories were originally studied by Cheng [8] on the abstract level of Lawvere theories. Piróg & Staton [26] formulated them in the more concrete setting of algebraic theories.

While Piróg & Staton state the correspondence between composite theories and distributive laws, they do not provide a proof, referring instead to Cheng. In her thesis, Zwart [35] gives a constructive version of this correspondence for the category Set, but she does not prove directly that the algebraic theory she constructs from a distributive law is indeed a composite theory.

Furthermore, the theory Zwart constructs is given via a set E_λ that contains all possible equations with interaction between the theories $\mathbb{S}$ and $\mathbb{T}$. While this axiomatisation does the job, it is neither elegant nor practical to work with. Composite theories can often be described in terms of a few simple distribution axioms. A classic example is the theory of rings, which is a composite of the theories of monoids and Abelian groups via the two 'times over plus' distribution axioms. A systematic approach to identify such a minimal set of distribution axioms for a composite theory would be far more practical than the set E_λ.

In this paper, we present a full and self-contained proof of the correspondence between composite theories $\mathbb{U}$ (of $\mathbb{T}$ after $\mathbb{S}$) and distributive laws $\lambda\colon ST \to TS$, where $\mathbb{S}$ and $\mathbb{T}$ are algebraic theories and S, T are their corresponding finitary Set-monads. Section 4 shows how to get a distributive law from a composite theory, and Section 5 shows how to construct a composite theory from a distributive law. The proof of the latter uses term rewriting techniques. In particular, we introduce *functor rewriting systems* in order to reason about strings of functors, and to obtain a separation of $\mathbb{U}$-terms.

In addition, in Section 6 we give criteria that ensure that a certain minimal set of distribution axioms $E' \subseteq E_\lambda$ suffices to axiomatise $\mathbb{U}$. The natural candidate for E' consists of equations in which the left-hand side is a term consisting of exactly one $\mathbb{S}$-operation symbol, which has exactly one $\mathbb{T}$-operation symbol among its arguments. We prove that if a term rewriting system corresponding to E' is terminating, then $E_{\mathbb{S}} \cup E_{\mathbb{T}} \cup E'$ axiomatises $\mathbb{U}$. To illustrate that this criterion is not trivially satisfied, we give an example in which E' does not terminate and indeed does not axiomatise $\mathbb{U}$. Finally, we show that we have termination if the right-hand sides of the equations in E' are of a certain form, and apply our results to establish presentations of some composite monads/theories.

2 Preliminaries

We assume that the reader is familiar with basic notions of category theory [3, 20, 28]. This section recalls basic definitions and results concerning monads, algebraic theories, and term rewriting systems, and fixes notation for the concepts we use in this paper.

2.1 Monads

Definition 1. *A* **monad** (M, η, μ) *on a category* C *is a triple consisting of an endofunctor* $M : \mathsf{C} \to \mathsf{C}$*, and two natural transformations, the* **unit** $\eta : \mathrm{id} \Rightarrow M$ *and the* **multiplication** $\mu : M^2 \Rightarrow M$ *that make* (1) *and* (2) *commute. For convenience, we often refer to a monad* (M, η, μ) *by its functor part* M*.*

$$\begin{tikzcd}
M \arrow[r, "M\eta"] \arrow[dr, equal] & M^2 \arrow[d, "\mu"] & M \arrow[l, "\eta M"'] \arrow[dl, equal] \\
& M &
\end{tikzcd} \tag{1}$$

$$\begin{tikzcd}
M^3 \arrow[r, "\mu M"] \arrow[d, "M\mu"'] & M^2 \arrow[d, "\mu"] \\
M^2 \arrow[r, "\mu"'] & M
\end{tikzcd} \tag{2}$$

Example 2. Here are some examples of Set-monads, where we always mean the finitary versions. For more details on these monads, see e.g. [13, §1.2.1].

- The *list* and *non-empty list* monads L and L^+, with $\eta^L_X(x) = \eta^{L^+}_X(x) = [x]$, and $\mu^L = \mu^{L^+}$ being concatenation.
- The *multiset* monad $\mathcal{M}$, with $\eta^{\mathcal{M}}(x) = \lbag x \rbag$ and $\mu^{\mathcal{M}}$ taking the union, adding multiplicities. Taking multiplicities in $\mathbb{Z}$ gives the *Abelian group* monad $\mathcal{A}$.
- The *distribution* monad $\mathcal{D}$, with $\eta^{\mathcal{D}}(x) = 1x$ and a weighted average of $\mu^{\mathcal{D}}$.
- The *reader* monad $R_A(X) = X^A$, where A is a finite set, with η^R the constant function and μ^R reading the same element twice.

Definition 3. *Given two monads (M, η^M, μ^M) and (T, η^T, μ^T) on a category C, a* **monad morphism** *from M to T is a natural transformation $\theta : M \Rightarrow T$ that makes (3) and (4) commute, , where $\theta\theta := \theta_T \cdot M\theta = T\theta \cdot \theta_M$ (called* horizontal composition*). If each component of θ is an isomorphism, we say that the two monads are* **isomorphic***.*

$$\begin{tikzcd}
& M \arrow[dd, "\theta"] \\
\mathrm{id} \arrow[ur, "\eta^M"] \arrow[dr, "\eta^T"'] & \\
& T
\end{tikzcd} \tag{3}$$

$$\begin{tikzcd}
M^2 \arrow[r, "\theta\theta"] \arrow[d, "\mu^M"'] & T^2 \arrow[d, "\mu^T"] \\
M \arrow[r, "\theta"'] & T
\end{tikzcd} \tag{4}$$

Definition 4. *Let (M, η, μ) be a monad on category C. An (Eilenberg-Moore) M-***algebra** *is a C-morphism $\alpha : MX \to X$ for some $X \in \mathsf{C}$, denoted (X, α) for short, such that (5) and (6) commute. An M-algebra* **homomorphism** *$f : (X, \alpha) \to (Y, \beta)$ between two M-algebras is a function $f : X \to Y$ such that (7) commutes. The category of M-algebras and M-algebra homomorphisms is denoted* $\mathbf{EM}(M)$ *and called the* **Eilenberg-Moore category** *of M.*

$$\begin{tikzcd}
X \arrow[r, "\eta_X"] \arrow[dr, equal] & MX \arrow[d, "\alpha"] \\
& X
\end{tikzcd} \tag{5}$$

$$\begin{tikzcd}
M^2X \arrow[r, "\mu_X"] \arrow[d, "M\alpha"'] & MX \arrow[d, "\alpha"] \\
MX \arrow[r, "\alpha"'] & X
\end{tikzcd} \tag{6}$$

$$\begin{tikzcd}
MX \arrow[r, "Mf"] \arrow[d, "\alpha"'] & MY \arrow[d, "\beta"] \\
X \arrow[r, "f"'] & Y
\end{tikzcd} \tag{7}$$

Definition 5. *Let S, T be monads. A* **distributive law** *$\lambda : ST \Rightarrow TS$ between monads is a natural transformation satisfying (8)-(11). A* **weak distributive law** *$\lambda : ST \Rightarrow TS$ is a natural transformation satisfying (9)-(11).*

$$\begin{tikzcd}
& T \arrow[dl, "\eta^S T"'] \arrow[dr, "T\eta^S"] & \\
ST \arrow[rr, "\lambda"] & & TS
\end{tikzcd} \tag{8}$$

$$\begin{tikzcd}
& S \arrow[dl, "S\eta^T"'] \arrow[dr, "\eta^T S"] & \\
ST \arrow[rr, "\lambda"] & & TS
\end{tikzcd} \tag{9}$$

$$\begin{tikzcd}
SST \arrow[r, "S\lambda"] \arrow[d, "\mu^S T"] & STS \arrow[r, "\lambda S"] & TSS \arrow[d, "T\mu^S"'] \\
ST \arrow[rr, "\lambda"] & & TS
\end{tikzcd} \tag{10}$$

$$\begin{tikzcd}
STT \arrow[r, "\lambda T"] \arrow[d, "S\mu^T"] & TST \arrow[r, "T\lambda"] & TTS \arrow[d, "\mu^T S"'] \\
ST \arrow[rr, "\lambda"] & & TS
\end{tikzcd} \tag{11}$$

A distributive law $\lambda : ST \to TS$ induces a monad structure on the functor TS as follows [5, §1]:

$$\Big(TS,\ \ \eta^{TS} := \big(\mathrm{id} \xrightarrow{\eta^T \eta^S} TS\big),\ \ \mu^{TS} := \big(TSTS \xrightarrow{T\lambda S} TTSS \xrightarrow{\mu^T \mu^S} TS\big)\Big) \tag{12}$$

The algebras for this composite monad are algebras that are simultaneously S-algebras and T-algebras. This is visible through the isomorphism $\mathbf{EM}(TS) \cong \mathbf{Alg}(\lambda)$ [5, §2], where the category $\mathbf{Alg}(\lambda)$ of λ-algebras is defined as follows:

Definition 6. *Given monads S, T and distributive law $\lambda : ST \to TS$, then the objects of the category $\mathbf{Alg}(\lambda)$ are triples (X, σ, τ), such that (X, σ) is an S-algebra and (X, τ) is a T-algebra, and the diagram on the right commutes. The morphisms of $\mathbf{Alg}(\lambda)$ are C-morphisms that are both S- and T-algebra homomorphisms.*

$$\begin{tikzcd}
STX \arrow[rr, "\lambda"] \arrow[d, "S\tau"'] & & TSX \arrow[d, "T\sigma"] \\
SX \arrow[dr, "\sigma"'] & & TX \arrow[dl, "\tau"] \\
& X &
\end{tikzcd}$$

2.2 Algebraic Theories

Definition 7. *An* **algebraic theory** *is a pair (Σ, E) consisting of an algebraic signature Σ and set of equations E over Σ defined as follows.*

- *An* **algebraic signature** *Σ is a set of operation symbols. Each $\mathsf{op}^{(n)} \in \Sigma$ has an arity $n \in \mathbb{N}$.*
- *The set $\mathcal{T}(\Sigma, X)$, also denoted $\Sigma^* X$, of Σ-***terms** *over a set X is defined inductively: elements in X are terms, and given terms $t_1, \ldots, t_n$ and $\mathsf{op}^{(n)} \in \Sigma$, then $\mathsf{op}(t_1, \ldots, t_n)$ is a term.*
- *An* **equation** *over a signature Σ is a pair (s, t) of Σ-terms.*

For the rest of this paper, we fix a set $\mathcal{V} = \{v_1, v_2, v_3, \ldots\}$ of variables. The subset of $\mathcal{V}$ appearing in a term t is denoted as $\mathsf{var}(t)$. Functions of the form $\upsilon : \mathcal{V} \to Y$ are called ***variable assignments***.

Notation 8. In this paper we make heavy use of substitutions. For readability, we pick from the following notations for substitutions, depending on context. Given terms $t(x_1, .., x_n)$ and $s_1, \ldots, s_n$, and variable assignment $h : \mathcal{V} \to \mathcal{T}(\Sigma, \mathcal{V})$ defined as $x_1 \mapsto s_{x_1}, \ldots, x_n \mapsto s_{x_n}$ and identity elsewhere, we denote the term t where each x_i is substituted with s_i (for $i = 1, \ldots, n$) by either $t[h]$, $t[s_1, ..., s_n]$, or $t[s_x/x]$ (or even $t[s_x]$) for short, where x ranges over all variables in t. Moreover, given a family of terms $(t_x[s_{x,y}/y])_{x \in X}$, we will simply write each term $t_x[s_y]$, as we can assume that each t_x has distinct variables by choosing the (say m) variables of t_{x_1} to be $y_1, \ldots, y_m$, the variables of t_{x_2} to start at y_{m+1}, and so on.

Definition 9. *The category $\mathbf{Alg}(\Sigma, E)$ consists of (Σ, E)-algebras and homomorphisms between them.*

- *A Σ-***algebra** *is a pair $(X, [\![\cdot]\!])$ consisting of a set X and a collection of interpretations: for each $\mathsf{op}^{(n)} \in \Sigma$, we have $[\![\mathsf{op}]\!] : X^n \to X$. Any function $f : X \to Y$ extends to a unique homomorphism, $[\![\cdot]\!]_f : \mathcal{T}(\Sigma, X) \to Y$,* *as given by equations* (13) *and* (14) *below. When* $f = \mathrm{id}_X$, *we omit the subscript.*

$$[\![x]\!]_f := f(x), \text{ and} \tag{13}$$

$$[\![\mathsf{op}(t_1,\dots,t_n)]\!]_f := [\![\mathsf{op}]\!]([\![t_1]\!]_f,\dots,[\![t_n]\!]_f). \tag{14}$$

- *A* (Σ,E)**-algebra** $(X,[\![\cdot]\!])$ *is a* Σ*-algebra whose* $[\![\cdot]\!]$ *satisfies all equations in* E, *i.e., for each* $(s,t)\in E$ *and all variable assignments* υ, $[\![s]\!]_\upsilon = [\![t]\!]_\upsilon$.
- *A* (Σ,E)*-algebra* **homomorphism** $f : (X,[\![\cdot]\!]) \to (X',[\![\cdot]\!]')$ *is a function* $f : X \to X'$ *such that* $f[\![\mathsf{op}]\!] = [\![\mathsf{op}]\!]' f^n$, *for all* $\mathsf{op}^{(n)} \in \Sigma$.

Given an algebraic theory $\mathbb{T} = (\Sigma_{\mathbb{T}}, E_{\mathbb{T}})$ and $\Sigma_{\mathbb{T}}$-terms s and t, we write $s =_{\mathbb{T}} t$ to denote that the equality $s = t$ is derivable from the axioms $E_{\mathbb{T}}$ in equational logic. The inference rules of equational logic are in [30, §8.1].

Definition 10. *There is a free-forgetful adjunction* $F\colon \mathsf{Set} \underset{\perp}{\rightleftarrows} \mathbf{Alg}(\Sigma,E) : U$.

- *The* **free** (Σ,E)**-algebra** *on set* X *is the* (Σ,E)*-algebra* $({}^{\mathcal{T}(\Sigma,X)}/_{=_{(\Sigma,E)}}, [\![\cdot]\!])$ *with carrier* $\mathcal{T}(\Sigma,X)$ *modulo* $=_{(\Sigma,E)}$. *The equivalence class of a term* t *is denoted* $\overline{t}^{(\Sigma,E)}$ *or* $\overline{t}$ *if the theory is clear from context. The interpretation of* $\mathsf{op}^{(n)} \in \Sigma_{\mathbb{T}}$ *is* $[\![\mathsf{op}]\!](\overline{t_1},\dots,\overline{t_n}) := \overline{\mathsf{op}(t_1,\dots,t_n)}$.
- *The* **free functor** $F : \mathsf{Set} \to \mathbf{Alg}(\Sigma,E)$ *sends* X *to its free* (Σ,E)*-algebra, and any function* $f : X \to Y$ *to* $Ff : FX \to FY$ *defined by* $Ff(\overline{t}) := \overline{t[f]}$.

The fact that F *is a well-defined functor is well-known and an account of it is provided in the extended version of the paper [30]. Composing the adjoint functors gives a monad* $(T := UF, \eta, \mu)$, *called the* **free algebra monad** *[20, VI.1]. The unit is* $\eta : x \mapsto \overline{x}$ *and the multiplication is* $\mu : \overline{t[\overline{t_i}/v_i]} \mapsto \overline{t[t_i/v_i]}$.

Definition 11 ([29, Def. 5, Lem. 8]). *An algebraic theory* (Σ,E) *is an* **algebraic presentation** *of a* Set*-monad* (M,η^M,μ^M) *if we have an isomorphism of monads* $(T,\eta^T,\mu^T) \cong (M,\eta^M,\mu^M)$, *where* T *is the free algebra monad of* (Σ,E). *An equivalent formulation is that both categories of algebras are concretely isomorphic*[4]*:* $\mathbf{EM}(M) \cong_{\mathsf{conc}} \mathbf{Alg}(\Sigma,E)$. *The former isomorphism relates the monads on a syntactic level, whereas the latter relates them semantically.*

Note that a monad can have multiple presentations.

Example 12. Here are algebraic presentations of the monads from Example 2.

- The *list* monad L is presented by the theory of *monoids*.
- The *non-empty list* monad L^+ is presented by the theory of *semigroups*.
- The *multiset* monad $\mathcal{M}$ is presented by the theory of *commutative monoids*.
- The *Abelian group* monad $\mathcal{A}$ is presented by the theory of *Abelian groups*.
- The *distribution* monad $\mathcal{D}$ is presented by the theory of *convex algebras* [15].
- The *reader* monad R_A is presented by the theory of *local states* [27] consisting of a single $|A|$-ary operation symbol, satisfying idempotence and diagonal equations (e.g. in the case $|A| = 2$: $a * a = a$ and $(a * b) * (c * d) = (a * d)$).

---

[4] “concrete” means that both functors of this isomorphism commute with the forgetful functors $\mathbf{EM}(M) \to \mathsf{Set}$ and $\mathbf{Alg}(\Sigma,E) \to \mathsf{Set}$. In other words it sends an M-algebra $(X, x : MX \to X)$ to a (Σ,E)-algebra with same carrier $(X,[\![\cdot]\!])$ and vice-versa.

2.3 Term Rewriting Systems

We only briefly explain the basic concepts and results of term rewriting systems (TRS) that we need in our proofs. For more background, we recommend the book "Term Rewriting Systems" by Terese [32].

Definition 13. *Given a signature Σ, a* **rewrite rule** $(l \to r)$ *is a pair of Σ-terms (l, r) such that l is not a variable, and all variables in the right occur also in the left:* $\mathsf{var}(l) \supseteq \mathsf{var}(r)$. *A* **term rewriting system** $\mathcal{R} = (\Sigma, R)$ *consists of a signature Σ and a set of rewrite rules R. The rewrite relation $\to_{\mathcal{R}}$ is the smallest relation on $\mathcal{T}(\Sigma, X)$ that contains $\mathcal{R}$ and is closed under substitution and under context.*[5] *We simply write $\to$ when $\mathcal{R}$ is clear from the context. The transitive and reflexive closure of $\to$ is written as $\twoheadrightarrow$. When all operation symbols in Σ have arity 1, then $\mathcal{R} = (\Sigma, R)$ is called a* **string rewriting system**.

Example 14. Let $\Sigma := \{0^{(0)}, s^{(1)}, +^{(2)}\}$ and $\mathcal{R} = \{x + 0 \to x,\ x + s(y) \to s(x + y)\}$. A rewrite sequence is for instance

$$s(s(0)) + s(0) \quad \to \quad s(s(s(0)) + 0) \quad \to \quad s(s(s(0))).$$

Definition 15. *Let $\mathcal{R} := (\Sigma, R)$ be a TRS.*

- $\mathcal{R}$ *is* **terminating** *or* **strongly normalising (SN)** *if every rewriting sequence is finite $t_0 \to t_1 \to \ldots \to t_n \not\to$.*
- $\mathcal{R}$ *is* **locally confluent** *or* **weak Church-Rosser** (WCR) *if for all terms t_1, t_2, t_3 with $t_2 \leftarrow t_1 \to t_3$, there exists a term t_4 with $t_2 \twoheadrightarrow t_4 \twoheadleftarrow t_3$.*
- $\mathcal{R}$ *is* **confluent** *or* **Church-Rosser** (CR) *if for all terms t_1, t_2, t_3 with $t_2 \twoheadleftarrow t_1 \twoheadrightarrow t_3$, there exists a term t_4 with $t_2 \twoheadrightarrow t_4 \twoheadleftarrow t_3$.*

$$\begin{tikzcd}
 & t_1 \arrow[dl] \arrow[dr] & \\
t_2 \arrow[dr, dotted, two heads] & (WCR) & t_3 \arrow[dl, dotted, two heads] \\
 & t_4 &
\end{tikzcd}$$

$$\begin{tikzcd}
 & t_1 \arrow[dl, two heads] \arrow[dr, two heads] & \\
t_2 \arrow[dr, dotted, two heads] & (CR) & t_3 \arrow[dl, dotted, two heads] \\
 & t_4 &
\end{tikzcd}$$

A term is called a *normal form*, if it cannot be rewritten any further. If a TRS is terminating (SN) and confluent (CR), then each term can be rewritten to a unique normal form.

A well-known result says that in the presence of termination, local confluence is enough to entail confluence.

Lemma 16 (Newman's Lemma). *If a TRS is terminating* (SN) *and locally confluent* (WCR)*, then it is also confluent* (CR).

Two common techniques to prove termination are the *polynomial interpretation* over $\mathbb{N}$ [32, §6.2.2] and the *multiset path order* [31]. The idea of polynomial interpretation over $\mathbb{N}$ is to choose a Σ-algebra $(\mathbb{N}, [\![\cdot]\!])$ where every interpretation $[\![\mathsf{op}]\!]$ is a monotone polynomial on $\mathbb{N}$. If each rule (l, r) of a system is strictly decreasing, $[\![l]\!] > [\![r]\!]$, then termination follows by well-foundedness of $\mathbb{N}$.

---

[5] For the definition of context, see [32, §2.1.1]

Example 17. The TRS in Example 14 is terminating. To see this, take as polynomial interpretation for example $[\![0]\!] = 1$, $[\![s(x)]\!] = x+1$, and $[\![x+y]\!] = x+2y+1$. These polynomials are monotone and every rule is strictly decreasing:

$$\begin{aligned}[\![x+0]\!] = x + 2\cdot 1 + 1 = x + 3 > x = [\![x]\!],\\ [\![x+s(y)]\!] = x + 2y + 3 > x + 2y + 2 = [\![s(x+y)]\!].\end{aligned}$$

The multiset path order method uses a decreasing sequence of multisets to show termination. We explain this briefly in the arXiv version of the paper [30].

A common technique for proving local confluence is to prove convergence of *critical pairs* [32, §2.7]. Informally, a critical pair is formed when two rewrite rules can be applied to the same term while overlapping on one or more function symbols, creating two different terms. A critical pair *converges* if the two mentioned terms can be rewritten to the same term.

Lemma 18 (Critical pair lemma). *A TRS is locally confluent* (WCR) *if and only if all its critical pairs converge.*

3 Composite Theories

We introduce the concept of *composite theories*. Our definition is slightly different from, but equivalent to, the original definition by Piróg & Staton [26, Def. 3] and equivalent formulations in Zwart's thesis [35, Def. 3.2, Prop. 3.4].

Definition 19. *Let $\mathbb{U}, \mathbb{S}, \mathbb{T}$ be algebraic theories. Suppose $\mathbb{U}$ contains $\mathbb{S}$ and $\mathbb{T}$, meaning $\Sigma_{\mathbb{S}}, \Sigma_{\mathbb{T}} \subseteq \Sigma_{\mathbb{U}}$ and $E_{\mathbb{S}}, E_{\mathbb{T}} \subseteq E_{\mathbb{U}}$.*

- *A $\mathbb{U}$-term is* **separated** *if it is of the form $t[s_x/x]$, where t is a $\mathbb{T}$-term and $\{s_x \mid x \in \mathsf{var}(t)\}$ is a family of $\mathbb{S}$-terms.*
- *Two separated terms $t[s_x]$ and $t'[s'_y]$ are* **equal modulo** $(\mathbb{S}, \mathbb{T})$ *if their TS-equivalence classes are equal in $TS\mathcal{V}$: $\overline{t[\overline{s_x}^{\mathbb{S}}]}^{\mathbb{T}} = \overline{t'[\overline{s'_y}^{\mathbb{S}}]}^{\mathbb{T}}$.*
- *$\mathbb{U}$ is a* **composite theory** *of $\mathbb{T}$ after $\mathbb{S}$ if every $\mathbb{U}$-term u is equal to a separated term $u =_{\mathbb{U}} t[s_x/x]$, that we call a* **separation** *of u, and for any two separated terms v, v', if $v =_{\mathbb{U}} v'$ then v and v' must be equal modulo* $(\mathbb{S}, \mathbb{T})$.

Lemma 20. *For any two separated terms $t[s_x/x]$ and $t'[s_y/y]$ in a composite theory, the following are equivalent:*

1. *$t[s_x/x]$ and $t'[s'_y/y]$ are equal modulo $(\mathbb{S}, \mathbb{T})$ in the sense of Definition 19.*
2. *$t[s_x/x]$ and $t'[s'_y/y]$ are equal modulo $(\mathbb{S}, \mathbb{T})$ in the sense of [35, Def. 3.2].* ■[6]

Example 21. Two $\mathbb{S}$-terms s and s' are equal modulo $(\mathbb{S}, \mathbb{T})$ if and only if $s =_{\mathbb{S}} s'$, and similarly for $\mathbb{T}$-terms.

---

[6] The symbol ■ denotes that the proof is in the extended version on arXiv [30].

Example 22. The prime example of a composite theory is the theory of rings $\mathbb{U} := \mathsf{Ring}$. It contains the theories $\mathbb{S} := \mathsf{Mon}$ of monoids and $\mathbb{T} := \mathsf{AbGrp}$ of Abelian groups. We recall their signatures to fix notation: $\Sigma_{\mathsf{Mon}} := \{\cdot^{(2)}, 1^{(0)}\}$ and $\Sigma_{\mathsf{AbGrp}} := \{0^{(0)}, +^{(2)}, -^{(1)}\}$. We sometimes omit the "multiplication" symbol $\cdot$ for simplicity. The signature of rings is given by $\Sigma_{\mathsf{Ring}} := \Sigma_{\mathsf{Mon}} \uplus \Sigma_{\mathsf{AbGrp}}$. The equations of rings are given by the equations of monoids, Abelian groups, and two distributivity axioms:

$$E_{\mathsf{Ring}} := E_{\mathsf{Mon}} \cup E_{\mathsf{AbGrp}} \cup \begin{Bmatrix} x(y+z) = (xy) + (xz), \\ (y+z)x = (yx) + (zx) \end{Bmatrix}.$$

A separated term $t[s_x/x]$ in Ring is an Abelian group term t, with monoid terms $\{s_x\}$ substituted for its variables. We give some examples of non-separated terms, of possible separations for them, and of equality modulo $(\mathsf{Mon}, \mathsf{AbGrp})$ between the separations.

The term $x(y+z)$ is non-separated. Possible separations are e.g. $xy + xz$ and $(xy + xz) + 0$. Both are equal modulo $(\mathsf{Mon}, \mathsf{AbGrp})$, as their monoid parts are identical and their Abelian group parts $t = (x_1 + x_2) + 0$ and $t' = x_1 + x_2$ are equal in the theory of Abelian groups.

The term $x \cdot 0$ is also non-separated. It is equal in Ring to the separated terms 0 and $(1 \cdot x) + (-(x \cdot 1))$. To see that these separations are equal modulo $(\mathsf{Mon}, \mathsf{AbGrp})$, notice that $1 \cdot x =_{\mathsf{Mon}} x \cdot 1$, and that the terms 0 and $x_1 + (-x_2)$ are equal in Abelian groups when $x_1 = x_2$. Thus: $\overline{0}^{\mathsf{AbGrp}} = \overline{(\overline{1 \cdot x}^{\mathsf{Mon}}) + (-(\overline{x \cdot 1}^{\mathsf{Mon}}))}^{\mathsf{AbGrp}}$.

We now show that distributive laws between monads correspond one-to-one to composite theories.

4 From Composite Theory to Distributive Law

We first show how to construct a distributive law from a given composite theory.

Theorem 23 ([35, Theorem 3.8]). *Let $\mathbb{S}, \mathbb{T}$ be algebraic theories with free algebra monads S, T respectively. Let $\mathbb{U}$ be a composite theory of $\mathbb{T}$ after $\mathbb{S}$, with free algebra monad U. Then the following defines a distributive law $\lambda : ST \Rightarrow TS$ such that $\mathbb{U}$ is an algebraic presentation of the resulting monad TS, where $t'[s'_x]$ is a separation of $s[t_x]$:*

$$\lambda_{\mathcal{V}} : ST\mathcal{V} \to TS\mathcal{V} : \quad \overline{s[\overline{t_x}^{\mathbb{T}}/x]}^{\mathbb{S}} \mapsto \overline{t'[\overline{s'_x}^{\mathbb{S}}/x]}^{\mathbb{T}}$$

Proof. Instead of directly checking the axioms for a distributive law, we prove an equivalent characterisation given by Beck [5, p.122]. That is, we claim that there exist a natural transformation $\mu^{TS} : TSTS \Rightarrow TS$ such that:

(i) $(TS, \eta^{TS} := \eta^T \eta^S, \mu^{TS})$ is a monad.
(ii) The natural transformations $\eta^T S$ and $T\eta^S$ are monad morphisms.
(iii) The middle unitary law holds: $\mu^{TS} \cdot T\eta^S \eta^T S = \mathrm{id}\, TS$.

It follows then that the monad $(TS, \eta^T\eta^S, \mu^{TS})$ does indeed come from a distributive law, which is given by: $\lambda = \mu^{TS} \cdot \eta^T ST\eta^S$. A simple but tedious calculation shows that indeed $\lambda(\overline{s[\overline{t_x}^{\mathbb{T}}/x]}^{\mathbb{S}}) = \overline{t'[\overline{s'_x}^{\mathbb{S}}/x]}^{\mathbb{T}}$. The details of this calculation are in the extended version [30].

To define μ^{TS}, we use the fact that the *functors* U and TS are isomorphic. Indeed, since $\mathbb{U}$ is a composite theory, every $\mathbb{U}$-term u has a separation $u =_{\mathbb{U}} t[s_x/x]$. Hence $\phi : U \Rightarrow TS$ and $\psi : TS \Rightarrow U$ given below are inverse natural transformations. Using ϕ, ψ, and the multiplication μ^U, we can then define μ^{TS}.

$$\phi(u) := \overline{t[\overline{s_x}^{\mathbb{S}}/x]}^{\mathbb{T}} \tag{15}$$

$$\psi(\overline{t[\overline{s_x}^{\mathbb{S}}/x]}^{\mathbb{T}}) := \overline{t[s_x/x]}^{\mathbb{U}} \tag{16}$$

$$\mu^{TS} := \big(TSTS \xrightarrow{\psi\psi} UU \xrightarrow{\mu^U} U \xrightarrow{\phi} TS\big). \tag{17}$$

Notice that ϕ is well-defined, as the choice of the separation $t[s_x/x]$ does not matter by equality modulo $(\mathbb{S}, \mathbb{T})$. To see that ψ is also well-defined, take $\overline{t[\overline{s_x}^{\mathbb{S}}/x]}^{\mathbb{T}} = \overline{t'[\overline{s'_x}^{\mathbb{S}}/x]}^{\mathbb{T}}$. The $\mathbb{T}$- and $\mathbb{S}$-proofs of that equality are also $\mathbb{U}$-proofs by definition of $\mathbb{U}$, implying that $\overline{t[\overline{s_x}^{\mathbb{U}}/x]}^{\mathbb{U}} = \overline{t'[\overline{s'_x}^{\mathbb{U}}/x]}^{\mathbb{U}}$ and hence that $\overline{t[s_x/x]}^{\mathbb{U}} = \overline{t'[s'_x/x]}^{\mathbb{U}}$ by applying $\mu^{\mathbb{U}}$ on both sides. The proofs of (i)-(iii) are in [30]. □

5 From Distributive Law to Composite Theory

We now show how to construct a composite theory from a given distributive law.

Theorem 24. *Let S, T be two monads algebraically presented by two algebraic theories $\mathbb{S}$ and $\mathbb{T}$, respectively. Let $\lambda : ST \Rightarrow TS$ be a distributive law. We define a set E_λ of equations and a theory $\mathbb{U}^\lambda$ as follows [35, Definition 3.8].*

$$E_\lambda := \Big\{ \big(s[t_x/x], t[s_y/y]\big) \mid \lambda_{\mathcal{V}}\big(\overline{s[\overline{t_x}^{\mathbb{T}}/x]}^{\mathbb{S}}\big) = \overline{t[\overline{s_y}^{\mathbb{S}}/y]}^{\mathbb{T}}\big) \Big\}.$$

$$\Sigma_{\mathbb{U}^\lambda} := \Sigma_{\mathbb{S}} \uplus \Sigma_{\mathbb{T}},$$

$$E_{\mathbb{U}^\lambda} := E_{\mathbb{S}} \cup E_{\mathbb{T}} \cup E_\lambda.$$

Then, $\mathbb{U}^\lambda$ is a composite theory of $\mathbb{T}$ after $\mathbb{S}$.

To prove Theorem 24, we observe that every $\mathbb{U}^\lambda$-term u can be assigned a regular set $\mathsf{type}(u)$ in $\{S, T\}^*\mathcal{V}$, expressing how u nests $\mathbb{S}$ and $\mathbb{T}$ operation symbols. We give an example below in Example 27. We obtain a TS-separated term by first mapping u to the equivalence class $\overline{u}$ in $\mathsf{type}(u)$, now viewed as a set. We then apply λ, μ^S and μ^T to $\overline{u}$ until we reach an equivalence class $\overline{t[\overline{s_x}^{\mathbb{S}}]}^{\mathbb{T}} \in TS\mathcal{V}$, where we use the axiom of choice to choose a representative $t[s_x]$. The axioms of the three natural transformations ensure that $\overline{t[\overline{s_x}^{\mathbb{S}}]}^{\mathbb{T}}$ does not depend on the order in which they were applied.

The termination of the procedure of applying λ, μ^S and μ^T and the uniqueness of $\overline{t[\overline{s_x}^{\mathbb{S}}]}^{\mathbb{T}}$ are intuitively clear, yet showing it formally is not trivial. In the following definitions we formalise the separation procedure that we described here. We then give a proof of termination using rewriting techniques. We denote string concatenation with "::".

Definition 25. *We define a function* $\mathsf{type} : \Sigma^*_{\mathbb{U}^\lambda}\mathcal{V} \to \{S,T\}^*\mathcal{V}$ *recursively:*

- *For* $v \in \mathcal{V}$, *then* $\mathsf{type}(v) := \mathcal{V}$.
- *For* $s[u_1,\ldots,u_n]$, *where* $s \in \mathcal{T}(\Sigma_{\mathbb{S}},\mathcal{V})$, *and* $u_1,\ldots,u_n \in \Sigma^*_{\mathbb{U}^\lambda}\mathcal{V}$ *do not have an* $\mathbb{S}$*-symbol as root, let* w *be longest word in the set* $\{\mathsf{type}(u_1),\ldots,\mathsf{type}(u_n)\}$, *then* $\mathsf{type}(s[u_1,\ldots,u_n]) := S :: w$.
- *The* $t[u_1,\ldots,u_n]$ *case, where* $u_1,\ldots,u_n$ *do not start with a* $\mathbb{T}$*-symbol, is dual.*

Informally, $\mathsf{type}(u)$ is the shortest string $w\mathcal{V}$ such that u belongs to an equivalence class in the set $w\mathcal{V}$. We will formally define this equivalence class in Definition 26 below. Furthermore, it can be seen that $\mathsf{type}(u)$ does not contain successive occurrences of S, similarly for T.

Definition 26. *For* $u \in \Sigma^*_{\mathbb{U}^\lambda}\mathcal{V}$ *and* $w \in \{S,T\}^*$ *such that* $\mathsf{type}(u)$ *is a substring*[7] *of* $w\mathcal{V}$, *we recursively define* $\overline{u}^w \in w\mathcal{V}$*:*

- *For* $v \in \mathcal{V}$, $\overline{v}^\varepsilon := v$, $\overline{v}^{S::w'} := \overline{\overline{v}^{w'}}^{\mathbb{S}}$, *and* $\overline{v}^{T::w'} := \overline{\overline{v}^{w'}}^{\mathbb{T}}$.
- *For* $s[u_1,\ldots,u_n]$ *where* $s \in \Sigma^*_{\mathbb{S}}\mathcal{V}$, *and* $u_1,\ldots,u_n \in \Sigma^*_{\mathbb{U}^\lambda}\mathcal{V}$ *that are either variables or have root symbols in* $\Sigma_{\mathbb{T}}$,

$$\overline{s[u_1,\ldots,u_n]}^{S::w'} := \overline{s[\overline{u_1}^{w'},\ldots,\overline{u_n}^{w'}]}^{\mathbb{S}}.$$

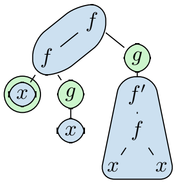
- *The* $t[u_1,\ldots,u_n]$ *case, where* $u_1,\ldots,u_n$ *do not start with a* $\mathbb{T}$*-symbol, is dual.*

If $\mathsf{type}(u)$ *is not a substring of* $w\mathcal{V}$, *then* $\overline{u}^w$ *is undefined.*

Example 27. Take the operations $f^{(2)}, f'^{(1)} \in \Sigma_{\mathbb{S}}$, and $g^{(1)} \in \Sigma_{\mathbb{T}}$. For $u := f(f(x,g(x)),g(f'(f(x,x))))$, we have

$$\mathsf{type}(u) = STS\mathcal{V}.$$

$$\overline{u}^{STS} = \overline{f(f(\overline{\overline{x}^{\mathbb{S}}}^{\mathbb{T}}, \overline{g(\overline{x}^{\mathbb{S}})}^{\mathbb{T}}), \overline{g(\overline{f'(f(x,x))}^{\mathbb{S}})}^{\mathbb{T}})}^{\mathbb{S}}.$$



Before we formalise the remainder of the separation procedure, we interpret functors and natural transformations as a term rewriting system.

Definition 28. *Let* $\Sigma := \{F_i \mid i \in I\}$ *be a finite set of (names of) functors, and* $\mathcal{R} := \{\alpha_j : w_j \to w'_j \mid w_j, w'_j \in \Sigma^*, j \in J\}$ *be a finite set of (names of) natural transformations. We call* $(\Sigma,\mathcal{R})$ *a* **functor rewriting system (FRS)**.

The name "functor rewriting system" is motivated by seeing each natural transformation $(\alpha : w \to w') \in \mathcal{R}$ as a rewrite rule on strings of functors in Σ^*. For all functor strings $w_0, w_1 \in \Sigma^*$, the natural transformation $w_0 \alpha w_1 : w_0 w w_1 \to w_0 w' w_1$ (sometimes called a *whiskering*) is seen as a rewrite step, with w_0 as left-context and w_1 as right-context. Note that the only valid rewrite steps are those resulting from natural transformations in $\mathcal{R}$. If the functors in Σ satisfy (semantic) identities like $FG = H$ that are not represented by some $\alpha \in \mathcal{R}$, then we do not allow rewrite steps that use this identity.

---

[7] We say that w is a substring of w' if w can be obtained by deleting zero or more letters from w'.

Remark 29. Kozen [19] introduced *rewrite categories* for applying rewriting concepts to categorical reasoning, including reasoning about monad compositions. A functor rewrite system $(\Sigma, \mathcal{R})$ is the rewrite category $(\Sigma^*, \mathcal{R})$. For Set-monads (S, μ^S, η^T) and (T, μ^T, η^T) and distributive law $\lambda\colon ST \to TS$, the FRS $\mathcal{R}^{sep}$ defined (below) in Definition 33 is the rewrite category $(\{S,T\}^*, \{\mu^S, \mu^T, \lambda\})$ viewed as a subcategory of the 2-category presented by $(\mathcal{O}, \mathcal{F}, \mathcal{R}, \mathcal{E})$ where $\mathcal{O} = \{\mathsf{Set}\}$, $\mathcal{F} = \{S, T\}$, $\mathcal{R} = \{\mu^S, \mu^T, \lambda\}$ and $\mathcal{E}$ consists of the equation (2) for μ^S and μ^T, and the distributive law axioms (10) and (11) involving λ and μ^S, μ^T. See also Section 7.1 for further discussion.

The functors and natural transformations in an FRS carry categorical structure in the form of commuting diagrams, allowing a variation of (local) confluence [19, §3.1].

Definition 30. *A functor rewriting system is (read* $\circlearrowleft$ *as "commuting")*

- WCR$\circlearrowleft$ *if for all* $w_0 \xleftarrow{\alpha} w \xrightarrow{\beta} w_1$ *there exists* $T_0 \overset{\gamma}{\twoheadrightarrow} w' \overset{\delta}{\twoheadleftarrow} T_1$ *s.t.* $\gamma\alpha = \delta\beta$.
- CR$\circlearrowleft$ *if for all* $w_0 \overset{\alpha}{\twoheadleftarrow} w \overset{\beta}{\twoheadrightarrow} w_1$ *there exists* $w_0 \overset{\gamma}{\twoheadrightarrow} w' \overset{\delta}{\twoheadleftarrow} w_1$ *s.t.* $\gamma\alpha = \delta\beta$.

There are equivalents to Newman's Lemma (Lemma 16) and the Critical Pair Lemma (Lemma 18). The proofs are in the extended version [30].

Lemma 31 (FRS Newman's lemma). *If a functor rewriting system is terminating* (SN) *and locally confluent-commuting* (WCR$\circlearrowleft$)*, then it is confluent-commuting* (CR$\circlearrowleft$). ■

Lemma 32 (FRS critical pair lemma). *A functor rewriting system is locally confluent-commuting* (WCR$\circlearrowleft$) *if and only if all critical pairs converge with a commuting diagram.* ■

We use the following FRS for our separation procedure.

Definition 33. *We define a functor rewriting system* $\mathcal{R}^{sep} = (\Sigma, R)$*, where* $\Sigma := \{S, T\}$ *and* $R := \{\lambda : ST \to TS,\ \mu^S : SS \to S,\ \mu^T : TT \to T\}$.

Lemma 34. $\mathcal{R}^{sep}$ *is terminating* (SN) *and confluent-commuting* (CR$\circlearrowleft$)*. Hence each functor string has a unique normal form in* $\mathcal{R}^{sep}$*.*

Proof. We show termination (SN) of $\mathcal{R}^{sep}$ using polynomial interpretation over $\mathbb{N}$. Let $[\![S]\!](x) := 2x + 1$ and $[\![T]\!](x) := x + 1$, which are indeed monotone in x. The three rewrite rules are strictly decreasing with respect to that order:

$$\begin{aligned}
[\![ST]\!](x) &= 2x + 3 > 2x + 2 = [\![TS]\!](x),\\
[\![SS]\!](x) &= 4x + 3 > 2x + 1 = [\![S]\!](x),\\
[\![TT]\!](x) &= x + 2 > x + 1 = [\![T]\!](x).
\end{aligned}$$

We now prove that $\mathcal{R}^{sep}$ is CR$\circlearrowleft$. Since we have termination (SN) it suffices to prove WCR$\circlearrowleft$ by Lemma 31. To invoke Lemma 32, we check that all critical pairs converge. Because we consider the objects purely syntactically as strings/words, we can enumerate all possible overlaps of left-hand sides of rules, giving rise to exactly 4 critical pairs, that indeed all converge:

$$\begin{tikzcd}
 & SST \arrow[dl, "S\lambda"'] \arrow[dr, "\mu^S T"] & \\
STS \arrow[d, "\lambda S"'] & (\lambda \text{ axiom}) & ST \arrow[ddl, "\lambda"] \\
TSS \arrow[dr, "T\mu^S"'] & (10) & \\
 & TS &
\end{tikzcd}
\quad
\begin{tikzcd}
 & STT \arrow[dl, "S\mu^T"'] \arrow[dr, "\lambda T"] & \\
ST \arrow[ddr, "\lambda"'] & (\lambda \text{ axiom}) & TST \arrow[d, "T\lambda"] \\
 & (11) & TTS \arrow[dl, "\mu^T S"] \\
 & TS &
\end{tikzcd}
\quad
\begin{tikzcd}
 & SSS \arrow[dl, "S\mu^S"'] \arrow[dr, "\mu^S S"] & \\
SS \arrow[dr, "\mu^S"'] & (S \text{ axiom})\ (2) & SS \arrow[dl, "\mu^S"] \\
 & S &
\end{tikzcd}
\quad
\begin{tikzcd}
 & TTT \arrow[dl, "T\mu^T"'] \arrow[dr, "\mu^T T"] & \\
TT \arrow[dr, "\mu^T"'] & (T \text{ axiom})\ (2) & TT \arrow[dl, "\mu^T"] \\
 & T &
\end{tikzcd}$$

□

We now have the required tools to formalise the separation procedure and show that every term in $\mathbb{U}^\lambda$ can be separated. The first step is to define a function sep that maps a $\mathbb{U}^\lambda$-term u to a separated term $\mathsf{sep}(u)$.

Definition 35. *For $u \in \Sigma^*_{\mathbb{U}^\lambda}$, we define $\mathsf{sep}(u)$ as follows. Let $w \in \{S, T\}^*$ be such that $\mathsf{type}(u) = w\mathcal{V}$. Let $\alpha \colon w \twoheadrightarrow w'$ be a $\mathcal{R}^{sep}$-rewrite sequence to the unique normal form w' of w in $\mathcal{R}^{sep}$. By the axiom of choice, there is a choice function $\rho_{w'\mathcal{V}}$ that selects a term representative $\rho_{w'\mathcal{V}}(c)$ for each equivalence class $c \in w'\mathcal{V}$. We define $\mathsf{sep}(u) := \rho_{w'\mathcal{V}}(\alpha_{\mathcal{V}}(\overline{u}^w))$.*

Remark 36. In general, we need the Axiom of Choice to obtain sep. However, if the theory $\mathbb{S}$ and $\mathbb{T}$ can be oriented[8] to give terminating and confluent TRSs, then we can make ρ select the unique normal form making sep constructive.

Lemma 37. *For all $u \in \Sigma^*_{\mathbb{U}^\lambda}$, $\mathsf{sep}(u)$ is a well-defined, separated $\mathbb{U}^\lambda$-term and $u =_{\mathbb{U}^\lambda} \mathsf{sep}(u)$.*

Proof. To see that $\mathsf{sep}(u)$ is well defined, note that if α and β are rewrite sequences $w \twoheadrightarrow w'$ from $w = \mathsf{type}(u)$ to its normal form w', then by CR$\circlearrowleft$, we have $\alpha = \beta$.

To see that $\mathsf{sep}(u)$ is separated, note that the normal form w' is equal to TS, T or S, since any other string will contain a reducible expression (redex). Hence $\alpha_{\mathcal{V}}(\overline{u}^w) \in TS\mathcal{V}, T\mathcal{V}$ or $S\mathcal{V}$, so any representative selected by $\rho_{w'\mathcal{V}}$ is separated.

To see that $u =_{\mathbb{U}^\lambda} \mathsf{sep}(u)$, recall that $\alpha \colon w \twoheadrightarrow w'$ is composed of λ, μ^S and μ^T, possibly applied within a context. By substitution and congruence rules, it suffices to prove that for all terms u, u' (of compatible type), if $\overline{u'}^{TS} = \lambda(\overline{u}^{ST})$ then $u =_{\mathbb{U}^\lambda} u'$, and similarly for μ^S and μ^T. That is, representatives of the input are $\mathbb{U}^\lambda$-equal to representatives of the output. For λ, this holds by definition of E_λ. For μ^S, if $u \in \overline{s[\overline{s_x}^{\mathbb{S}}]}^{\mathbb{S}}$ then $u =_{\mathbb{U}^\lambda} s[s_x]$, and if $u' \in \overline{s[s_x]}^{\mathbb{S}}$ then $u' =_{\mathbb{U}^\lambda} s[s_x]$. Hence by transitivity, $u =_{\mathbb{U}^\lambda} u'$. Similarly for μ^T. □

Lemma 38. *For all $\mathbb{S}$-terms s, $\mathsf{sep}(s) =_{\mathbb{S}} s$, for all $\mathbb{T}$-terms t, $\mathsf{sep}(t) =_{\mathbb{T}} t$, and for any separated term $t[s_x/x]$, $\mathsf{sep}(t[s_x/x])$ is equal to $t[s_x/x]$ modulo $(\mathbb{S}, \mathbb{T})$.*

---

[8] By *orientation*, we mean turning an equation $l = r$ into a rewrite rule, either from left to right $l \to r$ or right to left $l \leftarrow r$.

Proof. For an $\mathbb{S}$-term s, we have $\mathsf{type}(s) = S\mathcal{V}$ and $\overline{s}^{S} = \overline{s}^{\mathbb{S}}$. By definition $\mathsf{sep}(s) = \rho_{S\mathcal{V}}(\overline{s}^{\mathbb{S}})$ is a representative of $\overline{s}^{\mathbb{S}}$, hence $\mathsf{sep}(s) =_{\mathbb{S}} s$. The arguments for $\mathbb{T}$-terms and for separated terms $t[s_x/x]$ are similar. ☐

We now apply Lemma 37 to show that any two separated terms that are equal in $\mathbb{U}^{\lambda}$, are equal modulo $(\mathbb{S}, \mathbb{T})$.

Lemma 39. *Any two separated terms equal in $\mathbb{U}^{\lambda}$ are equal modulo $(\mathbb{S}, \mathbb{T})$.*

Proof. Suppose two separated terms $t_0[s_x/x]$ and $t'_0[s'_y/y]$ are equal in $\mathbb{U}^{\lambda}$. Let $\mathfrak{T}$ be a $\mathbb{U}^{\lambda}$-derivation tree of this equality $t_0[s_x/x] =_{\mathbb{U}^{\lambda}} t'_0[s_y/y]$ in equational logic. By an induction on the structure of $\mathfrak{T}$, we prove that for each equation $u = u'$ in $\mathfrak{T}$, $\mathsf{sep}(u)$ and $\mathsf{sep}(u')$ are equal modulo $(\mathbb{S}, \mathbb{T})$. By Lemma 38 and transitiviy of equality modulo $(\mathbb{S}, \mathbb{T})$, we then conclude that $t_0[s_x/x]$ and $t'_0[s'_y/y]$ are equal modulo $(\mathbb{S}, \mathbb{T})$.

The base cases are the Axiom and Reflexivity rules. The induction steps are the Symmetry, Transitivity, Congruence, and Substitution rules. We show only the cases of Congruence and Substitution here, as these are the only interesting cases. The full proof is in the extended version [30].

- Congruence: Given $\mathsf{op}^{(n)} \in \Sigma_{\mathbb{U}^{\lambda}}$, consider $\dfrac{u_1 = u'_1 \qquad \dots \qquad u_n = u'_n}{\mathsf{op}(u_1, \dots, u_n) = \mathsf{op}(u'_1, \dots, u'_n)}$

  Let $t_i[s_i] := \mathsf{sep}(u_i)$ and $t'_i[s'_i] := \mathsf{sep}(u'_i)$ for $i = 1, \dots, n$. The IH is that $\overline{t_i[\overline{s_i}^{\mathbb{S}}]}^{\mathbb{T}} = \overline{t'_i[\overline{s'_i}^{\mathbb{S}}]}^{\mathbb{T}}$. We consider the cases in which op is a $\mathbb{T}$-symbol or an $\mathbb{S}$-symbol separately.

  - Suppose $\mathsf{op} \in \Sigma_{\mathbb{T}}$. Here is a sketch of the reasoning:

    $$\begin{aligned}\overline{\mathsf{sep}(\mathsf{op}(u_1, \dots, u_n))}^{TS} &= \mu^{T}_{S\mathcal{V}}\left(\overline{\mathsf{op}\left(\overline{t_1[\overline{s_1}^{\mathbb{S}}]}^{\mathbb{T}}, \dots, \overline{t_n[\overline{s_n}^{\mathbb{S}}]}^{\mathbb{T}}\right)}^{\mathbb{T}}\right) \\ &= \mu^{T}_{S\mathcal{V}}\left(\overline{\mathsf{op}\left(\overline{t'_1[\overline{s'_1}^{\mathbb{S}}]}^{\mathbb{T}}, \dots, \overline{t'_n[\overline{s'_n}^{\mathbb{S}}]}^{\mathbb{T}}\right)}^{\mathbb{T}}\right) && \text{by IH} \\ &= \overline{\mathsf{sep}(\mathsf{op}(u'_1, \dots, u'_n))}^{TS}.\end{aligned}$$

    The first and third equalities are intuitively clear. The details can be found in the extended version [30].

  - Suppose $\mathsf{op} \in \Sigma_{\mathbb{S}}$. Here is a sketch of the reasoning:

    $$\begin{aligned}\overline{\mathsf{sep}(\mathsf{op}(u_1, \dots, u_n))}^{TS} &= T\mu^{S}_{\mathcal{V}} \cdot \lambda_{\mathcal{V}}\left(\overline{\mathsf{op}\left(\overline{t_1[\overline{s_1}^{\mathbb{S}}]}^{\mathbb{T}}, \dots, \overline{t_n[\overline{s_n}^{\mathbb{S}}]}^{\mathbb{T}}\right)}^{\mathbb{S}}\right) \\ &= T\mu^{S}_{\mathcal{V}} \cdot \lambda_{\mathcal{V}}\left(\overline{\mathsf{op}\left(\overline{t'_1[\overline{s'_1}^{\mathbb{S}}]}^{\mathbb{T}}, \dots, \overline{t'_n[\overline{s'_n}^{\mathbb{S}}]}^{\mathbb{T}}\right)}^{\mathbb{S}}\right) && \text{by IH} \\ &= \overline{\mathsf{sep}(\mathsf{op}(u'_1, \dots, u'_n))}^{TS}.\end{aligned}$$

    The first and third equalities are intuitively clear. The details can be found in the extended version [30].

- Substitution: Given a substitution f, consider $\dfrac{u = u'}{u[f] = u'[f]}$.

  Let $t[s_x] := \mathsf{sep}(u)$ and $t'[s'_x] := \mathsf{sep}(u')$. The IH is that $\overline{t[\overline{s_x}^{\mathbb{S}}]}^{\mathbb{T}} = \overline{t'[\overline{s'_x}^{\mathbb{S}}]}^{\mathbb{T}}$. We start by separating all terms in the image of f. This gives another substitution $g := \mathsf{sep} \cdot f$. We denote $t_y[s_z] := g(y)$ for all $y \in \mathsf{var}(u_1) \cup \mathsf{var}(u_2)$. Here is a sketch of the reasoning:

$$\begin{aligned} \overline{\mathsf{sep}(u[f])}^{TS} &= \mu^{TS}\left(\overline{t[s_x[t_y[s_z]]]}^{TSTS}\right) \\ &= \mu^{TS}\left(\overline{t'[s'_x[t_y[s_z]]]}^{TSTS}\right) && \text{by IH} \\ &= \overline{\mathsf{sep}(u'[f])}^{TS}. \end{aligned}$$

  The first and third equalities are intuitively clear. The details can be found in the extended version [30]. □

The proof of Theorem 24 now follows from Lemmas 37 and 39.

The next theorem was given in Zwart's thesis [35, Theorem 3.9] but not published elsewhere. We have updated the reasoning and obtained a much shorter proof using the shortcut $\mathbf{EM}(TS) \cong_{\mathsf{conc}} \mathbf{Alg}(\lambda)$.

Theorem 40. *Let S and T be the free algebra monads of algebraic theories $\mathbb{S}$ and $\mathbb{T}$. If there is a distributive law $\lambda : ST \Rightarrow TS$, then the monad $(TS, \eta^T\eta^S, \mu^T\mu^S \cdot T\lambda S)$ is presented algebraically by $\mathbb{U}^\lambda$.* ■

6 Axiomatisations of Composite Theories

In Theorem 24, we showed how to obtain an algebraic presentation $\mathbb{U}^\lambda$ of the composite monad arising from a distributive law $\lambda : ST \to TS$. However, the set of equations E_λ accounting for the interactions between $\mathbb{S}$- and $\mathbb{T}$-terms is maximal in the sense that it contains all possible equations that consist of representatives of some pair $(u, \lambda(u))$ in the graph of λ. In practice, we would like to have a minimal description of E_λ, such as the one for Ring in Example 22, which only adds two distribution axioms to the theories of monoids and Abelian groups.

In this section, we identify criteria on the shape of axioms that allow us to prove that certain minimal subsets of E_λ suffice to generate the whole of E_λ. We apply term rewriting methods for proving the necessary claims.

The shape of axioms will be described in terms of *layers*.

Definition 41. *Let $\mathbb{S}$ and $\mathbb{T}$ be two algebraic theories. Given a term $s[t_x/x] \in \Sigma^*_{\mathbb{S}}\Sigma^*_{\mathbb{T}}\mathcal{V}$, its ST-**layers** are described by the pair (m, n) of natural numbers where $m := \mathsf{depth}(s)$ and $n := \max\{\mathsf{depth}(t_x) \mid x \in \mathsf{var}(s)\}$, where depth denotes the maximal number of nested (possibly nullary) operation symbols. This corresponds to the inductively defined notion of depth of term trees where constants have depth 1, and variables depth 0. TS-layers are defined similarly for terms in $\Sigma^*_{\mathbb{T}}\Sigma^*_{\mathbb{S}}\mathcal{V}$.*

Example 42. We illustrate ST-layers in Ring (where $\mathbb{S} = \mathsf{Mon}, \mathbb{T} = \mathsf{AbGrp}$).

ST-Layers	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$	$(0,2)$	$(2,0)$
Examples	x	0	1	$x \cdot 0$	$x + 0$	$x \cdot 1$
	y	$x + y$	$x \cdot y$	$(x + y) \cdot (y + z)$	$(x + y) + z$	$x \cdot (y \cdot z)$

For the remainder of this section, we assume that $\mathbb{S}$, $\mathbb{T}$, λ, $E_{\mathbb{S}}$, $E_{\mathbb{T}}$, E_λ, and $\mathbb{U}^\lambda$ are as in Theorem 24.

Lemma 43. *For all $E' \subseteq E_\lambda$ such that for each $f^{(n)} \in \mathbb{S}$, $g^{(m)} \in \mathbb{T}$ and each $i \in \{1, \ldots, n\}$, E' contains one equation of the form $l = r$, where $l = f(x_1, \ldots, x_{i-1}, g(\vec{y}), x_{i+1}, \ldots, x_n)$ and $r \in \lambda_{\mathcal{V}}(\overline{\overline{l}^{\mathbb{T}}}^{\mathbb{S}})$, if the TRS $(\Sigma_{\mathbb{U}^\lambda} = \Sigma_{\mathbb{S}} \uplus \Sigma_{\mathbb{T}}, E')$ is terminating, then $E_{\mathbb{S}} \cup E_{\mathbb{T}} \cup E'$ generates the same congruence on $\mathbb{U}^\lambda$-terms as $E_{\mathbb{S}} \cup E_{\mathbb{T}} \cup E_\lambda$.* ■

Proof. Let us show why $(\Sigma_{\mathbb{U}^\lambda} = \Sigma_{\mathbb{S}} \uplus \Sigma_{\mathbb{T}}, E')$ is a TRS. First, no left-hand side is a variable by definition of E'. Second, $A := \mathsf{var}(s[t_x]) \supseteq \mathsf{var}(t[s_y])$ holds for all $(s[t_x], t[s_y]) \in E'$. This is the case since $\lambda_A : STA \to TSA : \overline{s[\overline{t_x}^{\mathbb{T}}]}^{\mathbb{S}} \mapsto \overline{t[\overline{s_y}^{\mathbb{S}}]}^{\mathbb{T}}$ forces the equivalence class of $t[s_y]$ to be in TSA and therefore to only use the variables in A.

Now let us argue why the congruence relation is left unchanged. Take an equation $(u, u') \in E_\lambda \cup E_{\mathbb{S}} \cup E_T$. The goal is to obtain this equation using only $E_{\mathbb{S}} \cup E_{\mathbb{T}} \cup E'$.

- First, using only equations in E', the $\mathbb{U}^\lambda$-terms u and u' can be separated. Indeed, we assume that the TRS $(\Sigma_{\mathbb{U}^\lambda}, E')$ is terminating, thus both u and u' can be rewritten to normal forms. The equations E' are exhaustive in the following sense: every term containing a $\Sigma_{\mathbb{T}}$-symbol below an $\Sigma_{\mathbb{S}}$-symbol is reducible (not in normal form). Thus the normal forms of u and u' must be in $\Sigma_{\mathbb{T}}^* \Sigma_S^* \mathcal{V}$, i.e., separated. Let us denote them $t[s_x/x]$ and $t'[s'_y/y]$.
- Since $\mathbb{U}^\lambda$ is a composite theory (proven in Theorem 24), and the separated normal forms $t[s_x/x]$ and $t'[s'_y/y]$ are $\mathbb{U}^\lambda$-equal, they must also be equal modulo $(\mathbb{S}, \mathbb{T})$. By equality modulo $(\mathbb{S}, \mathbb{T})$, we have a proof of $t[s_x/x] = t'[s'_y/y]$ using only equations from $E_{\mathbb{S}}$ and $E_{\mathbb{T}}$ (explicitly so when using the equivalent formulation (4) of equality modulo $(\mathbb{S}, \mathbb{T})$ in [35, Prop. 3.4]). □

In order to obtain an E' for Lemma 43, one can take equations of the form $l = \mathsf{sep}(l)$, but Lemma 43 also applies to other choices of r. As mentioned in Remark 36, if the theories $\mathbb{S}$ and $\mathbb{T}$ can be oriented to obtain a confluent and terminating TRS, then $\mathsf{sep}(l)$ can be chosen to be a normal form. For example, in [26], the theory of left-zero monoids and the theory with a unary idempotent operation were both oriented, allowing for a practical presentation of the composite theory that the authors called CUT.

Example 44. Let us retrieve the axiomatisation of Ring as given in Example 22, but starting from its corresponding distributive law $\lambda : L\mathcal{A} \to \mathcal{A}L$ [5, §4]. The set E will only contain equations whose left-hand side is among $(x+y)z$, $x(y+z)$, $0 \cdot x$, $x \cdot 0$, $(-x)y$, and $x(-y)$. For each of those, there are infinitely many choices for the right-hand side. For instance $(x \cdot 0, 0)$, $(x \cdot 0, 0+0)$, etc. Thankfully, there is an easy choice for the right-hand side r, because the theory Mon can be oriented,

$(xy)z \to x(yz)$, $1 \cdot x \to x$, and $x \cdot 1 \to x$, as can the theory AbGrp without the commutativity axiom. Not taking the commutativity axiom into account simply means that we have to choose one equation between $((x+y)z,\ xz+yz)$ and $((x+y)z,\ yz+xz)$. We end up with 6 equations:

$$\begin{aligned} (x+y)z &= xz+yz, & x\cdot 0 &= 0, & (-x)y &= -(xy),\\ z(x+y) &= zx+zy, & 0\cdot x &= 0, & x(-y) &= -(xy). \end{aligned}$$

Reducing from 6 to only the 2 equations of left and right distributivity can be done using automated tools. In our case, we used Prover9 [23] and obtained the result instantaneously [30, §8.7].

Note that if $E' \subseteq E_\lambda$ is not terminating, then the conclusion is not guaranteed to hold. The example below exhibits a situation where the set E' of equations as defined in Lemma 43 is not enough to generate all of the E_λ equations.

Example 45. We show that the subset of equations of E_λ where all left-hand sides have layers $(1,1)$ is not always sufficient (together with $E_{\mathbb{S}}$ and $E_{\mathbb{T}}$) to generate all E_λ equations obtained from a distributive law λ. This example is an extension of the well-known non-terminating TRS $ab \to bbaa$ [32, Ex.2.3.9].

Consider the theories $\mathbb{S}$ and $\mathbb{T}$, with signatures $\Sigma_{\mathbb{S}} := \{a^{(1)}\}$ and $\Sigma_{\mathbb{T}} := \{b^{(1)}\}$, and equations $E_{\mathbb{S}} := \{aaa = aa\}$ and $E_{\mathbb{T}} := \{bbb = bb\}$. We use some string rewriting notations, such as aax or a^2x as shorthand for $a(a(x))$, etc. The set of equivalence classes of $\mathbb{S}$ is $S\mathcal{V} = \{\overline{a^2x}^{\mathbb{S}}, \overline{ax}^{\mathbb{S}}, \overline{x}^{\mathbb{S}} \mid x \in \mathcal{V}\}$. Similarly, $T\mathcal{V} = \{\overline{b^2x}^{\mathbb{T}}, \overline{bx}^{\mathbb{T}}, \overline{x}^{\mathbb{T}} \mid x \in \mathcal{V}\}$. We define a mapping

$$\begin{aligned} \lambda\colon ST\mathcal{V} &\to TS\mathcal{V}\\ \overline{a^n\overline{b^mx}^{\mathbb{T}}}^{\mathbb{S}} &\mapsto \overline{b^2\overline{a^2x}^{\mathbb{S}}}^{\mathbb{T}}, && \text{for } n,m \in \{1,2\}\\ \overline{a^n\overline{x}^{\mathbb{T}}}^{\mathbb{S}} &\mapsto \overline{\overline{a^nx}^{\mathbb{S}}}^{\mathbb{T}}, && \text{for } n \in \{1,2\}\\ \overline{\overline{b^nx}^{\mathbb{T}}}^{\mathbb{S}} &\mapsto \overline{b^n\overline{x}^{\mathbb{S}}}^{\mathbb{T}}, && \text{for } n \in \{1,2\}\\ \overline{\overline{x}^{\mathbb{T}}}^{\mathbb{S}} &\mapsto \overline{\overline{x}^{\mathbb{S}}}^{\mathbb{T}} \end{aligned}$$

We show that λ is a distributive law:

- Unit law (8): $\lambda_{\mathcal{V}}(S\eta^T_{\mathcal{V}}(\overline{a^nx}^{\mathbb{S}})) = \lambda_{\mathcal{V}}(\overline{a^n\overline{x}^{\mathbb{T}}}^{\mathbb{S}}) = \overline{\overline{a^nx}^{\mathbb{S}}}^{\mathbb{T}} = \eta^T_{S\mathcal{V}}(\overline{a^nx}^{\mathbb{S}})$.
- Unit law (9): $\lambda_{\mathcal{V}}(\eta^S_{T\mathcal{V}}(\overline{b^nx}^{\mathbb{T}})) = \lambda_{\mathcal{V}}(\overline{\overline{b^nx}^{\mathbb{T}}}^{\mathbb{S}}) = \overline{b^n\overline{x}^{\mathbb{S}}}^{\mathbb{T}} = T\eta^S_{\mathcal{V}}(\overline{b^nx}^{\mathbb{T}})$.
- Multiplication law (10): We only show the case for $n, m, k \geqslant 1$. Other cases can be easily verified in a similar manner.

$$\begin{tikzcd}
\overline{a^n\overline{a^m\overline{b^kx}^{\mathbb{T}}}^{\mathbb{S}}}^{\mathbb{S}} \in SST \arrow[r, "S\lambda"] \arrow[d, "\mu^S T"'] & \overline{a^nb^2\overline{\overline{a^2x}^{\mathbb{S}}}^{\mathbb{T}}}^{\mathbb{S}} \in STS \arrow[r, "\lambda S"] & \overline{b^2\overline{a^2\overline{a^2x}^{\mathbb{S}}}^{\mathbb{S}}}^{\mathbb{T}} \in TSS \arrow[d, "T\mu^S"] \\
\overline{a^{n+m}\overline{b^kx}^{\mathbb{T}}}^{\mathbb{S}} \in ST \arrow[rr, "\lambda"'] & & \overline{b^2\overline{a^2x}^{\mathbb{S}}}^{\mathbb{T}} = \overline{b^2\overline{a^4x}^{\mathbb{S}}}^{\mathbb{T}} \in TS
\end{tikzcd}$$

- Multiplication law (11): Analogous to the previous point.

From Theorem 24, defining the set E_λ of distributivity equations as below ensures that $E_\mathbb{S} \cup E_\mathbb{T} \cup E_\lambda$ is an axiomatization of the composite theory $\mathbb{U}^\lambda$.

$$\begin{aligned} E_\lambda = \{&a^n b^m x = b^2 a^2 x \mid m, n \geqslant 1, x \in \mathcal{V}\} \cup \\ \{&a^n x = a^n x, b^n x = b^n x \mid n \in \{0, 1, 2\}, x \in \mathcal{V}\} \end{aligned}$$

The subset of equations of E_λ that have left-hand side with ST-layers $(1, 1)$ is $E' = \{ab = b^2 a^2\}$. However, we claim that $E_\mathbb{S} \cup E_\mathbb{T} \cup E'$ cannot derive all equations in E_λ. Indeed, we observe that the distributivity equation $aab =_{E_\lambda} bbaa$ cannot be derived. Trying to do so leaves us stuck in a loop: (we underline the part where an equation is applied)

$$\begin{aligned} a\underline{ab} =_{E'} \underline{ab}baa =_{E'} bba\underline{ab}aa =_{E'} bbabb\underline{aaaa} =_{E_\mathbb{S}} bb\underline{ab}baa \\ =_{E'} \underline{bbbb}aabaa =_{E_\mathbb{T}} bbaabaa =_{E'} \ldots \text{ (loop)} \end{aligned}$$

It is not hard to see that there are no other ways of proving $aab = bbaa$ in $E_\mathbb{S} \cup E_\mathbb{T} \cup E'$. Hence $E_\mathbb{S} \cup E_\mathbb{T} \cup E'$ does not generate the same congruence as $E_\mathbb{S} \cup E_\mathbb{T} \cup E_\lambda$. In line with Lemma 43, the above indeed also shows that E', when viewed as a TRS, is not terminating. Note that Lemma 43 only says that termination is a sufficient condition for a (1,1)-axiomatisation. It does not exclude that in some composite theories, the set of equations $E_\mathbb{S} \cup E_\mathbb{T} \cup E'$ might axiomatise $\mathbb{U}^\lambda$ even in presence of non-termination.

The next lemma identifies a class of equations where termination of the TRS $(\Sigma_\mathbb{U} = \Sigma_\mathbb{S} \uplus \Sigma_\mathbb{T}, E')$ is guaranteed. These are equations in which the right-hand sides have layers $(n, 1)$, which is inspired from similar results for string rewriting obtained by Zantema & Geser [34].

Lemma 46. *Let $\mathbb{S}$ and $\mathbb{T}$ be two algebraic theories. Let R be a set rules of the form $s[t_x/x] \to t[s_y/y]$. Let $Z = \{t_x \mid t_x$ is a variable$\}$, i.e., all $z \in Z$ occur directly below an $\mathbb{S}$-operation in $s[t_x/x]$. If each $s[t_x/x]$ has ST-layers $(1, 1)$, each $t[s_y/y]$ has TS-layers $(n, 1)$ for some n not fixed, and each s_y is linear*[9] *in Z, then R is terminating.* ■

Example 47. We give some axiomatisations of composite theories resulting from distributive laws in the literature:

1. Let $R(X) = X^A$ be the reader monad, with $A = \{a_1, \ldots, a_n\}$. There is a distributive law of the finite distribution monad $\mathcal{D}$ over R, $\lambda : \mathcal{D}R \to R\mathcal{D}$, that sends $p_1 h_1 + \ldots + p_n h_n$ to $(a \mapsto p_1 h_1(a) + \ldots + p_n h_n(a))$ [13, Example 1.34]. Recall that R is presented algebraically by a single operation $f^{(n)}$ with two equations Example 12, and $\mathcal{D}$ is presented by convex algebras. The distribution axioms as described in Lemma 43 are in our case, for each $p \in [0, 1]$

$$\begin{aligned} f(x_1, \ldots, x_n) \oplus_p y = f(x_1 \oplus_p y, \ldots, x_n \oplus_p y). \\ x \oplus_p f(y_1, \ldots, y_n) = f(x \oplus_p y_1, \ldots, x \oplus_p y_n). \end{aligned}$$

---

[9] *Linear* in a TRS sense, i.e. variables appearing *at most* once.

We see that the right-hand sides of these equations have layers $(1, 1)$ and both equations satisfy the linearity requirement of Lemma 46, thus ensuring termination. Hence by Theorem 24 and Lemma 43, the above equations together with the equations for f and for convex algebras present the composite monad on $R\mathcal{D}$ induced by λ. Furthermore, we notice that each of the above equations can be derived from the other one using the axioms of convex algebras. Therefore, we only need to include one of them for each p.

2. There is a distributive law of multisets over distributions $\lambda\colon \mathcal{M}\mathcal{D} \to \mathcal{D}\mathcal{M}$ called the *parallel multinomial law* in [16], see also [9, 11] and [13, Ex. 1.37]. It sends e.g. $\Lbag px_1 + (1-p)x_2, y \Rbag$ to $p\Lbag x_1, y\Rbag + (1-p)\Lbag x_2, y\Rbag$, which can be expressed in the syntax of convex algebras and commutative monoids as

   $$(x_1 \oplus_p x_2) \cdot y = (x_1 \cdot y) \oplus_p (x_2 \cdot y).$$

   By Theorem 24, Lemma 43 and Lemma 46 these equations (one for each $p \in [0, 1]$), together with the axioms of convex algebras and commutative monoids, present the composite monad on $\mathcal{D}\mathcal{M}$ induced by λ.
3. There is a distributive law $\lambda\colon L^+L^+ \to L^+L^+$ for the non-empty list monad over itself [22]. It sends a list of lists to the singleton list containing the list of all heads: $[[a, b], [c], [d, e, f]] \mapsto [[a, c, d]]$. We get the following distributivity axioms for the composite theory:

   $$\begin{aligned} a * (b \star c) &= a * b \\ (a \star b) * c &= a * c. \end{aligned}$$

   Again, the equations satisfy the conditions for Lemma 46, and our results imply that the above equations together with the semigroup axioms for $*$ and $\star$ present the composite monad on L^+L^+ induced by λ.

7 Conclusion

In this paper, we proved the correspondence between composite theories of $\mathbb{T}$ after $\mathbb{S}$ and distributive laws $\lambda\colon ST \to TS$. Furthermore, we gave sufficient criteria for when a minimal set $E' \subseteq E_\lambda$ of distribution equations, along with $E_\mathbb{S}$ and $E_\mathbb{T}$, axiomatises the composite theory.

The set E' itself is unlikely to turn many heads, as distributive laws are often informally described in the literature in terms of such simple distribution axioms. The surprise, however, comes from the fact that E' is not always enough (see Example 45). This is a possible pitfall similar to the 'simplicity' of the various false distributive laws of the powerset monad over itself [18].

7.1 Related work

In Kozen's work on rewrite categories, he proves that distributive laws yield composite monads in [19, Section 4.2], by showing that crucial properties correspond to TS being a terminal object in the rewrite category with $\mu^S, \mu^T, \eta^S, \eta^T, \lambda$.

However, we cannot apply these results to prove Theorem 24 since they do not involve composite theories. Another difference with Kozen's approach is that we do not include the monad units in $\mathcal{R}^{sep}$ (Definition 33). By omitting the units, we obtain unique normal forms in $\mathcal{R}^{sep}$ in the classic rewriting sense, but no terminal object in the corresponding rewrite category. This allows our reasoning to follow classic rewrite arguments more closely.

A result akin to Theorem 40 appears in the literature on polygraphs [2, 3.3.6 Theorem]. Polygraphs are generalisations of graphs that can serve as presentations of categories. The notion of distributive law between categories presented by polygraphs seems related to the notion of distributive law between Lawvere Theories as described by Cheng [8], but the precise connection is not explained in [2] and remains to be explored.

7.2 Future work

There are several directions for future work. We showed that termination of E' (as TRS) is sufficient for $E_{\mathbb{S}} \cup E_{\mathbb{T}} \cup E'$ to axiomatise the composite theory (Lemma 43), and that taking equations in E' to have layers $(1,1) \to (n,1)$ ensures termination (Lemma 46). We would like to identify other criteria for termination, and make more use of term rewriting techniques. We speculate that one could allow layers $(1,1) \to (2,2)$ in which some symbol in the left-hand side is absent from the right-hand side in order to avoid problems such as in Example 45 with $ab \to bbaa$.

In light of negative results concerning monad compositions [18, 33, 36, 10], there has been much interest in understanding the limits of monad composition. Positive results using algebraic methods were given in [9]. Another approach has been to generalise to so-called weak distributive laws [12, 13]. Presentations of monads arising from the composition of monads via a weak distributive law, in particular monads for nondeterminism and probabilities, have been given in [6, 14]. These presentations are obtained by adding a simple distribution axiom to the two underlying theories, similar to our results in Section 6, but the resulting theory is no longer a composite theory as the essential uniqueness modulo $(\mathbb{S}, \mathbb{T})$ is not guaranteed to hold. Another future line of work would be to extend the current correspondence to weak distributive laws [12, 13] thereby giving a definition of *weak composite theories*. Such a correspondence would allow for a more thorough study of weak distributive laws on the algebraic level, and could perhaps lead to no-go theorems for weak distributive laws.

Alternatively, the current correspondence could also be extended to account for multi-sorted algebraic theories, and by such means defining *multi-sorted distributive law*.

Acknowledgments: We thank all anonymous reviewers for their valuable feedback and suggestions. Aloïs Rosset and Jörg Endrullis received funding from the Netherlands Organization for Scientific Research (NWO) under the Innovational Research Incentives Scheme Vidi (project. No. VI.Vidi.192.004).

References

1. Aczel, P., Adámek, J., Milius, S., Velebil, J.: Infinite trees and completely iterative theories: a coalgebraic view. Theoretical Computer Science **300**(1-3), 1–45 (2003). https://doi.org/10.1016/S0304-3975(02)00728-4
2. Ara, D., Burroni, A., Guiraud, Y., Malbos, P., Métayer, F., Mimram, S.: Polygraphs: From rewriting to higher categories. CoRR **abs/2312.00429** (2023). https://doi.org/10.48550/ARXIV.2312.00429
3. Awodey, S.: Category Theory. Oxford Logic Guides, Ebsco Publishing (2006)
4. Barr, M., Wells, C.: Toposes, Triples and Theories. Comprehensive studies in mathematics, Springer New York (1985)
5. Beck, J.: Distributive laws. In: Seminar on triples and categorical homology theory. p. 119–140. Springer (1969)
6. Bonchi, F., Sokolova, A., Vignudelli, V.: The theory of traces for systems with nondeterminism and probability. In: 34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2019. pp. 1–14. IEEE (2019). https://doi.org/10.1109/LICS.2019.8785673
7. Borceux, F.: Handbook of categorical algebra. 2, Encyclopedia of Mathematics and its Applications, vol. 51. Cambridge University Press, Cambridge (1994), categories and structures
8. Cheng, E.: Distributive laws for Lawvere theories. Compositionality (2020). https://doi.org/https://doi.org/10.32408/compositionality-2-1
9. Dahlqvist, F., Parlant, L., Silva, A.: Layer by layer – combining monads. In: Fischer, B., Uustalu, T. (eds.) Theoretical Aspects of Computing (ICTAC 2018). Lect. Notes Comp. Sci., vol. 11187. Springer (2018). https://doi.org/10.1007/978-3-030-02508-3_9
10. Dahlqvist, F., Neves, R.: Compositional semantics for new paradigms: probabilistic, hybrid and beyond (2018). https://doi.org/10.48550/ARXIV.1804.04145
11. Dash, S., Staton, S.: A monad for probabilistic point processes. In: Spivak, D.I., Vicary, J. (eds.) Proceedings of the 3rd Annual International Applied Category Theory Conference 2020, Cambridge, USA, 6-10th July 2020. Electronic Proceedings in Theoretical Computer Science, vol. 333, pp. 19–32. Open Publishing Association (2021). https://doi.org/10.4204/EPTCS.333.2
12. Garner, R.: The Vietoris monad and weak distributive laws. Applied Categorical Structures **28**(2), 339–354 (2020). https://doi.org/10.1007/s10485-019-09582-w
13. Goy, A.: On the compositionality of monads via weak distributive laws. (Compositionnalité des monades par lois de distributivité faibles). Ph.D. thesis, University of Paris-Saclay, France (2021), https://tel.archives-ouvertes.fr/tel-03426949
14. Goy, A., Petrisan, D.: Combining probabilistic and non-deterministic choice via weak distributive laws. In: Hermanns, H., Zhang, L., Kobayashi, N., Miller, D. (eds.) LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science. pp. 454–464. ACM (2020). https://doi.org/10.1145/3373718.3394795
15. Jacobs, B.: Convexity, duality and effects. In: Calude, C.S., Sassone, V. (eds.) Theoretical Computer Science - 6th IFIP TC 1/WG 2.2 International Conference, TCS 2010. Proceedings. IFIP Advances in Information and Communication Technology, vol. 323, pp. 1–19. Springer (2010). https://doi.org/10.1007/978-3-642-15240-5_1
16. Jacobs, B.: From multisets over distributions to distributions over multisets. In: Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer

Science. LICS '21, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1109/LICS52264.2021.9470678

17. Jacobs, B., Silva, A., Sokolova, A.: Trace semantics via determinization. Journal of Computer and System Sciences **81**(5), 859–879 (2015). https://doi.org/10.1016/j.jcss.2014.12.005
18. Klin, B., Salamanca, J.: Iterated covariant powerset is not a monad. In: Staton, S. (ed.) Proceedings of the Thirty-Fourth Conference on the Mathematical Foundations of Programming Semantics, MFPS 2018. Electronic Notes in Theoretical Computer Science, vol. 341, pp. 261–276. Elsevier (2018). https://doi.org/10.1016/j.entcs.2018.11.013
19. Kozen, D.: Natural transformations as rewrite rules and monad composition. Log. Methods Comput. Sci. **15**(1) (2019). https://doi.org/10.23638/LMCS-15(1:1)2019
20. MacLane, S.: Categories for the Working Mathematician, Graduate Texts in Mathematics, vol. 5. Springer-Verlag, New York (1971)
21. Manes, E.: Algebraic Theories, Graduate Texts in Mathematics, vol. 26. Springer (1976). https://doi.org/10.1007/978-1-4612-9860-1
22. Manes, E., Mulry, P.: Monad compositions. i: General constructions and recursive distributive laws. Theory and Applications of Categories **18**, 172–208 (2007)
23. McCune, W.: Release of prover9. In: Mile high conference on quasigroups, loops and nonassociative systems, Denver, Colorado (2005)
24. Moggi, E.: Notions of computation and monads. Information and Computation **93**(1), 55–92 (1991). https://doi.org/10.1016/0890-5401(91)90052-4, selections from 1989 IEEE Symposium on Logic in Computer Science
25. Parlant, L.: Monad composition via preservation of algebras. Ph.D. thesis, University College London, UK (2020), https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.819930
26. Piróg, M., Staton, S.: Backtracking with cut via a distributive law and left-zero monoids. Journal of Functional Programming **27**, e17 (2017). https://doi.org/10.1017/S0956796817000077
27. Plotkin, G.D., Power, J.: Notions of computation determine monads. In: Nielsen, M., Engberg, U. (eds.) Foundations of Software Science and Computation Structures, 5th International Conference, FOSSACS 2002. Proceedings. LNCS, vol. 2303, pp. 342–356. Springer (2002). https://doi.org/10.1007/3-540-45931-6_24
28. Riehl, E.: Category Theory in Context. Aurora: Dover Modern Math Originals, Dover Publications (2017)
29. Rosset, A., Hansen, H.H., Endrullis, J.: Algebraic presentation of semifree monads. In: Hansen, H.H., Zanasi, F. (eds.) Coalgebraic Methods in Computer Science. pp. 110–132. Springer International Publishing, Cham (2022)
30. Rosset, A., Zwart, M., Hansen, H.H., Endrullis, J.: Correspondence between composite theories and distributive laws (2024). https://doi.org/10.48550/ARXIV.2404.00581
31. Schneider-Kamp, P., Thiemann, R., Annov, E., Codish, M., Giesl, J.: Proving termination using recursive path orders and SAT solving. In: Konev, B., Wolter, F. (eds.) Frontiers of Combining Systems, 6th International Symposium, FroCoS 2007, Liverpool, UK, September 10-12, 2007, Proceedings. Lecture Notes in Computer Science, vol. 4720, pp. 267–282. Springer (2007). https://doi.org/10.1007/978-3-540-74621-8_18
32. Terese: Term rewriting systems, Cambridge tracts in theoretical computer science, vol. 55. Cambridge University Press (2003)

33. Varacca, D., Winskel, G.: Distributing probabililty over nondeterminism. Mathematical Structures in Computer Science **16**, 87–113 (2006). https://doi.org/10.1017/S0960129505005074
34. Zantema, H., Geser, A.: A complete characterization of termination of 0^{p} 1^{q}-> 1^{r} 0^{s}. Appl. Algebra Eng. Commun. Comput. **11**(1), 1–25 (2000). https://doi.org/10.1007/S002009900019
35. Zwart, M.: On the Non-Compositionality of Monads via Distributive Laws. Ph.D. thesis, Department of Computer Science, University of Oxford (September 2020)
36. Zwart, M., Marsden, D.: No-Go Theorems for Distributive Laws. Logical Methods in Computer Science **18** (Jan 2022). https://doi.org/10.46298/lmcs-18(1:13)2022