Logical Methods in Computer Science
Volume 19, Issue 4, 2023, pp. 9:1–9:47
https://lmcs.episciences.org/

Submitted Jul. 11, 2022
Published Nov. 03, 2023

THE POWER-SET CONSTRUCTION FOR TREE ALGEBRAS

ACHIM BLUMENSATH

Masaryk University Brno
e-mail address: blumens@fi.muni.cz

ABSTRACT. We study power-set operations on classes of trees and tree algebras. Our main result consists of a distributive law between the tree monad and the upwards-closed power-set monad, in the case where all trees are assumed to be *linear*. For *non-linear* ones, we prove that such a distributive law does not exist.

1. INTRODUCTION

The main approaches to formal language theory are based on automata, logic, and algebra. Each comes with their own strengths and weaknesses and thereby complements the other two. In the present article we focus on the algebraic approach, which is well-known for producing proofs that are often simpler than automaton-based ones, if not as elementary and at the cost of yielding worse complexity bounds. Algebraic methods are especially successful at deriving structural results about classes of languages. In particular, they are the method of choice when deriving characterisations of subclasses of regular languages. A prominent example of such a result is the Theorem of Schützenberger [Sch65] stating that a language is first-order definable if, and only if, its syntactic monoid is aperiodic. By now algebraic language theory is well-developed for a wide variety of settings and types of languages, including finite words, infinite words, and finite trees.

In recent years several groups have started to work on a category-theoretic unification of algebraic language theory [Boj, UACM17, Boj20, Blu20, Blu21]. The motivations include both the wish to simplify the existing theories and the need to generalise them to new settings, like infinite trees or data words. Here, we are interested in the case of languages of infinite trees, where an algebraic language theory has so far been missing. We continue the technical development of the framework presented in [Blu20, Blu21] by integrating a power-set operation. (To be precise, we use the upwards-closed power set since our framework is based on ordered sets.) Such an operation has numerous uses in language theory: for instance, when introducing regular expressions, for determinisation, or when proving closure under projections. We will present two such applications in Sections 5 and 6 below.

There are several ways to formalise languages of infinite trees. Most of the choices involved do not make much of a difference, but we isolate one design choice that does: a framework built on linear trees is much better behaved than one using possibly non-linear

Key words and phrases: tree algebras, power-set functor, distributive law.

DOI:10.46298/LMCS-19(4:9)2023

ones. This continues a trend already established in [Blu21] indicating that non-linear trees are more complicated than linear ones.

The main technical result needed for an integration of the power-set operation is a theorem stating that this operation can be lifted to the category of algebras under consideration. In category-theoretical lingo this means we have to establish a distributive law between the power-set monad and the monad our algebras are based on. Note that there has been recent renewed interest in distributive laws also in other parts of category theory (see, e.g., [GPA21, ZM22]), but the focus there is on different settings and, in particular, different functors.

We start in Section 2 by presenting our category-theoretical framework for infinite trees. Furthermore, we define the power-set operation we will be investigating, and we recall the notion of a distributive law, which will be central to our work. Section 3 contains a general derivation of such laws for a certain kind of polynomial monad, including the monad for linear trees, and a proof that the same is not possible for non-linear trees. The heart of the article is Section 4 where we will derive a partial result for non-linear trees that sometimes can be used as a substitute for a full distributive law. Finally, Sections 5 and 6 contain two applications: the first one is a simplified proof of a recently published result on substitutions for tree languages; while the second one describes how regular expressions can be defined using power sets of non-linear trees.

2. Monads for trees

In algebraic language theory one uses tools from algebra to study sets K of labelled objects. In the monadic framework from [Blu20, Blu21] these take the form $K \subseteq \mathbb{M}\Sigma$ where Σ is some alphabet and $\mathbb{M}$ is a suitable monad mapping a given set X to a set $\mathbb{M}X$ of X-labelled objects of a certain kind. Here we are mostly interested in three such monads: (i) the monad $\mathbb{R}$ of *rooted directed graphs;* (ii) the monad $\mathbb{T}$ of *linear trees;* and (iii) the monad $\mathbb{T}^\times$ of *possibly non-linear trees.* One of our results is that the latter two behave quite differently.

Fix a countably infinite set X of *variables* and let Ξ be the set of all finite subsets of X. As in [Blu21], we will be working in the category Pos^Ξ, the category of Ξ-sorted partial orders with monotone maps as morphisms. Thus, the objects are families $A = (A_\xi)_{\xi\in\Xi}$ where each sort A_ξ is equipped with a partial order, and the morphisms $f : A \to B$ are families $f = (f_\xi)_{\xi\in\Xi}$ of monotone maps $f_\xi : A_\xi \to B_\xi$. From this point on, we will use the terms 'set' and 'function' as a short-hand for 'ordered Ξ-sorted set' and 'order-preserving Ξ-sorted function'. For simplicity, we will frequently identify a sorted set $A = (A_\xi)_{\xi\in\Xi}$ with its disjoint union $A = \sum_{\xi\in\Xi} A_\xi$. Using this point of view, a morphism $f : A \to B$ corresponds to a sort-preserving and order-preserving function between the corresponding unions.

Given a set A, we consider *A-labelled, rooted, directed graphs* which are (possibly infinite) directed graphs with a distinguished vertex called the *root* such that every vertex is reachable by some directed path from the root. The edges of such graphs are labelled by elements of X and the vertices by elements of A in such a way that a vertex with label $a \in A_\xi$ has exactly one outgoing edge for each variable $x \in \xi$ and this edge is labelled by x. If there is an edge from v to u with label x, we call u the *x-successor* of v. We denote the set of vertices of a graph g by $\mathrm{dom}(g)$. Usually, we identify a graph g with the function $g : \mathrm{dom}(g) \to A$ mapping vertices of g to their labels. We can regard $\mathrm{dom}(g)$ as a set in Pos^Ξ by equipping it with the trivial order and by assigning sort ξ to a vertex v if ξ is the set of labels of the edges leaving v. Then $g : \mathrm{dom}(g) \to A$ is sort-preserving and order-preserving.

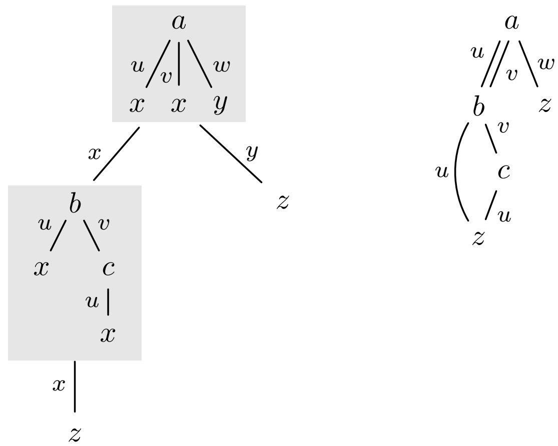


Figure 1: The flattening operation: g and $\mathrm{flat}(g)$ (edge directions not shown to reduce noise)

Definition 2.1. Let $A \in \mathsf{Pos}^{\Xi}$.

(a) For a sort $\xi \in \Xi$, we denote by $\mathbb{R}_\xi A \in \mathsf{Pos}$ the set of all $(A+\xi)$-*labelled rooted directed graphs* g (up to isomorphism) where

- the elements of ξ are called *variables* and have sort $\emptyset$,
- each variable $x \in \xi$ occurs at least once in g, and
- the root of g is *not* labelled by a variable.

The ordering on $\mathbb{R}_\xi A$ is defined componentwise:

$$g \leq h \quad :\text{iff} \quad \mathrm{dom}(g) = \mathrm{dom}(h) \quad \text{and} \quad g(v) \leq h(v)\,, \quad \text{for all } v \in \mathrm{dom}(g)\,.$$

(We assume that the ordering on ξ is just the identity.) We set

$$\mathbb{R}A := (\mathbb{R}_\xi A)_{\xi \in \Xi} \in \mathsf{Pos}^{\Xi}.$$

If $f : A \to B$ is a function, then $\mathbb{R}f : \mathbb{R}A \to \mathbb{R}B$ is the function that applies f to each label of the given graph (leaving the labels not in A unchanged).

(b) The *flattening function* $\mathrm{flat} : \mathbb{R}\mathbb{R}A \to \mathbb{R}A$ maps an $(\mathbb{R}A+\xi)$-labelled digraph g to the $(A+\xi)$-labelled digraph $\mathrm{flat}(g)$ that is obtained (see Figure 1) from the disjoint union of all digraphs $g(v)$, for $v \in \mathrm{dom}(g)$, by

- deleting from each component $g(v)$ every vertex labelled by a variable $x \in X$ and
- replacing every edge of $g(v)$ leading to such a vertex by an edge to the root of $g(u_x)$, where u_x is the x-successor of v in g.

The *singleton function* $\mathrm{sing} : A \to \mathbb{R}A$ maps an element $a \in A_\xi$ to the digraph g consisting of a root labelled by a and $|\xi|$ successors labelled by the variables in ξ.

(c) For $g \in \mathbb{R}A$, we denote by $\mathrm{dom}_0(g)$ the set of all vertices $v \in \mathrm{dom}(g)$ that are labelled by an element in A. ⌟

It is straightforward to check that $\mathbb{R}$ forms a monad. (Each of the three equations can be proved by exhibiting a label-preserving bijection between the respective domains.)

Proposition 2.2. $\langle \mathbb{R}, \mathrm{flat}, \mathrm{sing} \rangle$ *forms a monad on* Pos^{Ξ}*.*

The functors $\mathbb{T}$ and $\mathbb{T}^{\times}$ can now be derived from $\mathbb{R}$.

Definition 2.3. (a) For a set A, we denote by $\mathbb{T}^\times A \subseteq \mathbb{R}A$ the subset of all rooted graphs that are trees, and by $\mathbb{T}A \subseteq \mathbb{T}^\times A$ the subset consisting of all trees where every variable x appears exactly once. We call the elements of $\mathbb{T}A$ *linear trees over* A and those of $\mathbb{T}^\times A$ *non-linear trees.*

For finite trees in $\mathbb{T}^\times A$, we will frequently use the usual term notation like

$$a(x, b(y, x))\,, \quad \text{for } a, b \in A\,,\ x, y \in X\,.$$

(b) We denote the functions $\mathbb{T}\mathbb{T}A \to \mathbb{T}A$ and $A \to \mathbb{T}A$ induced by, respectively, $\mathrm{flat} : \mathbb{R}\mathbb{R}A \to \mathbb{R}A$ and $\mathrm{sing} : A \to \mathbb{R}A$ also by flat and sing. In cases where we want to distinguish between these versions, we add the functor as a superscript: $\mathrm{flat}^{\mathbb{R}}$, $\mathrm{flat}^{\mathbb{T}}$, etc.

(c) We denote the category of all $\mathbb{R}$-algebras by $\mathsf{Alg}(\mathbb{R})$, and similarly for the other monads. ⌟

The variants of flat and sing for the functor $\mathbb{T}^\times$ will be defined in a later section as $\mathbb{T}^\times$ does not form a submonad of $\mathbb{R}$. (The family of sets $\mathbb{T}^\times A$ is not closed under flat.)

The fact that $\mathbb{T}$ is a monad now follows directly from the fact that it is a restriction of $\mathbb{R}$. To see this, we need the notion of a morphism of monads.

Definition 2.4. Let $\langle \mathbb{M}, \mu, \varepsilon\rangle$ and $\langle \mathbb{N}, \nu, \eta\rangle$ be monads.

(a) A natural transformation $\varrho : \mathbb{M} \Rightarrow \mathbb{N}$ is a *morphism of monads* if

$$\eta = \varrho \circ \varepsilon \quad \text{and} \quad \nu \circ \varrho \circ \mathbb{M}\varrho = \varrho \circ \mu\,.$$

In this case we say that $\mathbb{M}$ is a *reduct* of $\mathbb{N}$.

(b) Let $\varrho : \mathbb{M} \Rightarrow \mathbb{N}$ be a morphism of monads and $\mathfrak{A} = \langle A, \pi\rangle$ an $\mathbb{N}$-algebra. The *ϱ-reduct* of $\mathfrak{A}$ is the $\mathbb{M}$-algebra $\mathfrak{A}|_\varrho := \langle A, \pi \circ \varrho\rangle$. If ϱ is understood, we also speak of the *$\mathbb{M}$-reduct* of $\mathfrak{A}$. ⌟

The following lemma is frequently useful to prove that a functor forms a monad. The proof is straightforward.

Lemma 2.5. *Let $\mathbb{M}$ and $\mathbb{N}$ be functors, $\mu : \mathbb{M}\mathbb{M} \Rightarrow \mathbb{M}$, $\nu : \mathbb{N}\mathbb{N} \Rightarrow \mathbb{N}$, $\varepsilon : \mathrm{Id} \Rightarrow \mathbb{M}$, $\eta : \mathrm{Id} \Rightarrow \mathbb{N}$ natural transformations, and let $\varrho : \mathbb{M} \Rightarrow \mathbb{N}$ be a natural transformation satisfying*

$$\eta = \varrho \circ \varepsilon \quad \text{and} \quad \nu \circ \varrho \circ \mathbb{M}\varrho = \varrho \circ \mu\,.$$

(a) *Suppose that ϱ is a monomorphism. If $\langle \mathbb{N}, \nu, \eta\rangle$ is a monad, then so is $\langle \mathbb{M}, \mu, \varepsilon\rangle$ and $\varrho : \mathbb{M} \Rightarrow \mathbb{N}$ is a morphism of monads.*

(b) *Suppose that ϱ is an epimorphism and that $\mathbb{M}$ preserves epimorphisms. If $\langle \mathbb{M}, \mu, \varepsilon\rangle$ is a monad, then so is $\langle \mathbb{N}, \nu, \eta\rangle$ and $\varrho : \mathbb{M} \Rightarrow \mathbb{N}$ is a morphism of monads.*

Corollary 2.6. *$\langle \mathbb{T}, \mathrm{flat}, \mathrm{sing}\rangle$ forms a monad on Pos^Ξ.*

Since our algebras are ordered it is natural to add meets (and joins) as operations. We start by defining a monad just for meets and then add it to our algebras via a standard construction based on so-called distributive laws. In this and the next section we only consider the monads $\mathbb{R}$ and $\mathbb{T}$. The more complicated case of $\mathbb{T}^\times$ will be dealt with separately in Section 4 below.

Definition 2.7. Let $A \in \mathsf{Pos}^\Xi$.

(a) For $X \subseteq A$, we write

$$\Uparrow X := \{\, a \in A \mid a \geq x \text{ for some } x \in X \,\}\,,$$

$$\text{and} \quad \Downarrow X := \{\, a \in A \mid a \leq x \text{ for some } x \in X \,\}\,.$$

For single elements $x \in A$, we omit the braces and simply write $\Uparrow x$ and $\Downarrow x$.

(b) The *(upward) power set* $\mathbb{U}A$ of A is the ordered set with domains

$$\mathbb{U}_\xi A := \{\, I \subseteq A_\xi \mid I \text{ is upwards closed} \,\}, \quad \text{for } \xi \in \Xi ,$$

and ordering

$$I \leq J \quad :\text{iff} \quad I \supseteq J , \quad \text{for } I, J \in \mathbb{U}_\xi A .$$

For a function $f : A \to B$, we define $\mathbb{U}f : \mathbb{U}A \to \mathbb{U}B$ by

$$\mathbb{U}f(I) := \Uparrow f[I] , \quad \text{for } I \in \mathbb{U}A .$$

(c) The *(downward) power set* $\mathbb{D}A$ of A is the ordered set with domains

$$\mathbb{D}_\xi A := \{\, I \subseteq A_\xi \mid I \text{ is downwards closed} \,\}, \quad \text{for } \xi \in \Xi ,$$

and ordering

$$I \leq J \quad :\text{iff} \quad I \subseteq J , \quad \text{for } I, J \in \mathbb{D}_\xi A .$$

For a function $f : A \to B$, we define $\mathbb{D}f : \mathbb{D}A \to \mathbb{D}B$ by

$$\mathbb{D}f(I) := \Downarrow f[I] , \quad \text{for } I \in \mathbb{D}A . \qquad \lrcorner$$

In the following we will state and prove most results only for the functor $\mathbb{U}$. The case of $\mathbb{D}$ can be handled in exactly the same way. First, let us note that it is straightforward to check that $\mathbb{U}$ forms a monad on Pos^Ξ.

Proposition 2.8. *The functor* $\mathbb{U} : \mathsf{Pos}^\Xi \to \mathsf{Pos}^\Xi$ *forms a monad where the multiplication*

$$\text{union} : \mathbb{U}\mathbb{U}A \to \mathbb{U}A : X \mapsto \bigcup X$$

is given by taking the union and the singleton function

$$\text{pt} : A \to \mathbb{U}A : a \mapsto \Uparrow\{a\}$$

is given by the principal filter operation.

Example 2.9. The algebras for the monad $\mathbb{U}$ are exactly those of the form $\langle A, \inf\rangle$ where A is a completely ordered set. A function $f : A \to B$ preserves arbitrary meets if, and only if, it is a morphism $\langle A, \inf\rangle \to \langle B, \inf\rangle$ of the corresponding $\mathbb{U}$-algebras. The same holds for $\mathbb{D}$ and suprema. $\lrcorner$

To show that $\mathbb{U}$ lifts to a monad on $\mathsf{Alg}(\mathbb{R})$, we use a standard technique based on distributive laws [Bec69]. Let us recall the basic definitions and results.

Definition 2.10. Let $\langle \mathbb{M}, \mu, \varepsilon\rangle$ and $\langle \mathbb{N}, \nu, \eta\rangle$ be monads. A natural transformation $\delta : \mathbb{M}\mathbb{N} \Rightarrow \mathbb{N}\mathbb{M}$ is a *distributive law* if

$$\begin{aligned} \delta \circ \mu &= \mathbb{N}\mu \circ \delta \circ \mathbb{M}\delta , & \delta \circ \varepsilon &= \mathbb{N}\varepsilon , \\ \delta \circ \mathbb{M}\nu &= \nu \circ \mathbb{N}\delta \circ \delta , & \delta \circ \mathbb{M}\eta &= \eta . \end{aligned}$$

$$
\begin{tikzcd}
\mathbb{M}\mathbb{M}\mathbb{N}A \arrow[r, "\mathbb{M}\delta"] \arrow[d, "\mu"'] & \mathbb{M}\mathbb{N}\mathbb{M}A \arrow[r, "\delta"] & \mathbb{N}\mathbb{M}\mathbb{M}A \arrow[d, "\mathbb{N}\mu"] \\
\mathbb{M}\mathbb{N}A \arrow[rr, "\delta"'] & & \mathbb{N}\mathbb{M}A
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbb{N}A \arrow[d, "\varepsilon"'] \arrow[dr, "\mathbb{N}\varepsilon"] & \\
\mathbb{M}\mathbb{N}A \arrow[r, "\delta"'] & \mathbb{N}\mathbb{M}A
\end{tikzcd}
$$

$$
\begin{tikzcd}
\mathbb{M}\mathbb{N}\mathbb{N}A \arrow[r, "\delta"] \arrow[d, "\mathbb{M}\nu"'] & \mathbb{N}\mathbb{M}\mathbb{N}A \arrow[r, "\mathbb{N}\delta"] & \mathbb{N}\mathbb{N}\mathbb{M}A \arrow[d, "\nu"] \\
\mathbb{M}\mathbb{N}A \arrow[rr, "\delta"'] & & \mathbb{N}\mathbb{M}A
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbb{M}A \arrow[d, "\mathbb{M}\eta"'] \arrow[dr, "\eta"] & \\
\mathbb{M}\mathbb{N}A \arrow[r, "\delta"'] & \mathbb{N}\mathbb{M}A
\end{tikzcd}
$$

$\lrcorner$

We can use distributive laws to lift a monad from the base category to the category of algebras.

Definition 2.11. Let $\langle \mathbb{M}, \mu, \varepsilon \rangle$ and $\langle \mathbb{N}, \nu, \eta \rangle$ be monads on some category $\mathcal{C}$ and let $\mathbb{V} : \mathsf{Alg}(\mathbb{M}) \to \mathcal{C}$ be the forgetful functor mapping an $\mathbb{M}$-algebra to its universe.

(a) We say that a monad $\langle \hat{\mathbb{N}}, \hat{\nu}, \hat{\eta} \rangle$ is a *lift* of $\mathbb{N}$ to the category of $\mathbb{M}$-algebras if

$$\mathbb{V} \circ \hat{\mathbb{N}} = \mathbb{N} \circ \mathbb{V}\,, \quad \mathbb{V}\hat{\nu} = \nu\,, \quad \mathbb{V}\hat{\eta} = \eta\,.$$

(b) The *Kleisli category* $\mathsf{Free}(\mathbb{N})$ of $\mathbb{N}$ is the full subcategory of $\mathsf{Alg}(\mathbb{N})$ induced by all free $\mathbb{N}$-algebras. The *free functor* $\mathbb{F}_{\mathbb{N}} : \mathcal{C} \to \mathsf{Free}(\mathbb{N})$ maps an object $C \in \mathcal{C}$ to the free $\mathbb{N}$-algebra generated by C, that is,

$$\begin{aligned} \mathbb{F}_{\mathbb{N}}C &:= \langle \mathbb{N}C, \nu \rangle\,, && \text{for objects } C \in \mathcal{C}\,, \\ \mathbb{F}_{\mathbb{N}}\varphi &:= \mathbb{N}\varphi\,, && \text{for } \mathcal{C}\text{-morphisms } \varphi : A \to B\,. \end{aligned}$$

(c) An *extension* of $\mathbb{M}$ to $\mathsf{Free}(\mathbb{N})$ is a monad $\langle \widehat{\mathbb{M}}, \hat{\mu}, \hat{\varepsilon} \rangle$ on $\mathsf{Free}(\mathbb{N})$ satisfying

$$\widehat{\mathbb{M}} \circ \mathbb{F}_{\mathbb{N}} = \mathbb{F}_{\mathbb{N}} \circ \mathbb{M}\,, \quad \hat{\mu} = \mathbb{F}_{\mathbb{N}}\mu\,, \quad \hat{\varepsilon} = \mathbb{F}_{\mathbb{N}}\varepsilon\,.$$

$\lrcorner$

Theorem 2.12 [Bec69]**.** *Let $\langle \mathbb{M}, \mu, \varepsilon \rangle$ and $\langle \mathbb{N}, \nu, \eta \rangle$ be monads. There exist bijections between the following objects:*

(1) *distributive laws $\delta : \mathbb{M}\mathbb{N} \Rightarrow \mathbb{N}\mathbb{M}$;*
(2) *liftings $\hat{\mathbb{N}}$ of $\mathbb{N}$ to the category of $\mathbb{M}$-algebras;*
(3) *extensions $\widehat{\mathbb{M}}$ of $\mathbb{M}$ to the Kleisli category* $\mathsf{Free}(\mathbb{N})$*;*
(4) *functions κ such that*
  (M1) *$\langle \mathbb{N}\mathbb{M}, \kappa, \eta \circ \varepsilon \rangle$ is a monad,*
  (M2) *the functions $\mathbb{N}\varepsilon$ and η induce morphisms of monads $\mathbb{N} \Rightarrow \mathbb{N}\mathbb{M}$ and $\mathbb{M} \Rightarrow \mathbb{N}\mathbb{M}$,*
  (M3) *κ satisfies the* middle unit law: $\kappa \circ \mathbb{N}(\varepsilon \circ \eta) = \mathrm{id}\,.$

3. Polynomial functors

It is not hard to manually find a distributive law between $\mathbb{U}$ and the monads $\mathbb{R}$ and $\mathbb{T}$, but it is not that much more difficult to prove a much more general result. The monads used in language theory, including $\mathbb{R}$, $\mathbb{T}$, and $\mathbb{T}^{\times}$, construct sets of labelled objects. The following definition captures the general form of such a monad.

Definition 3.1. A functor $\mathbb{F} : \mathsf{Pos}^{\Xi} \to \mathsf{Pos}^{\Xi}$ is *polynomial* if it is of the following form. For objects $A \in \mathsf{Pos}^{\Xi}$,

$$\mathbb{F}A = \sum_{i \in I} A^{D_i} ,$$

for some fixed sequence $(D_i)_{i \in I}$ of sets with $I, D_i \in \mathsf{Set}^{\Xi}$. Hence, an element of $\mathbb{F}A$ is of the form $\langle i, s \rangle$ with $i \in I$ and $s : D_i \to A$ sort-preserving. The sort of $\langle i, s \rangle$ is the sort of i. We usually omit the first component from the notation and simply write s. The set $\mathrm{dom}(s) := D_i$ is the called *domain* of s.

The ordering on $\mathbb{F}A$ is defined componentwise:

$$\langle i, s \rangle \leq \langle j, t \rangle \quad \text{iff} \quad i = j \quad \text{and} \quad s(v) \leq t(v) , \quad \text{for all } v \in \mathrm{dom}(s) .$$

Finally, $\mathbb{F}$ acts on morphisms by relabelling, that is,

$$\mathbb{F}f(s) := f \circ s : \mathrm{dom}(s) \to B , \quad \text{for } f : A \to B .$$ ⌟

Remark 3.2. (a) Note that the functors $\mathbb{R}$, $\mathbb{T}$, and $\mathbb{T}^{\times}$ are polynomial since

$$\mathbb{R}A = \sum_{g \text{ graph}} A^{\mathrm{dom}_0(g)} ,$$

where the sum ranges over all countable unlabelled graphs, i.e., the set $\mathbb{R}1$. The same holds for the other two functors.

(b) As one can see from the above expression, our notation for domains is not entirely consistent. What we call $\mathrm{dom}(s)$ for elements of a polynomial functor, is called $\mathrm{dom}_0(g)$ for graphs $g \in \mathbb{R}A$. ⌟

As observed in [SN] we can describe natural transformations between polynomial functors in the following way.

Lemma 3.3. *Let $\mathbb{F}X = \sum_{i \in I} X^{D_i}$ and $\mathbb{G}X = \sum_{j \in J} X^{E_j}$ be polynomial functors. There exists a one-to-one correspondence between natural transformations*

$$\alpha : \mathbb{F} \Rightarrow \mathbb{G}$$

and families of functions (in Set^{Ξ})

$$\alpha' : I \to J \quad \textit{and} \quad \alpha''_i : E_{\alpha'(i)} \to D_i , \quad \textit{for } i \in I .$$

This correspondence is given by the equation

$$\alpha(\langle i, s \rangle) = \langle \alpha'(i), t \rangle \quad \textit{with} \quad t(v) = s(\alpha''_i(v)) , \quad \textit{for } v \in E_{\alpha'(i)} .$$

Proof. The above equations induce a function mapping α', α''_i to α. This function is clearly injective. Hence, it remains to show surjectivity. Let $\alpha : \mathbb{F} \Rightarrow \mathbb{G}$ be a natural transformation. We start by recovering the function $\alpha' : I \to J$. Let 1 be a set with exactly 1 element $*_\xi$ of each sort ξ. Then $1^{D_i} = 1^{E_j}$ is a 1-element set. Hence, there are bijections between $\mathbb{F}1$ and I and between $\mathbb{G}1$ and J. In particular, the component $\alpha_1 : \mathbb{F}1 \to \mathbb{G}1$ of α induces a function $\alpha' : I \to J$. Given some set A, let $u : A \to 1$ be the unique function. For $\langle i, s \rangle \in \mathbb{F}A$ it follows that

$$\mathbb{G}u(\alpha_A(\langle i, s \rangle)) = \alpha_1(\mathbb{F}u(\langle i, s \rangle)) = \alpha_1(\langle i, *_\xi \rangle) = \langle \alpha'(i), *_\xi \rangle ,$$

where ξ is the sort of $\langle i, s \rangle$. This implies that

$$\alpha_A(\langle i, s \rangle) = \langle \alpha'(i), t \rangle , \quad \text{for some } t : E_{\alpha'(i)} \to A .$$

It thus remains to construct the functions $\alpha''_i : E_{\alpha'(i)} \to D_i$. We have just shown that $\alpha : \mathbb{F} \Rightarrow \mathbb{G}$ induces a natural transformation $X^{D_i} \Rightarrow X^{E_{\alpha'(i)}}$. It is therefore sufficient to show that every natural transformation $\beta : X^D \Rightarrow X^E$ (in Pos) corresponds to a function $\beta'' : E \to D$ (in Set^Ξ) such that

$$\beta(s) = t \quad \text{where} \quad t(v) = s(\beta''(v))\,.$$

We set

$$\beta'' := \beta_D(\mathrm{id}_D) \in D^E.$$

Given $s \in A^D$ and $v \in E$, it then follows that

$$\beta_A(s)(v) = \beta_A(s^D(\mathrm{id}))(v) = s^E(\beta_D(\mathrm{id}))(v) = s^E(\beta'')(v) = s(\beta''(v))\,,$$

as desired. □

We will need the following notation for relations between elements of polynomial functors.

Definition 3.4. Let $\mathbb{F} : \mathsf{Pos}^\Xi \to \mathsf{Pos}^\Xi$ be a functor, A, B sets, and $p : A \times B \to A$ and $q : A \times B \to B$ the two projections.

(a) The *lift* of a relation $\theta \subseteq A \times B$ is the relation $\theta^{\mathbb{F}} \subseteq \mathbb{F}A \times \mathbb{F}B$ defined by

$$s\ \theta^{\mathbb{F}}\ t \quad :\text{iff} \quad \mathbb{F}p(u) = s \text{ and } \mathbb{F}q(u) = t\,, \quad \text{for some } u \in \mathbb{F}\theta\,.$$

(b) We set $\simeq_{\mathrm{sh}} := \theta^{\mathbb{F}}$ for $\theta := A \times B$. If $s \simeq_{\mathrm{sh}} t$, we say that s and t have *the same shape.* ⌟

Remark 3.5. (a) For a polynomial functor $\mathbb{F}X = \sum_{i \in I} X^{D_i}$ and $s \in \mathbb{F}A$, $t \in \mathbb{F}B$, we have

$$s \simeq_{\mathrm{sh}} t \quad \text{iff} \quad s \in A^{D_i} \quad \text{and} \quad t \in B^{D_i}, \quad \text{for the same index } i \in I\,.$$

(This implies that s and t have the same sort, namely that of i.) Then

$$s\ \theta^{\mathbb{F}}\ t \quad \text{iff} \quad s \simeq_{\mathrm{sh}} t \quad \text{and} \quad s(v)\ \theta\ t(v)\,, \quad \text{for all } v \in \mathrm{dom}(s) = \mathrm{dom}(t)\,.$$

(b) In particular, two graphs $g, h \in \mathbb{R}A$ have the same shape if they have the same underlying graph, the same sort, and the same labelling with variables. Only the labelling with elements of A may differ. ⌟

The goal of this section is to derive a distributive law between certain polynomial functors and the monad $\mathbb{U}$. Our proof closely follows similar work from [Jac04, GPA21, BKS]. The differences are mainly technical and immaterial. The only part of the following that can be considered original seems to be

- the notion of linearity in Definition 3.14,
- Theorem 3.22, which states that the distributive law we present is unique, and
- Theorem 3.23, which states that there is no distributive law for non-linear monads.

Our existence proof is based on the characterisation in terms of extensions to the Kleisli category. We start by developing a few tools to construct such extensions. The first observation is that we can reduce the number of conditions we have to check.

Lemma 3.6. *Let $\langle \mathbb{M}, \mu, \varepsilon \rangle$ and $\langle \mathbb{N}, \nu, \eta \rangle$ be monads on $\mathcal{C}$ and let $\widehat{\mathbb{M}} : \mathsf{Free}(\mathbb{N}) \to \mathsf{Free}(\mathbb{N})$ be a functor satisfying*

- $\widehat{\mathbb{M}} \circ \mathbb{F}_{\mathbb{N}} = \mathbb{F}_{\mathbb{N}} \circ \mathbb{M}\,,$
- $\mathbb{F}_{\mathbb{N}}\mu \circ \widehat{\mathbb{M}}\widehat{\mathbb{M}}\varphi = \widehat{\mathbb{M}}\varphi \circ \mathbb{F}_{\mathbb{N}}\mu\,,$ *for every morphism $\varphi : A \to B$ of* $\mathsf{Free}(\mathbb{N})\,,$

- $\mathbb{F}_{\mathbb{N}}\varepsilon \circ \varphi = \widehat{\mathbb{M}}\varphi \circ \mathbb{F}_{\mathbb{N}}\varepsilon\,,$ *for every morphism* $\varphi : A \to B$ *of* $\mathsf{Free}(\mathbb{N})$,

then $\langle \widehat{\mathbb{M}}, \mathbb{F}_{\mathbb{N}}\mu, \mathbb{F}_{\mathbb{N}}\varepsilon \rangle$ *is an extension of* $\langle \mathbb{M}, \mu, \varepsilon \rangle$ *to* $\mathsf{Free}(\mathbb{N})$.

Proof. Our assumptions immediately imply that

$$\mathbb{F}_{\mathbb{N}}\mu : \widehat{\mathbb{M}}\widehat{\mathbb{M}} \Rightarrow \widehat{\mathbb{M}} \quad \text{and} \quad \mathbb{F}_{\mathbb{N}}\varepsilon : \mathrm{Id} \Rightarrow \widehat{\mathbb{M}}$$

are natural transformations. Hence, we only have to check the monad laws for $\langle \widehat{\mathbb{M}}, \mathbb{F}_{\mathbb{N}}\mu, \mathbb{F}_{\mathbb{N}}\varepsilon \rangle$.

$$\begin{aligned}
\mathbb{F}_{\mathbb{N}}\mu \circ \widehat{\mathbb{M}}\mathbb{F}_{\mathbb{N}}\mu &= \mathbb{F}_{\mathbb{N}}\mu \circ \mathbb{F}_{\mathbb{N}}\mathbb{M}\mu = \mathbb{F}_{\mathbb{N}}(\mu \circ \mathbb{M}\mu) = \mathbb{F}_{\mathbb{N}}(\mu \circ \mu) = \mathbb{F}_{\mathbb{N}}\mu \circ \mathbb{F}_{\mathbb{N}}\mu\,, \\
\mathbb{F}_{\mathbb{N}}\mu \circ \mathbb{F}_{\mathbb{N}}\varepsilon &= \mathbb{F}_{\mathbb{N}}(\mu \circ \varepsilon) = \mathrm{id}\,, \\
\mathbb{F}_{\mathbb{N}}\mu \circ \widehat{\mathbb{M}}\mathbb{F}_{\mathbb{N}}\varepsilon &= \mathbb{F}_{\mathbb{N}}\mu \circ \mathbb{F}_{\mathbb{N}}\mathbb{M}\varepsilon = \mathbb{F}_{\mathbb{N}}(\mu \circ \mathbb{M}\varepsilon) = \mathrm{id}\,.
\end{aligned}$$ □

Note that the action of $\widehat{\mathbb{M}}$ on objects is already completely determined by the requirement that $\widehat{\mathbb{M}} \circ \mathbb{F}_{\mathbb{N}} = \mathbb{F}_{\mathbb{N}} \circ \mathbb{M}$. Hence, we only have to find a suitable definition of $\widehat{\mathbb{M}}$ on morphisms $\varphi : \mathbb{N}A \to \mathbb{N}B$. For the functor $\mathbb{N} = \mathbb{U}$, we adapt a construction from [Jac04, Gar20, GPA21] based on the category of relations. Note that every morphism $\varphi : \mathbb{U}A \to \mathbb{U}B$ of $\mathbb{U}$-algebras is uniquely determined by its restriction $f : A \to \mathbb{U}B$ to A. The key idea is to use the following encoding of such functions.

Definition 3.7. (a) We denote the (sort-wise) power set of $A \in \mathsf{Pos}^{\Xi}$ by $\mathscr{P}(A) \in \mathsf{Pos}^{\Xi}$.

(b) A *span* is a pair of morphisms $A \leftarrow^{p} R \to^{q} B$ with the same domain. We call a span $A \leftarrow^{p} R \to^{q} B$ *injective* if

$$p(c) = p(c') \quad \text{and} \quad q(c) = q(c') \quad \text{implies} \quad c = c'\,,$$

and we call it *closed* if, for all $a \in A$, $b \in B$, and $c \in R$,

$$\begin{aligned}
&a \leq p(c) \quad \text{implies} \quad p^{-1}(a) \cap q^{-1}(q(c)) \neq \emptyset\,, \\
&b \geq q(c) \quad \text{implies} \quad q^{-1}(b) \cap p^{-1}(p(c)) \neq \emptyset\,.
\end{aligned}$$

(c) The function $f : A \to \mathscr{P}(B)$ (not necessarily monotone) *represented* by a span $A \leftarrow^{p} R \to^{q} B$ is given by

$$f(a) := q[p^{-1}(a)]\,, \quad \text{for } a \in A\,.$$

(d) The *graph* of a function $f : A \to \mathbb{U}B$ is the relation

$$G(f) := \{\, \langle a, b \rangle \in A \times B \mid b \in f(a) \,\}\,,$$

and the *representation* of f is the span $A \leftarrow G(f) \to B$ consisting of the two projections. ⌟

Lemma 3.8. *The correspondence between a function $A \to \mathbb{U}B$ and its representation forms a bijection between* (I) *the set of all functions $A \to \mathbb{U}B$ in Pos^{Ξ} and* (II) *the set of all spans $A \leftarrow R \to B$ that are injective and closed.*

Proof. Let $f : A \to \mathbb{U}B$ be a function with representation $A \leftarrow^{p} G(f) \to^{q} B$. This span is injective as every pair is uniquely determined by the values of its two components. To see that it is also closed, suppose that $b \geq q(c)$, for some $b \in B$ and $c \in G(f)$. By definition of $G(f)$, we have $c = \langle a', b' \rangle$ with $b' \in f(a')$. As $f(a')$ is upwards closed, $b \geq q(c) = b'$ implies $b \in f(a')$. Hence, $\langle a', b \rangle \in G(f)$ and

$$\langle a', b \rangle \in q^{-1}(b) \cap p^{-1}(a') = q^{-1}(b) \cap p^{-1}[p(c)] \neq \emptyset\,.$$

Similarly, suppose that $a \leq p(c)$. Then $c = \langle a', b' \rangle$ with $b' \in f(a')$ and $a \leq a'$. As f is monotone, it follows that $f(a) \supseteq f(a')$. In particular, $b' \in f(a)$. Hence, $\langle a, b' \rangle \in G(f)$ and

$$\langle a, b' \rangle \in p^{-1}(a) \cap q^{-1}(b') = p^{-1}(a) \cap q^{-1}(q(c)) \neq \emptyset\,.$$

Conversely, consider an injective, closed span $A \leftarrow^p R \rightarrow^q B$ and let $f : A \to \mathscr{P}(B)$ be the function it represents. We have to show that f is monotone and that $f(a)$ is upwards closed, for each $a \in A$. For monotonicity, let $a \leq a'$ and $b' \in f(a')$. We have to show that $b' \in f(a)$. By definition of f, there is some $c \in R$ with $p(c) = a'$ and $q(c) = b'$. Then $a \leq p(c)$ implies that there is some $d \in p^{-1}(a) \cap q^{-1}(q(c))$. Consequently, $b' = q(c) = q(d) \in q[p^{-1}(a)] = f(a)$.

To show that $f(a)$ is upwards closed, suppose that $b \geq b' \in f(a) = q[p^{-1}(a)]$. Then we can find some element $c \in R$ with $p(c) = a$ and $q(c) = b'$. Hence, $b \geq q(c)$ and closedness implies that we can find some element $c' \in q^{-1}(b) \cap p^{-1}[p(c)]$. It follows that $q(c') = b$ and $p(c') = p(c) = a$. Consequently, $b \in q[p^{-1}(a)] = f(a)$.

To conclude the proof, we have to show that these two operations are inverse to each other. Given a function $f : A \to \mathbb{U}B$, let g be the function represented by $A \leftarrow^p G(f) \rightarrow^q B$. Then

$$g(a) = q[p^{-1}(a)] = \{\, b \mid \langle a, b \rangle \in G(f) \,\} = \{\, b \mid b \in f(a) \,\} = f(a)\,.$$

Conversely, consider an injective, closed span $A \leftarrow^p R \rightarrow^q B$, let $f : A \to \mathbb{U}B$ be the function it represents, and let $A \leftarrow^u G(f) \rightarrow^v B$ be the representation of f. Then

$$\begin{aligned}
G(f) &= \{\, \langle a, b \rangle \mid b \in f(a) \,\} \\
&= \{\, \langle a, b \rangle \mid b \in q[p^{-1}(a)] \,\} \\
&= \{\, \langle a, b \rangle \mid c \in R\,,\ b = q(c)\,,\ p(c) = a \,\} \\
&= \{\, \langle p(c), q(c) \rangle \mid c \in R \,\}\,.
\end{aligned}$$

Since the span $A \leftarrow^p R \rightarrow^q B$ is injective, it follows that the function $\langle p, q \rangle : R \to G(f)$ is a bijection that commutes with the two projections. Thus, the two spans $A \leftarrow^p R \rightarrow^q B$ and $A \leftarrow^u G(f) \rightarrow^v B$ are isomorphic. ☐

We can compose spans by performing a pullback.

Lemma 3.9. *Let $f : A \to \mathbb{U}B$ and $g : B \to \mathbb{U}C$ be represented by, respectively, $A \leftarrow^p R \rightarrow^q B$ and $B \leftarrow^u S \rightarrow^v C$. Then the function*

$$\mathrm{union} \circ \mathbb{U}g \circ f : A \to \mathbb{U}C$$

is represented by $A \leftarrow^{p \circ k} T \rightarrow^{v \circ l} C$, where $R \leftarrow^k T \rightarrow^l S$ is the pullback of $R \rightarrow^q B \leftarrow^u S$.

$$\begin{tikzcd}
& & T \arrow[dl, "k"'] \arrow[dr, "l"] & & \\
& R \arrow[dl, "p"'] \arrow[dr, "q"] & & S \arrow[dl, "u"'] \arrow[dr, "v"] & \\
A & & B & & C
\end{tikzcd}$$

Proof. Note that the pullback in Pos^{Ξ} is given by

$$T = \{\, \langle r, s \rangle \mid q(r) = u(s) \,\}$$

and k and l are the respective projections. For $a \in A$, we therefore have

$$\begin{aligned}(\mathrm{union} \circ \mathbb{U}g \circ f)(a) &= \bigcup \{\, g(b) \mid b \in f(a) \,\} \\ &= \bigcup \big\{\, v[u^{-1}(b)] \bigm| b \in q[p^{-1}(a)] \,\big\} \\ &= \big\{\, c \in C \bigm| c = v(s)\,,\ u(s) = q(r)\,,\ p(r) = a \,\big\} \\ &= \big\{\, c \in C \bigm| c = v(s)\,,\ \langle r, s \rangle \in T\,,\ p(r) = a \,\big\} \\ &= \big\{\, v(l(t)) \bigm| t \in T\,,\ p(k(t)) = a \,\big\} \\ &= (v \circ l)[(p \circ k)^{-1}(a)]\,. \qquad \square\end{aligned}$$

It remains to prove that polynomial functors satisfy the conditions in Lemma 3.6. We start by taking a look at how such a functor operates on spans.

Lemma 3.10. *Let $\mathbb{M} : \mathsf{Pos}^{\Xi} \to \mathsf{Pos}^{\Xi}$ be a polynomial functor.*

(a) $\mathbb{M}$ *preserves injective and closed spans.*
(b) $\mathbb{M}$ *preserves pullbacks.*
(c) $s \simeq_{\mathrm{sh}} \mathbb{M}f(s)\,,$ *for all $s \in \mathbb{M}A\,,\ f : A \to B\,.$*

Proof. (a) Let $A \leftarrow^{p} R \rightarrow^{q} B$ be injective and closed and let $\mathbb{M}A \leftarrow^{\mathbb{M}p} \mathbb{M}R \rightarrow^{\mathbb{M}q} \mathbb{M}B$ be its image under $\mathbb{M}$.

For injectivity, consider elements $s, t \in \mathbb{M}R$. Then

$$\begin{aligned}&\mathbb{M}p(s) = \mathbb{M}p(t) \quad \text{and} \quad \mathbb{M}q(s) = \mathbb{M}q(t) \\ \Rightarrow\ & p(s(v)) = p(t(v)) \quad \text{and} \quad q(s(v)) = q(t(v))\,, \quad \text{for all } v\,, \\ \Rightarrow\ & s(v) = t(v)\,, \quad \text{for all } v\,, \\ \Rightarrow\ & s = t\,.\end{aligned}$$

For closedness, suppose that $s \geq \mathbb{M}q(t)$. Then

$$s(v) \geq q(t(v))\,, \quad \text{for all } v\,.$$

Hence, we can fix elements $c_v \in q^{-1}[s(v)] \cap p^{-1}[p(t(v))]$. Setting $t'(v) := c_v$, it follows that $t' \in \mathbb{M}R$ and

$$\begin{aligned}&q(t'(v)) = s(v) \quad \text{and} \quad p(t'(v)) = p(t(v))\,, \quad \text{for all } v\,, \\ \Rightarrow\ & \mathbb{M}q(t') = s \quad \text{and} \quad \mathbb{M}p(t') = \mathbb{M}p(t) \\ \Rightarrow\ & (\mathbb{M}q)^{-1}(s) \cap (\mathbb{M}p)^{-1}[\mathbb{M}p(t)] \neq \emptyset\,.\end{aligned}$$

Similarly, suppose that $s \leq \mathbb{M}p(t)$. Then $s(v) \leq p(t(v))$, for all v. Hence, we can fix elements $c_v \in p^{-1}[s(v)] \cap q^{-1}[q(t(v))]$. Setting $t'(v) := c_v$, it follows that $t' \in \mathbb{M}R$ and

$$\begin{aligned}&p(t'(v)) = s(v) \quad \text{and} \quad q(t'(v)) = q(t(v))\,, \quad \text{for all } v\,, \\ \Rightarrow\ & \mathbb{M}p(t') = s \quad \text{and} \quad \mathbb{M}q(t') = \mathbb{M}q(t) \\ \Rightarrow\ & (\mathbb{M}p)^{-1}(s) \cap (\mathbb{M}q)^{-1}[\mathbb{M}q(t)] \neq \emptyset\,.\end{aligned}$$

(b) Let $A \leftarrow^{p} P \rightarrow^{q} B$ be the pullback of $A \rightarrow^{f} C \leftarrow^{g} B$. Then

$$P = \{\, \langle a, b \rangle \mid f(a) = g(b) \,\}$$

and p and q are the respective projections. Similarly, the pullback of $\mathbb{M}A \to^{\mathbb{M}f} \mathbb{M}C \leftarrow^{\mathbb{M}g} \mathbb{M}B$ is

$$\begin{aligned} Q &:= \{ \langle s,t \rangle \mid \mathbb{M}f(s) = \mathbb{M}g(t) \} \\ &= \{ \langle s,t \rangle \mid f(s(v)) = g(t(v)) \text{ for all } v \} \\ &= \{ \langle s,t \rangle \mid \langle s(v), t(v) \rangle \in P \text{ for all } v \}. \end{aligned}$$

Consequently, the map $\langle \mathbb{M}p, \mathbb{M}q \rangle : \mathbb{M}(A \times B) \to \mathbb{M}A \times \mathbb{M}B$ induces a bijection between $\mathbb{M}P$ and Q.

(c) Setting $r := s$, $p := \mathrm{id}$, and $q := f$, we obtain $\mathbb{M}p(r) = s$ and $\mathbb{M}q(r) = \mathbb{M}f(s)$. □

Lemma 3.11. *Let $\mathbb{M}$ be a polynomial functor. If $A \leftarrow^p R \to^q B$ represents $f : A \to \mathbb{U}B$, then its image $\mathbb{M}A \leftarrow^{\mathbb{M}p} \mathbb{M}R \to^{\mathbb{M}q} \mathbb{M}B$ under $\mathbb{M}$ represents $F : \mathbb{M}A \to \mathbb{U}\mathbb{M}B$ where*

$$F(s) = \{ t \in \mathbb{M}B \mid t \in^{\mathbb{M}} \mathbb{M}f(s) \}.$$

Proof. We have shown in Lemma 3.10 that polynomial functors preserve injective closed spans. For $s \in \mathbb{M}A$, it therefore follows that

$$\begin{aligned} F(s) &= \mathbb{M}q\big[(\mathbb{M}p)^{-1}(s)\big] \\ &= \{ t \in \mathbb{M}B \mid r \in \mathbb{M}R,\ t = \mathbb{M}q(r),\ s = \mathbb{M}p(r) \} \\ &= \{ t \in \mathbb{M}B \mid r \in \mathbb{M}R,\ t(v) = q(r(v)),\ s(v) = p(r(v)),\ \text{for all } v \} \\ &= \{ t \in \mathbb{M}B \mid t(v) \in q[p^{-1}(s(v))],\ \text{for all } v \} \\ &= \{ t \in \mathbb{M}B \mid t(v) \in f(s(v)),\ \text{for all } v \} \\ &= \{ t \in \mathbb{M}B \mid t \in^{\mathbb{M}} \mathbb{M}f(s) \}. \end{aligned}$$ □

We obtain the following proof that every polynomial functor $\mathbb{M}$ on Pos^{Ξ} has an extension to $\mathsf{Free}(\mathbb{U})$.

Proposition 3.12. *Every polynomial functor $\mathbb{M}$ on Pos^{Ξ} induces a functor $\widehat{\mathbb{M}}$ on $\mathsf{Free}(\mathbb{U})$ satisfying*

$$\widehat{\mathbb{M}} \circ \mathbb{F}_{\mathbb{N}} = \mathbb{F}_{\mathbb{N}} \circ \mathbb{M}.$$

This functor maps a morphism $\varphi : \mathbb{U}A \to \mathbb{U}B$ to

$$\widehat{\mathbb{M}}\varphi(x) := \mathbb{M}q[(\mathbb{M}p)^{-1}[x]],$$

where $A \leftarrow^p R \to^q B$ is the span representing the morphism $\varphi \circ \mathrm{pt}$.

Proof. As we have already explained above, for objects we are forced to set

$$\widehat{\mathbb{M}}\langle \mathbb{U}A, \mathrm{union} \rangle := \langle \mathbb{U}\mathbb{M}A, \mathrm{union} \rangle.$$

For a morphism $\varphi : \langle \mathbb{U}A, \mathrm{union} \rangle \to \langle \mathbb{U}B, \mathrm{union} \rangle$ of free $\mathbb{U}$-algebras we define $\widehat{\mathbb{M}}\varphi$ as follows. Let $A \leftarrow^p G(\varphi) \to^q B$ be the representation of $\varphi \circ \mathrm{pt} : A \to \mathbb{U}B$, and let $\hat{\varphi} : \mathbb{M}A \to \mathbb{U}\mathbb{M}B$ be the function represented by the span $\mathbb{M}A \leftarrow^{\mathbb{M}p} \mathbb{M}G(\varphi) \to^{\mathbb{M}q} \mathbb{M}B$. Then we set

$$\widehat{\mathbb{M}}\varphi := \mathrm{union} \circ \mathbb{U}\hat{\varphi}.$$

We claim that this defines the desired functor $\widehat{\mathbb{M}}$.

First, let us prove that $\widehat{\mathbb{M}}$ is a functor $\mathsf{Free}(\mathbb{U}) \to \mathsf{Free}(\mathbb{U})$. Clearly, $\widehat{\mathbb{M}}$ maps free $\mathbb{U}$-algebras to free $\mathbb{U}$-algebras. Furthermore, by the above definition $\widehat{\mathbb{M}}\varphi$ is the free extension of $\hat{\varphi} : \mathbb{M}A \to \mathbb{U}\mathbb{M}B$ to a morphism $\mathbb{U}\mathbb{M}A \to \mathbb{U}\mathbb{M}B$ of $\mathbb{U}$-algebras. Hence, we only have to show that

$$\widehat{\mathbb{M}}(\varphi \circ \psi) = \widehat{\mathbb{M}}\varphi \circ \widehat{\mathbb{M}}\psi\,.$$

Let $\mathbb{M}B \leftarrow^{\mathbb{M}p} \mathbb{M}G(\varphi) \to^{\mathbb{M}q} \mathbb{M}C$ and $\mathbb{M}A \leftarrow^{\mathbb{M}u} \mathbb{M}G(\psi) \to^{\mathbb{M}v} \mathbb{M}B$ be the representations of $\hat{\varphi}$ and $\hat{\psi}$. By Lemma 3.9, the morphism

$$\mathrm{union} \circ \mathbb{U}\hat{\varphi} \circ \hat{\psi} : \mathbb{M}A \to \mathbb{U}\mathbb{M}C$$

is then represented by $\mathbb{M}A \leftarrow^{\mathbb{M}u \circ k} P \to^{\mathbb{M}q \circ l} \mathbb{M}C$ where $\mathbb{M}G(\psi) \leftarrow^{k} P \to^{l} \mathbb{M}G(\varphi)$ is the pullback of $\mathbb{M}G(\psi) \to^{\mathbb{M}v} \mathbb{M}B \leftarrow^{\mathbb{M}p} \mathbb{M}G(\varphi)$. Since $\mathbb{M}$ preserves pullbacks, we have $P = \mathbb{M}P'$, $k = \mathbb{M}k'$, and $k = \mathbb{M}k'$ where $G(\psi) \leftarrow^{k'} P' \to^{l'} G(\varphi)$ is the pullback of $G(\psi) \to^{v} B \leftarrow^{p} G(\varphi)$. Furthermore, it follows by Lemma 3.9 that $A \leftarrow^{u \circ k'} P' \to^{q \circ l'} C$ represents $\varphi \circ \psi$. Consequently, $\widehat{\mathbb{M}}(\varphi \circ \psi) \circ \mathrm{pt}$ is also represented by $\mathbb{M}A \leftarrow^{\mathbb{M}u \circ k} P \to^{\mathbb{M}q \circ l} \mathbb{M}C$ and we have

$$\widehat{\mathbb{M}}(\varphi \circ \psi) \circ \mathrm{pt} = \mathrm{union} \circ \mathbb{U}\hat{\varphi} \circ \hat{\psi} = \widehat{\mathbb{M}}\varphi \circ \widehat{\mathbb{M}}\psi \circ \mathrm{pt}\,.$$

As $\widehat{\mathbb{M}}(\varphi \circ \psi)$ and $\widehat{\mathbb{M}}\varphi \circ \widehat{\mathbb{M}}\psi$ are morphisms of $\mathbb{U}$-algebras, which are determined by their restriction to the range of pt, it follows that

$$\widehat{\mathbb{M}}(\varphi \circ \psi) = \widehat{\mathbb{M}}\varphi \circ \widehat{\mathbb{M}}\psi\,.$$

To conclude the proof, it remains to show that $\widehat{\mathbb{M}} \circ \mathbb{F}_{\mathbb{U}} = \mathbb{F}_{\mathbb{U}} \circ \mathbb{M}$. For objects $A \in \mathsf{Pos}^{\Xi}$, this is obvious from the definition. Hence, consider a function $f : A \to B$ and set $\varphi := \mathbb{U}f$. Let $A \leftarrow^{p} G(\mathbb{U}f) \to^{q} B$ be the span representing $\mathbb{U}f$. Then $\hat{\varphi} : \mathbb{M}A \to \mathbb{U}\mathbb{M}B$ is represented by $\mathbb{M}A \leftarrow^{\mathbb{M}p} \mathbb{M}G(\mathbb{U}f) \to^{\mathbb{M}q} \mathbb{M}B$. By Lemma 3.11, it follows that

$$\begin{aligned}
\widehat{\mathbb{M}}\mathbb{U}f(I) &= \bigcup \mathbb{U}\hat{\varphi}(I) \\
&= \bigcup {\Uparrow}\{\, \hat{\varphi}(s) \mid s \in I \,\} \\
&= \bigcup {\Uparrow}\big\{\, \{\, t \mid t \in^{\mathbb{M}} \mathbb{M}(\mathbb{U}f \circ \mathrm{pt})(s) \,\} \bigm| s \in I \,\big\} \\
&= {\Uparrow}\big\{\, t \bigm| s \in I\,,\ t \in^{\mathbb{M}} \mathbb{M}(\mathrm{pt} \circ f)(s) \,\big\} \\
&= {\Uparrow}\big\{\, t \bigm| s \in I\,,\ t(v) \in (\mathrm{pt} \circ f)(s(v)) \text{ for all } v \,\big\} \\
&= {\Uparrow}\big\{\, t \bigm| s \in I\,,\ t(v) \geq f(s(v)) \text{ for all } v \,\big\} \\
&= {\Uparrow}\big\{\, t \bigm| s \in I\,,\ t \geq \mathbb{M}f(s) \,\big\} \\
&= {\Uparrow}\big\{\, \mathbb{M}f(s) \bigm| s \in I \,\big\} \\
&= \mathbb{U}\mathbb{M}f(I)\,.
\end{aligned}$$

□

To find the desired distributive law for polynomial monads, it remains to prove the two remaining conditions of Lemma 3.6. To do so, we have to make additional assumptions on our monad: we require that the multiplication $\mathbb{M}\mathbb{M} \Rightarrow \mathbb{M}$ does not duplicate labels. We will call such monads *linear.* Before we can give the formal definition, we need to take a look at the special form the multiplication morphism for a polynomial functor takes.

Remark 3.13. Let $\langle \mathbb{M}, \mu, \varepsilon \rangle$ be a monad with a polynomial functor $\mathbb{M}X = \sum_{i \in I} X^{D_i}$. Note that the composition $\mathbb{M} \circ \mathbb{M}$ is also a polynomial functor. A straightforward computation yields

$$\mathbb{MM}X = \sum_{i \in I} \sum_{g : D_i \to I} X^{\sum_{v \in D_i} \operatorname{dom}(g(v))} .$$

Thus $\mathbb{MM}X = \sum_{j \in J} X^{E_j}$ where

$$J := \sum_{i \in I} I^{D_i} \quad \text{and} \quad E_{\langle i,g \rangle} := \sum_{v \in D_i} D_{g(v)} .$$

Note that the identity functor Id is polynomial, since

$$\mathrm{Id}(A) = \sum_{\xi \in \Xi} A^{1_\xi} ,$$

where 1_ξ is a set with a single element, which has sort ξ. Therefore, we can apply Lemma 3.3 to the natural transformations $\mu : \mathbb{MM} \Rightarrow \mathbb{M}$ and $\varepsilon : \mathrm{Id} \Rightarrow \mathbb{M}$ and we obtain induced maps

$$\begin{aligned} &\varepsilon' : \Xi \to I , && \varepsilon''_\xi : D_{\varepsilon'(\xi)} \to 1_\xi , && \text{for } \xi \in \Xi , \\ &\mu' : J \to I , && \mu''_j : D_{\mu'(j)} \to E_j , && \text{for } j \in J . \end{aligned}$$

With our conventions regarding polynomial functors, we can write the latter as

$$\mu''_s : \operatorname{dom}(\mu(s)) \to \sum_{v \in \operatorname{dom}(s)} \operatorname{dom}(s(v)) , \quad \text{for } s \in \mathbb{MM}A .$$

⌟

Definition 3.14. Let $\langle \mathbb{M}, \mu, \varepsilon \rangle$ be a monad where $\mathbb{M}$ is polynomial and let μ', μ''_j, ε', and ε''_j be the functions corresponding to the natural transformations $\mu : \mathbb{MM} \Rightarrow \mathbb{M}$ and $\varepsilon : \mathrm{Id} \Rightarrow \mathbb{M}$ as above. We call $\langle \mathbb{M}, \mu, \varepsilon \rangle$ *linear* if, for all indices j, the maps μ''_j are injective and the maps ε''_j are bijective. ⌟

Example 3.15. The monads $\mathbb{R}$ and $\mathbb{T}$ are linear since each vertex of $\operatorname{flat}(g)$ corresponds to exactly one vertex of exactly one component $g(v)$. The monad $\mathbb{T}^\times$ (defined below) on the other hand is not linear, since its multiplication duplicates labels: substituting $b(z)$ for x in $a(x, x)$ creates two copies of b. ⌟

Remark 3.16. Concerning terminology, the notion of a linear monad is not a priori related to that of a linear tree. But note that a submonad $\mathbb{T}^0$ of $\mathbb{T}^\times$ is linear in the above sense if, and only if, it is a submonad of $\mathbb{T}$. ⌟

For linear monads, we can now establish the missing identities. We start with a technical lemma.

Lemma 3.17. *Let $\langle \mathbb{M}, \mu, \varepsilon \rangle$ be a linear monad on Pos^Ξ.*

(a) $s \simeq_{\mathrm{sh}} t$ *and* $s(v) \simeq_{\mathrm{sh}} t(v)$, *for all* $v \in \operatorname{dom}(s)$, *implies* $\mu(s) \simeq_{\mathrm{sh}} \mu(t)$, *for* $s \in \mathbb{MM}A$ *and* $t \in \mathbb{MM}B$.

(b) $s \simeq_{\mathrm{sh}} \mu(t)$ *implies* $s = \mu(s')$, *for some* s' *with* $s' \simeq_{\mathrm{sh}} t$ *and* $s'(v) \simeq_{\mathrm{sh}} t(v)$.

Proof. Let $\mu''_j : \operatorname{dom}(\mu(s)) \to \sum_v \operatorname{dom}(s(v))$ be the injective map induced by μ.

(a) Let $p^* : \mathbb{M}A \to 1$, $q^* : \mathbb{M}B \to 1$, $p : A \to 1$, and $q : B \to 1$. By assumption, we have

$$\mathbb{M}p^*(s) = \mathbb{M}q^*(t) \quad \text{and} \quad \mathbb{M}p(s(v)) = \mathbb{M}q(t(v)) , \quad \text{for all } v .$$

For $w \in \operatorname{dom}(\mu(s))$ with $\mu''_j(w) = \langle v, u \rangle$ it follows that

$$p\big(\mu(s)(w)\big) = p\big(s(v)(u)\big) = q\big(t(v)(u)\big) = q\big(\mu(t)(w)\big) ,$$

as desired.

(b) Choose $s' \in \mathbb{MMA}$ such that $s' \simeq_{\mathrm{sh}} t$, $s'(v) \simeq_{\mathrm{sh}} t(v)$, for all v, and

$$s'(v)(u) := \begin{cases} s\big((\mu''_j)^{-1}(v,u)\big) & \text{if } \langle v,u\rangle \in \operatorname{rng} \mu''_j\,, \\ \text{arbitrary} & \text{otherwise}\,. \end{cases}$$

Then we have

$$s(w) = s'(v)(u)\,, \quad \text{for } \mu''_j(w) = \langle v,u\rangle\,,$$

which, by definition of μ''_j, implies that $\mu(s') = s$. □

Lemma 3.18. *Let $\langle \mathbb{M}, \mu, \varepsilon\rangle$ be a linear monad on Pos^{Ξ}, $\widehat{\mathbb{M}}$ its extension to $\mathsf{Free}(\mathbb{U})$ from Proposition 3.12, and let $\varphi : \mathbb{U}A \to \mathbb{U}B$ be a morphism of free $\mathbb{U}$-algebras.*

(a) $\widehat{\mathbb{M}}\varphi \circ \mathbb{U}\varepsilon = \mathbb{U}\varepsilon \circ \varphi$.

(b) $\widehat{\mathbb{M}}\varphi \circ \mathbb{U}\mu = \mathbb{U}\mu \circ \widehat{\mathbb{M}}\widehat{\mathbb{M}}\varphi$.

Proof. (a) Given a morphism $\varphi : \mathbb{U}A \to \mathbb{U}B$ between free $\mathbb{U}$-algebras, set $\varphi_0 := \varphi \circ \mathrm{pt}$ and let $A \leftarrow^p G(\varphi_0) \rightarrow^q B$ be the span representing it. For $I \in \mathbb{U}A$ it then follows that

$$\widehat{\mathbb{M}}\varphi(I) := \mathbb{M}q[(\mathbb{M}p)^{-1}[I]]\,.$$

Since $\mathbb{M}$ is linear we furthermore have

$$\begin{aligned} & \varepsilon(a) \leq \varepsilon(a') \\ \text{iff}\quad & a = \varepsilon(a)(v) \leq \varepsilon(a')(v) = a'\,, \quad \text{for all } v \in \operatorname{dom}(\varepsilon(a)) = \{*\}\,, \\ \text{iff}\quad & a \leq a' \end{aligned}$$

Hence,

$$\begin{aligned} \widehat{\mathbb{M}}\varphi(\mathbb{U}\varepsilon(I)) &= \widehat{\mathbb{M}}\big(\Uparrow\{\, \varepsilon(a) \mid a \in I\,\}\big) \\ &= \mathbb{M}q\big[(\mathbb{M}p)^{-1}\big[\Uparrow\{\, \varepsilon(a) \mid a \in I\,\}\big]\big] \\ &= \mathbb{M}q\big[\big\{\, s \in \mathbb{M}G(\varphi_0) \bigm| \mathbb{M}p(s) \geq \varepsilon(a)\,,\ a \in I\,\big\}\big] \\ &= \big\{\, \mathbb{M}q(\varepsilon(c)) \bigm| \varepsilon(c) \in \mathbb{M}G(\varphi_0)\,,\ \mathbb{M}p(\varepsilon(c)) \geq \varepsilon(a)\,,\ a \in I\,\big\} \\ &= \big\{\, \varepsilon(b) \bigm| \langle a', b\rangle \in G(\varphi_0)\,,\ a' \geq a\,,\ a \in I\,\big\} \\ &= \big\{\, \varepsilon(b) \bigm| b \in \varphi_0(a)\,,\ a \in I\,\big\} \\ &= \mathbb{U}\varepsilon(\operatorname{union}(\mathbb{U}\varphi_0(I))) \\ &= \mathbb{U}\varepsilon(\varphi(I))\,. \end{aligned}$$

(b) Given a morphism $\varphi : \mathbb{U}A \to \mathbb{U}B$ between free $\mathbb{U}$-algebras, set $\varphi_0 := \varphi \circ \mathrm{pt}$ and let $A \leftarrow^p G(\varphi_0) \rightarrow^q B$ be the span representing it. It then follows that

$$\begin{aligned} \widehat{\mathbb{M}}\varphi(I) &:= \mathbb{M}q[(\mathbb{M}p)^{-1}[I]]\,, && \text{for } I \in \mathbb{UM}A\,, \\ \widehat{\mathbb{M}}\widehat{\mathbb{M}}\varphi(I) &:= \mathbb{MM}q\big[(\mathbb{MM}p)^{-1}[I]\big]\,, && \text{for } I \in \mathbb{UMM}A\,. \end{aligned}$$

We start by proving that, for $r \in \mathbb{M}G(\varphi_0)$ and $s \in \mathbb{M}A$,

$$\mathbb{M}p(r) \geq s \quad \text{implies} \quad \mathbb{M}p(r') = s \text{ for some } r' \leq r\,.$$

To see this, consider a position $v \in \operatorname{dom}(r)$. Then

$$r(v) = \langle a_v, b_v\rangle \in G(\varphi_0) \quad \text{and} \quad s(v) = a'_v \leq a_v\,.$$

Hence, $b_v \in f(a_v) \geq f(a'_v)$ implies $b_v \in f(a'_v)$. Setting

$$r' \simeq_{\mathrm{sh}} r \quad \text{and} \quad r'(v) := \langle a'_v, b_v \rangle\,,$$

we obtain $r' \in \mathbb{M}G(\varphi_0)$, $r' \leq r$, and $\mathbb{M}p(r') = s$.

To conclude the proof, note that

$$\begin{aligned}
\mathbb{U}\mu(\widehat{\mathbb{M}}\widehat{\mathbb{M}}\varphi(I)) &= \mathbb{U}\mu\big(\mathbb{M}\mathbb{M}q\big[(\mathbb{M}\mathbb{M}p)^{-1}[I]\big]\big) \\
&= \mathbb{U}\mu\big(\mathbb{M}\mathbb{M}q\big[\{\, r \in \mathbb{M}\mathbb{M}G(\varphi_0) \mid \mathbb{M}\mathbb{M}p(r) \in I \,\}\big]\big) \\
&= \mathbb{U}\mu\big(\big\{\, t \in \mathbb{M}\mathbb{M}B \bigm| \langle s(v)(u), t(v)(u)\rangle \in G(\varphi_0)\,,\ s \in I \,\big\}\big) \\
&= {\Uparrow}\big\{\, \mu(t) \bigm| r(v)(u) \in G(\varphi_0)\,,\ r(v)(u) = \langle s(v)(u), t(v)(u)\rangle\,,\ s \in I \,\big\} \\
&= {\Uparrow}\big\{\, \mu(t) \bigm| r \in \mathbb{M}\mathbb{M}G(\varphi_0)\,,\ \mathbb{M}\mathbb{M}p(r) = s\,,\ \mathbb{M}\mathbb{M}q(r) = t\,,\ s \in I \,\big\} \\
&= {\Uparrow}\big\{\, \mu(\mathbb{M}\mathbb{M}q(r)) \bigm| r \in \mathbb{M}\mathbb{M}G(\varphi_0)\,,\ \mathbb{M}p(\mu(r)) = \mu(s)\,,\ s \in I \,\big\} \\
&= {\Uparrow}\big\{\, \mathbb{M}q(\mu(r)) \bigm| r \in \mathbb{M}\mathbb{M}G(\varphi_0)\,,\ \mathbb{M}p(\mu(r)) = \mu(s)\,,\ s \in I \,\big\} \\
&= {\Uparrow}\big\{\, \mathbb{M}q(r') \bigm| r' \in \mathbb{M}G(\varphi_0)\,,\ \mathbb{M}p(r') = \mu(s)\,,\ s \in I \,\big\} \\
&= {\Uparrow}\big\{\, \mathbb{M}q(r') \bigm| r' \in \mathbb{M}G(\varphi_0)\,,\ \mathbb{M}p(r') \geq \mu(s)\,,\ s \in I \,\big\} \\
&= {\Uparrow}\mathbb{M}q\big[(\mathbb{M}p)^{-1}[\mathbb{U}\mu(I)]\big] \\
&= \mathbb{M}q\big[(\mathbb{M}p)^{-1}[\mathbb{U}\mu(I)]\big] \\
&= \widehat{\mathbb{M}}\varphi(\mathbb{U}\mu(I))\,,
\end{aligned}$$

where we have used implicit universal quantification over u and v and where the eight step follows by Lemma 3.17 (b) and the nineth step by the above claim. ☐

Theorem 3.19. *Let $\mathbb{M}$ be a linear monad on Pos^{Ξ}. The functions $\mathrm{dist}_A : \mathbb{M}\mathbb{U}A \to \mathbb{U}\mathbb{M}A$ defined by*

$$\mathrm{dist}_A(t) := \{\, s \in \mathbb{M}A \mid s \in^{\mathbb{M}} t \,\}$$

form a distributive law $\mathbb{M}\mathbb{U} \Rightarrow \mathbb{U}\mathbb{M}$.

Proof. By (the proof of) Theorem 2.12, we can obtain the desired distributive law from an extension $\widehat{\mathbb{M}}$ of $\mathbb{M}$ to $\mathsf{Free}(\mathbb{U})$ by setting

$$\delta := \mathbb{V}\widehat{\mathbb{M}}\mathrm{id} \circ \mathrm{pt}\,,$$

where $\mathbb{V} : \mathsf{Free}(\mathbb{U}) \to \mathsf{Pos}^{\Xi}$ is the forgetful functor. Note that the span representing the identity $\mathrm{id} : \mathbb{U}A \to \mathbb{U}A$ is $A \leftarrow^{\mathrm{id}} A \to^{\mathrm{pt}} \mathbb{U}A$. For $t \in \mathbb{M}\mathbb{U}A$, it therefore follows that

$$\begin{aligned}
\delta(t) &= \widehat{\mathbb{M}}\mathrm{id}(\mathrm{pt}(t)) \\
&= \mathbb{M}\mathrm{id}[(\mathbb{M}\mathrm{pt})^{-1}[{\Uparrow}\{t\}]] \\
&= \{\, \mathbb{M}\mathrm{id}(s) \mid \mathbb{M}\mathrm{pt}(s) \geq t \,\} \\
&= \{\, s \mid \mathrm{pt}(s(v)) \subseteq t(v) \text{ for all } v \,\} \\
&= \{\, s \mid s(v) \in t(v) \text{ for all } v \,\} \\
&= \{\, s \mid s \in^{\mathbb{M}} t \,\}\,.
\end{aligned}$$

☐

Corollary 3.20. *The functions* dist *from above form distributive laws $\mathbb{T}\mathbb{U} \Rightarrow \mathbb{U}\mathbb{T}$ and $\mathbb{R}\mathbb{U} \Rightarrow \mathbb{U}\mathbb{R}$.*

Remark 3.21. The distributive law dist above was first stated in [Jac04] for functors (not monads) on Set preserving weak pullbacks. Our proof follows basically the same lines, except that we cannot use the algebra of relations for Pos, so we have to resort to direct calculations in several places. See also [GPA21, BKS] for similar arguments. ⌟

We can strengthen this theorem in two ways: (I) the distributive law dist is unique and (II) there is no distributive law for non-linear monads. We start with the former.

Theorem 3.22. *Let $\mathbb{M}$ be a polynomial monad on Pos^{Ξ} and $\delta : \mathbb{MU} \Rightarrow \mathbb{UM}$ a distributive law. Then $\delta = \mathrm{dist}$.*

Proof. ($\supseteq$) Since δ is monotone, we have

$$\begin{aligned}
\delta(t) &\leq \inf \{\, \delta(s) \mid s \geq t \,\} \\
&\leq \inf \big\{\, \delta(\mathbb{M}\mathrm{pt}(r)) \bigm| \mathbb{M}\mathrm{pt}(r) \geq t \,\big\} \\
&= \inf \big\{\, \mathrm{pt}(r) \bigm| \mathbb{M}\mathrm{pt}(r)(v) \geq t(v) \text{ for all } v \,\big\} \\
&= \inf \big\{\, \mathrm{pt}(r) \bigm| \mathrm{pt}(r(v)) \geq t(v) \text{ for all } v \,\big\} \\
&= \bigcup \big\{\, \mathrm{pt}(r) \bigm| \mathrm{pt}(r(v)) \subseteq t(v) \text{ for all } v \,\big\} \\
&= \bigcup \big\{\, \mathrm{pt}(r) \bigm| r(v) \in t(v) \text{ for all } v \,\big\} \\
&= \Uparrow\{\, r \mid r \in^{\mathbb{M}} t \,\} \\
&= \mathrm{dist}(t)\,.
\end{aligned}$$

($\subseteq$) Suppose that $s \in \delta(t)$ for $t \in \mathbb{MU}A$. To prove that $s \in \mathrm{dist}(t)$ it is sufficient to show that $s(v) \in t(v)$, for all v. Hence, fix $v \in \mathrm{dom}(t)$ and let $\theta : A \to [2]$ be the map with

$$\theta(a) := \begin{cases} 1 & \text{if } a \in t(v)\,, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathbb{MU}\theta(t)(v) = \mathbb{U}\theta(t(v)) = \{1\}$. Since $[2]$ is well-ordered, we can find some $r \in \mathbb{M}[2]$ such that $\mathbb{MU}\theta(t) = \mathbb{M}\mathrm{pt}(r)$. It follows that

$$\mathbb{UM}\theta(\delta(t)) = \delta(\mathbb{MU}\theta(t)) = \delta(\mathbb{M}\mathrm{pt}(t)) = \mathrm{pt}(r)\,.$$

Consequently,

$$\theta(s(v)) = \mathbb{M}\theta(s)(v) \geq r(v) = 1 \quad \text{implies} \quad s(v) \in t(v)\,. \qquad \square$$

As a consequence, we obtain the following strengthening of Theorem 3.19.

Theorem 3.23. *Let $\langle \mathbb{M}, \mu, \varepsilon \rangle$ be a polynomial monad on Pos^{Ξ}. There exists a distributive law $\delta : \mathbb{MU} \Rightarrow \mathbb{UM}$ if, and only if, $\mathbb{M}$ is linear.*

Proof. ($\Leftarrow$) has already been proved in Theorem 3.19.

($\Rightarrow$) Suppose that $\mathbb{M}$ is not linear and let μ', μ''_j, ε', and ε''_j be the functions corresponding to the natural transformations $\mu : \mathbb{MM} \Rightarrow \mathbb{M}$ and $\varepsilon : \mathrm{Id} \Rightarrow \mathbb{M}$ as in the definition of linearity. By Theorem 3.22, it is sufficient to show that dist is not a distributive law. For a contradiction, suppose otherwise.

By assumption, there is some index j such that μ''_j is not injective or ε''_j not bijective. First, assume that $\mu''_j : D_{\mu'(j)} \to E_j$ is not injective, for some index j. Then there are two positions $u, v \in D_{\mu'(j)}$ with $\mu''_j(u) = \mu''_j(v)$. Set $w := \mu''_j(u)$, Let A be a set with at least two elements a and b of the same sort as these positions (and trivial ordering), and let $s \in \mathbb{MMUM}A$ be such that $\mathrm{dom}(s) = E_j$,

$$s(w) := \{\varepsilon(a), \varepsilon(b)\} \quad \text{and} \quad s(x) = \{\varepsilon(c_x)\}\,, \quad \text{for all } x \neq w\,.$$

By Theorem 2.12, $\langle \mathbb{UM}A, \mathbb{U}\mu \circ \mathrm{dist}\rangle$ is an $\mathbb{M}$-algebra with product $\pi := \mathbb{U}\mu \circ \mathrm{dist}$. Note that

$$\begin{aligned}
&\big\{\, \langle t(u), t(v)\rangle \bigm| t \in \mathrm{dist}(\mu(s)) \,\big\} \\
={}& \big\{\, \langle t(u), t(v)\rangle \bigm| t \in^{\mathbb{M}} \mu(s) \,\big\} \\
={}& \big\{\, \langle p, q\rangle \bigm| p \in \mu(s)(u),\ q \in \mu(s)(v) \,\big\} \\
={}& \big\{\, \langle p, q\rangle \bigm| p, q \in s(w) \,\big\} \\
={}& \big\{\langle \varepsilon(a), \varepsilon(a)\rangle,\ \langle \varepsilon(a), \varepsilon(b)\rangle,\ \langle \varepsilon(b), \varepsilon(a)\rangle,\ \langle \varepsilon(b), \varepsilon(b)\rangle\big\}\,.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\big\{\, t(w) \bigm| t \in \mathrm{dist}(\mathbb{M}\pi(s)) \,\big\} \\
={}& \big\{\, t(w) \bigm| t \in^{\mathbb{M}} \mathbb{M}\pi(s) \,\big\} \\
={}& \big\{\, p \bigm| p \in \pi(s(w)) \,\big\} \\
={}& \big\{\, p \bigm| p \in \mathbb{U}\mu(\mathrm{dist}(s(w))) \,\big\} \\
={}& \big\{\mu(\varepsilon(a)), \mu(\varepsilon(b))\big\} \\
={}& \{a, b\}\,.
\end{aligned}$$

Since every $t \in \mathrm{dist}(\mu(s))$ is of the form $t = \mathbb{M}\varepsilon(t_0)$, for some $t_0 \in \mathbb{M}A$, it follows that

$$\begin{aligned}
&\big\{\, \langle t(u), t(v)\rangle \bigm| t \in \mathbb{U}\mu(\mathrm{dist}(\mu(s))) \,\big\} \\
={}& \big\{\, \big\langle \mu(t(u)), \mu(t(v))\big\rangle \bigm| t \in \mathrm{dist}(\mu(s)) \,\big\} \\
={}& \big\{\langle a, a\rangle,\ \langle a, b\rangle,\ \langle b, a\rangle,\ \langle b, b\rangle\big\}\,.
\end{aligned}$$

But

$$\begin{aligned}
&\big\{\, \langle t(u), t(v)\rangle \bigm| t \in \mathbb{U}\mu(\mathrm{dist}(\mathbb{M}\pi(s))) \,\big\} \\
={}& \big\{\, \langle t(w), t(w)\rangle \bigm| t \in \mathrm{dist}(\mathbb{M}\pi(s)) \,\big\} \\
={}& \big\{\langle a, a\rangle,\ \langle b, b\rangle\big\}\,.
\end{aligned}$$

Thus $\pi(\mu(s)) \neq \pi(\mathbb{M}\pi(s))$. A contradiction.

It remains to consider the case where ε''_j is not bijective, for some j. Then there is some sort ξ such that, for every element a of sort ξ, the domain $D := \mathrm{dom}(\varepsilon(a))$ is either empty or of size at least 2. Let $A := \{a, b\}$ be a set with two elements of sort ξ and the trivial ordering. If D is empty, we set $s := \varepsilon(a)$ and $t := \varepsilon(b)$. Then

$$\mathrm{dom}(\varepsilon(s)) = \emptyset = \mathrm{dom}(\varepsilon(t)) \quad \text{implies} \quad \varepsilon(s) = \varepsilon(t)\,.$$

Hence, $s = \mu(\varepsilon(s)) = \mu(\varepsilon(t)) = t$. A contradiction.

Consequently, D must have at least two elements and $\varepsilon(a) : D \to \{a\}$ is the constant function with value a. Note that $A \in \mathbb{U}A$ and

$$\begin{aligned}
\mathbb{U}\varepsilon(A) &= \big\{\varepsilon(a), \varepsilon(b)\big\} = \{\, s \mid s : D \to \{a, b\} \text{ a constant function} \,\}\,, \\
\mathrm{dist}(\varepsilon(A)) &= \{\, s \mid s \in^{\mathbb{M}} \varepsilon(A) \,\} = \{\, s \mid s : D \to \{a, b\} \,\}\,.
\end{aligned}$$

As $|D| > 1$, there exist non-constant functions $D \to \{a, b\}$. This implies that $\mathrm{dist} \circ \varepsilon \neq \mathbb{U}\varepsilon$, a violation of one of the axioms of a distributive law. □

Remark 3.24. (a) We did not make essential use of the fact that we are working with ordered sets. All results of this section also hold in the category Set^{Ξ}.

(b) In the literature one can find many cases where there is no distributive law between some variant of the power-set monad and some other monad. In particular, there is no such law between the power-set monad and itself. As a workaround there has been a lot of recent work (see, e.g., [Gar20, GPA21]) on so-called *weak distributive laws* which satisfy the axioms for a distributive law, except possibly for $\delta \circ \varepsilon = \mathbb{N}\varepsilon$. A closer look at the proofs above reveals that our results also hold for weak distributive laws if we replace linearity with the weaker condition that only the functions μ''_j are injective. If we call such a monad *weakly linear* it follows in particular that there is a weak distributive law $\delta : \mathbb{M}\mathbb{U} \Rightarrow \mathbb{U}\mathbb{M}$ if, and only if, $\mathbb{M}$ is weakly linear.

(c) In light of the above theorem, it is unsurprising that all known distributive laws for variants of the power-set monad require some form of linearity, although it is frequently expressed in terms of which equations the free algebra satisfies, instead of using properties of the monad multiplication.

For instance, there is a distributive law [MM07] in Set between so-call 'commutative monads' (like the power-set monad) and finitary term monads (which are linear in our sense). Similarly, there is a distributive law [MM08] between certain monads and quotients of finitary term monads by linear equations (i.e., term equations where every variable appears exactly once on each side).

In [ZM22] a variety of non-existence results for distributive laws between quotients of finitary term monads is proved. In many of the cases, one of the assumptions is that there is some term s satisfying the equation $s(x, \dots, x) = x$ (which is non-linear).

It seems that much of the existing theory could be unified if the results of this section (which also apply to monads that are non-finitary) could be generalised from linear polynomial monads to suitable 'linear' quotients of such monads. ⌟

4. Non-linear trees

It is time to properly define our third monad, that of non-linear trees, and to prove its limited compatibility with the power-set monad. Unfortunately, this turns out to be much more complex than the case of linear trees. In fact, as we have seen in Theorem 3.23, there does not exist a distributive law between $\mathbb{T}^\times$ and $\mathbb{U}$. We will therefore forego distributive laws and directly prove the existence of a lift of $\mathbb{U}$ to the class of free $\mathbb{T}^\times$-algebras, a partial result that is sufficient for many applications. We start by defining the monad structure of $\mathbb{T}^\times$.

Definition 4.1. (a) We denote the *unravelling* (in the usual graph-theoretic sense) of a graph $g \in \mathbb{R}_\xi A$ by $\mathrm{un}_0(g) \in \mathbb{R}_\xi A$. That is, $\mathrm{un}_0(g)$ is the graph whose vertices consist of all finite paths of g that start at the root and there is an edge between two such paths if the second one is the corresponding prolongation of the first one.

(b) We define $\mathrm{flat}^\times : \mathbb{T}^\times\mathbb{T}^\times A \to \mathbb{T}^\times A$ and $\mathrm{sing}^\times : A \to \mathbb{T}^\times A$ by

$$\mathrm{flat}^\times := \mathrm{un}_0 \circ \mathrm{flat} \quad \text{and} \quad \mathrm{sing}^\times := \mathrm{sing}\,. \qquad ⌟$$

This gives us the desired monad structure for $\mathbb{T}^\times$. The proof is straightforward.

Lemma 4.2. $\langle \mathbb{T}^\times, \mathrm{flat}^\times, \mathrm{sing}^\times \rangle$ *is a monad.*

In contrast to $\mathbb{T}$, the monad $\mathbb{T}^\times$ is *not* a submonad of $\mathbb{R}$. Instead it is a *quotient.*

Lemma 4.3. $\mathrm{un}_0 : \mathbb{R} \Rightarrow \mathbb{T}^\times$ *is a morphism of monads.*

Proof. We have to check that

$$\mathrm{sing}^\times = \mathrm{un}_0 \circ \mathrm{sing} \quad \text{and} \quad \mathrm{flat}^\times \circ \mathrm{un}_0 \circ \mathbb{R}\mathrm{un}_0 = \mathrm{un}_0 \circ \mathrm{flat}\,.$$

The first equation immediately follows form the fact that $\mathrm{un}_0(\mathrm{sing}(a)) = \mathrm{sing}(a)$. For the second one, note that the vertices of $\mathrm{un}_0(\mathrm{flat}(g))$ correspond to the finite paths of $\mathrm{flat}(g)$, while those of $\mathrm{un}_0(\mathrm{flat}(\mathrm{un}_0(\mathbb{R}\mathrm{un}_0(g))))$ correspond to those of $\mathrm{flat}(\mathrm{un}_0(\mathbb{R}\mathrm{un}_0(g)))$. Furthermore, every path α in a graph of the form $\mathrm{flat}(h)$ corresponds to a path $(v_n)_n$ of h and a family of paths β_n of $h(v_n)$ such that α can be identified with the concatenation $\beta_0\beta_1\dots$. Finally, a path in $\mathrm{un}_0(h)$ is the same as a path in h. Consequently, each path of $\mathrm{flat}(\mathrm{un}_0(\mathbb{R}\mathrm{un}_0(g)))$ corresponds to (i) a path of g together with (ii) a family of paths in some components $g(v)$ as above. This correspondence induces a bijection between

$$\mathrm{dom}(\mathrm{un}_0(\mathrm{flat}(g))) \quad \text{and} \quad \mathrm{dom}(\mathrm{un}_0(\mathrm{flat}(\mathrm{un}_0(\mathbb{R}\mathrm{un}_0))))\,.$$

As this bijection preserves the labelling it follows that

$$\mathrm{un}_0(\mathrm{flat}(g)) = \mathrm{un}_0(\mathrm{flat}(\mathrm{un}_0(\mathbb{R}\mathrm{un}_0)))\,. \qquad \square$$

The fact that there is no distributive law for $\mathbb{T}^\times$ follows directly from Theorem 3.23 since $\mathbb{T}^\times$ is not linear. This means that our main goal is unreachable. But having a distributive law between $\mathbb{T}^\times$ and $\mathbb{U}$ would be very useful. For instance, it is needed when introducing regular expressions for infinite trees. Therefore we will try to find a useable workaround, something weaker than an actual distributive law that nevertheless covers the applications we have in mind. The rest of this section is meant to get an overview over our options in this regard, and to probe the dividing line between the possible and the impossible.

Remark 4.4. We have already mentioned above that, for cases where there is no distributive law, there is the notion of a weak distributive law which often can be used instead. Unfortunately, this does not work in our case since the problem above is the monad multiplication, not the unit. ($\mathbb{T}^\times$ is not even weakly linear.) ⌟

4.1. Infinite sorts. We start with some technical remarks considering sorts. Below we will need to deal with trees with infinitely many different variables, that is, we have to work in the category $\mathsf{Pos}^{\mathscr{P}(X)}$ instead of Pos^{Ξ}. It is straightforward to extend the monads $\mathbb{R}$, $\mathbb{T}$, and $\mathbb{T}^\times$ to this more general setting. Hence, let us consider the following situation: we are given two sets $\Delta \subseteq \Gamma$ of sorts and a monad $\mathbb{M}$ on Pos^Γ. The following technical tools allow us to translate between the associated categories Pos^Δ and Pos^Γ.

Definition 4.5. Let $\Delta \subseteq \Gamma$ be sets of sorts.

(a) The *extension* of $A = (A_\xi)_{\xi\in\Delta} \in \mathsf{Pos}^\Delta$ to Pos^Γ is the set $A^\uparrow \in \mathsf{Pos}^\Gamma$ defined by

$$A^\uparrow_\xi := \begin{cases} A_\xi & \text{if } \xi \in \Delta\,, \\ \emptyset & \text{otherwise}\,. \end{cases}$$

(b) The *restriction* of $A = (A_\xi)_{\xi\in\Gamma} \in \mathsf{Pos}^\Gamma$ to Pos^Δ is the set $A|_\Delta := (A_\xi)_{\xi\in\Delta}$. Similarly, for a function $f : A \to B$ in Pos^Γ, we denote by $f|_\Delta : A|_\Delta \to B|_\Delta$ the restriction to Δ. Finally, for an $\mathbb{M}$-algebra $\mathfrak{A} = \langle A, \pi\rangle$, we set

$$\mathfrak{A}|_\Delta := \langle A|_\Delta, \pi|_\Delta \circ (\mathbb{M}i)|_\Delta\rangle\,,$$

where $i : (A|_\Delta)^\uparrow \to A$ is the inclusion map.

(c) The *restriction* of a functor $\mathbb{M} : \mathsf{Pos}^\Gamma \to \mathsf{Pos}^\Gamma$ to Pos^Δ is the functor $\mathbb{M}|_\Delta : \mathsf{Pos}^\Delta \to \mathsf{Pos}^\Delta$ defined by

$$\mathbb{M}|_\Delta A := (\mathbb{M}(A^\uparrow))|_\Delta\,. \qquad \lrcorner$$

Example 4.6. Let $\Delta := \{\emptyset, \{x\}\} \subseteq \Xi$, for some fixed $x \in X$. The monad $\mathbb{T}|_\Delta$ is isomorphic to the functor

$$\mathbb{M}\langle X_0, X_1\rangle = \langle X_1^* X_0 + X_1^\omega,\ X_1^+\rangle$$

(up to renaming of the sorts for readability) whose algebras are (ordered) ω-semigroups $\langle S_0, S_1, \pi\rangle$. The restriction $\mathbb{M}|_{\{1\}} X_1 = X_1^+$ is the monad for (ordered) semigroups, while $\mathbb{M}|_{\{0\}} X_0 = X_0$ is just the identity monad. Given an ω-semigroup $\mathfrak{S} = \langle S_0, S_1, \pi\rangle$, the corresponding restrictions are the associated semigroup $\mathfrak{S}|_{\{1\}} = \langle S_1, \pi_1\rangle$ and the set $\mathfrak{S}|_{\{0\}} = \langle S_0, \mathrm{id}\rangle$. ⌟

Let us quickly check that these definitions make sense.

Lemma 4.7. *Let $\langle \mathbb{M}, \mu, \varepsilon\rangle$ be a monad on Pos^Γ.*

(a) *$\mathbb{M}|_\Delta$ forms a monad with multiplication $(\mu \circ \mathbb{M}i)|_\Delta$ and unit map $\varepsilon|_\Delta$.*
(b) *If $\mathfrak{A}$ is an $\mathbb{M}$-algebra, then $\mathfrak{A}|_\Delta$ is an $\mathbb{M}|_\Delta$-algebra.*

Proof. To improve readability, let us denote the functor $(-)|_\Delta$ by R and the functor $(-)^\uparrow$ by E. Then $\mathbb{M}|_\Delta = R \circ \mathbb{M} \circ E$. We denote the inclusion $ER \Rightarrow \mathrm{Id}$ by i and the identity function $\mathrm{Id} \Rightarrow RE$ by e. One can show that $E \dashv R$ is an adjunction with unit e and counit i, but for our purposes it is sufficient to note that we have the following equalities

$$i \circ Ee = \mathrm{id} \quad \text{and} \quad Ri \circ e = \mathrm{id}\,,$$

whose proofs are trivial.

(a) We have to check three axioms.

$$\begin{aligned}
R(\mu \circ \mathbb{M}i) \circ R\varepsilon &= R(\mu \circ \mathbb{M}i) \circ R\varepsilon \circ e \\
&= R(\mu \circ \varepsilon \circ i) \circ e \\
&= Ri \circ e \\
&= \mathrm{id}\,,
\end{aligned}$$

$$\begin{aligned}
R(\mu \circ \mathbb{M}i) \circ \mathbb{M}|_\Delta R\varepsilon &= R(\mu \circ \mathbb{M}i) \circ \mathbb{M}|_\Delta (R\varepsilon \circ e) \\
&= R\big(\mu \circ \mathbb{M}i \circ \mathbb{M}ER\varepsilon \circ \mathbb{M}Ee\big) \\
&= R\big(\mu \circ \mathbb{M}(i \circ ER\varepsilon \circ Ee)\big) \\
&= R\big(\mu \circ \mathbb{M}(\varepsilon \circ i \circ Ee)\big) \\
&= R\mathbb{M}(\mathrm{id} \circ \mathbb{M}\mathrm{id}) \\
&= \mathrm{id}\,,
\end{aligned}$$

$$\begin{aligned}
R(\mu \circ \mathbb{M}i) \circ R(\mu \circ \mathbb{M}i) &= R\big(\mu \circ \mathbb{M}i \circ \mu \circ \mathbb{M}i\big) \\
&= R\big(\mu \circ \mu \circ \mathbb{M}\mathbb{M}i \circ \mathbb{M}i\big) \\
&= R\big(\mu \circ \mathbb{M}\mu \circ \mathbb{M}(\mathbb{M}i \circ i)\big) \\
&= R\big(\mu \circ \mathbb{M}(\mu \circ \mathbb{M}i \circ i\big) \\
&= R\big(\mu \circ \mathbb{M}(i \circ ER(\mu \circ \mathbb{M}i))\big) \\
&= R\big(\mu \circ \mathbb{M}i \circ \mathbb{M}ER(\mu \circ \mathbb{M}i)\big) \\
&= R(\mu \circ \mathbb{M}i) \circ \mathbb{M}|_\Delta R(\mu \circ \mathbb{M}i)\,.
\end{aligned}$$

(b) Note that the product has the correct type since

$$R(\pi \circ \mathbb{M}i) : R\mathbb{M}ERA \to RA \quad \text{and} \quad \mathbb{M}|_\Delta(A|_\Delta) = R\mathbb{M}ERA\,.$$

For the axioms of an $\mathbb{M}|_\Delta$-algebra, we have

$$\begin{aligned}
R(\pi \circ \mathbb{M}i) \circ R\varepsilon &= R(\pi \circ \mathbb{M}i \circ \varepsilon) \\
&= R(\pi \circ \varepsilon \circ i) \\
&= Ri \\
&= \mathrm{id}\,, \\
R(\pi \circ \mathbb{M}i) \circ \mathbb{M}|_\Delta R(\pi \circ \mathbb{M}i) &= R\big(\pi \circ \mathbb{M}(i \circ ER(\pi \circ \mathbb{M}i))\big) \\
&= R\big(\pi \circ \mathbb{M}(\pi \circ \mathbb{M}i \circ i)\big) \\
&= R\big(\pi \circ \mu \circ \mathbb{M}(\mathbb{M}i \circ i)\big) \\
&= R\big(\pi \circ \mathbb{M}i \circ \mu \circ \mathbb{M}i\big) \\
&= R(\pi \circ \mathbb{M}i) \circ R(\mu \circ \mathbb{M}i)\,.
\end{aligned}$$

□

In the remainder of this section, we work in the category Pos^{Ξ_+} where $\Xi_+ := \mathscr{P}(\omega)$. The functors $\mathbb{R}$, $\mathbb{T}$, and $\mathbb{T}^\times$ have canonical extensions to this category, which we will denote by the same letters to keep notation readable.

4.2. The action on the variables. The problem with finding a distributive law for $\mathbb{T}^\times$ is that this monad is not linear. Its multiplication contains an unravelling operation un_0 which is used to duplicate arguments for variables appearing multiple times. To continue we need a variant of this operation that also modifies the variables of the given graph.

Definition 4.8. Let $g \in \mathbb{R}_\zeta A$ be a graph.

(a) For a surjective function $\sigma : \zeta \to \xi$, we denote by ${}^\sigma g \in \mathbb{R}_\xi A$ the graph obtained from g by replacing each variable x by $\sigma(x)$.

(b) We set

$$\mathrm{un}(g) := \langle \sigma, t \rangle\,,$$

where t is the tree obtained from the unravelling $\mathrm{un}_0(g)$ by renaming the variables so that each of them appears exactly once (note that this changes the sort) and σ is the function such that ${}^\sigma t = \mathrm{un}_0(g)$. (To make this well-defined, we can fix a standard well-ordering on the domain, say, the length-lexicographic one, and we number the variables in increasing order with respect to this ordering, i.e., if $v_0 <_{\mathrm{llex}} v_1 <_{\mathrm{llex}} \dots$ is an enumeration of all vertices labelled by a variable, we set $t(v_i) := x_i$, where $x_0, x_1, \dots$ is some fixed sequence of variables.)

(c) We denote by $\mathbb{T}^{\circ}A$ the set of trees $t \in \mathbb{T}^{\times}A$ such that $\mathrm{un}(t) = \langle \mathrm{id}, t \rangle$. Let $\iota : \mathbb{T}^{\circ} \Rightarrow \mathbb{T}^{\times}$ be the inclusion. (In actual calculations we will frequently omit ι to keep the notation simple.) ⌟

Remark 4.9. Note that the operation un can introduce infinitely many different variables. This is the reason why we have to work in $\mathsf{Pos}^{\Xi_{+}}$. ⌟

Example 4.10. $\mathrm{un}(a(x, y, x)) = \langle \sigma, a(x_0, x_1, x_2) \rangle$ where the function σ maps x_0, x_1, x_2 to x, y, x. Then ${}^{\sigma}a(x_0, x_1, x_2) = a(x, y, z)$. ⌟

To make sense of the type of the above operations, we introduce the following monad where every element is annotated by some function renaming the variables.

Definition 4.11. (a) We define a functor $\mathbb{X} : \mathsf{Pos}^{\Xi_{+}} \to \mathsf{Pos}^{\Xi_{+}}$ as follows. For $A \in \mathsf{Pos}^{\Xi_{+}}$, we set

$$\mathbb{X}_{\xi}A := \{\, \langle \sigma, a \rangle \mid a \in A_{\zeta}\,,\ \sigma : \zeta \to \xi \text{ surjective} \,\}\,.$$

We define the order on $\mathbb{X}_{\xi}A$ by

$$\langle \sigma, a \rangle \leq \langle \tau, b \rangle \quad :\text{iff} \quad \sigma = \tau \quad \text{and} \quad a \leq b\,.$$

For a morphism $f : A \to B$, we define $\mathbb{X}f : \mathbb{X}A \to \mathbb{X}B$ by

$$\mathbb{X}f(\langle \sigma, a \rangle) := \langle \sigma, f(a) \rangle\,.$$

(b) We define functions $\mathrm{comp} : \mathbb{X}\mathbb{X}A \to \mathbb{X}A$ and $\mathrm{in} : A \to \mathbb{X}A$ by

$$\mathrm{comp}(\langle \tau, \langle \sigma, a \rangle \rangle) := \langle \tau \circ \sigma, a \rangle \quad \text{and} \quad \mathrm{in}(a) := \langle \mathrm{id}, a \rangle\,.$$ ⌟

Lemma 4.12. *$\langle \mathbb{X}, \mathrm{comp}, \mathrm{in} \rangle$ and $\langle \mathbb{T}^{\circ}, \mathrm{flat}, \mathrm{sing} \rangle$ are monads.*

The set $\mathbb{T}^{\times}A$ carries a canonical structure of a $\mathbb{X}$-algebra.

Definition 4.13. For $\langle \sigma, t \rangle \in \mathbb{X}\mathbb{T}^{\times}A$, we define the *reconstitution operation*

$$\mathrm{re}(\langle \sigma, t \rangle) := {}^{\sigma}t \in \mathbb{T}^{\times}A\,.$$

We denote its restriction to $\mathbb{X}\mathbb{T}^{\circ}$ by $\mathrm{re}_0 := \mathrm{re} \circ \mathbb{X}\iota : \mathbb{X}\mathbb{T}^{\circ} \Rightarrow \mathbb{T}^{\times}$. ⌟

The unravelling operation on trees can now be formalised using the following two natural transformations.

Lemma 4.14. *The inclusion morphism $\iota : \mathbb{T}^{\circ} \Rightarrow \mathbb{T}^{\times}$ is a morphism of monads. The functions*

$$\mathrm{un} : \mathbb{T}^{\times} \Rightarrow \mathbb{X}\mathbb{T}^{\circ}, \quad \mathrm{re}_0 : \mathbb{X}\mathbb{T}^{\circ} \Rightarrow \mathbb{T}^{\times}, \quad \textit{and} \quad \mathrm{re} : \mathbb{X}\mathbb{T}^{\times} \Rightarrow \mathbb{T}^{\times}$$

form natural transformations satisfying the following equations.

(a) $\mathrm{re}_0 \circ \mathrm{un} = \mathrm{id}$
(b) $\mathrm{un} \circ \mathrm{re} = \mathrm{comp} \circ \mathbb{X}\mathrm{un}$
(c) $\mathrm{un} \circ \iota = \mathrm{in}$
(d) $\mathrm{re}_0 \circ \mathrm{comp} = \mathrm{re} \circ \mathbb{X}\mathrm{re}_0$
(e) $\mathrm{flat}^{\times} \circ \mathrm{re}_0 = \mathrm{re} \circ \mathbb{X}(\mathrm{flat}^{\times} \circ \iota)$
(f) $\mathrm{re} \circ \mathrm{in} = \mathrm{id}$
(g) $\mathrm{un} \circ \mathrm{re}_0 = \mathrm{id}$

Proof. The fact that ι is a morphism of monads is straightforward. To see that un is natural, it is sufficient to note that

$$\mathrm{un}(t) = \langle \sigma, s\rangle \quad \text{iff} \quad \mathrm{un}(\mathbb{T}^\times f(t)) = \langle \sigma, \mathbb{T}^\circ f(s)\rangle\,,$$

for every function $f : A \to B$. For re, we have

$$\begin{aligned}\mathbb{T}^\times f(\mathrm{re}(\langle \sigma, t\rangle)) &= \mathbb{T}^\times f({}^\sigma t)\\ &= {}^\sigma\big(\mathbb{T}^\times f(t)\big)\\ &= \mathrm{re}(\langle \sigma, \mathbb{T}^\times f(t)\rangle) = \mathrm{re}(\mathbb{X}\mathbb{T}^\times f(\langle \sigma, t\rangle))\,.\end{aligned}$$

Since $\mathrm{re}_0 = \mathrm{re} \circ \mathbb{X}\iota$, this implies that re_0 is natural as well.

(a) Note that $\mathrm{re}_0 \circ \mathrm{un} = \mathrm{id}$ holds since

$$\mathrm{un}(t) = \langle \sigma, s\rangle \quad \text{implies} \quad {}^\sigma s = t\,, \quad \text{for trees } t \in \mathbb{T}^\times A\,.$$

(b) Suppose that $\mathrm{un}(t) = \langle \sigma, s\rangle$ and $\mathrm{un}({}^\tau t) = \langle \rho, r\rangle$. Then

$${}^{\tau\circ\sigma} s = {}^\tau t = {}^\rho r\,.$$

In particular, s and r only differ in the labelling of the variables. But $s, r \in \mathbb{T}^\circ A$ implies that the variables appear in the same order in both trees. Hence, $s = r$ and it follows that $\tau \circ \sigma = \rho$. Consequently,

$$\begin{aligned}\mathrm{un}(\mathrm{re}(\langle \tau, t\rangle)) &= \langle \rho, r\rangle\\ &= \langle \tau \circ \sigma, s\rangle = \mathrm{comp}(\langle \tau, \langle \sigma, s\rangle\rangle) = \mathrm{comp}(\mathbb{X}\mathrm{un}(\langle \tau, t\rangle))\,.\end{aligned}$$

(c)–(f) We have

$$\begin{aligned}\mathrm{un}(\iota(t)) &= \langle \mathrm{id}, t\rangle = \mathrm{in}(t)\,,\\ \mathrm{re}_0\big(\mathrm{comp}\big(\langle \sigma, \langle \tau, t\rangle\rangle\big)\big) &= \mathrm{re}_0(\langle \sigma \circ \tau, t\rangle)\\ &= {}^{\sigma\circ\tau}\iota(t)\\ &= {}^\sigma({}^\tau\iota(t))\\ &= {}^\sigma\mathrm{re}_0(\langle \tau, t\rangle)\\ &= \mathrm{re}\big(\big\langle \sigma, \mathrm{re}_0(\langle \tau, t\rangle)\big\rangle\big) = \mathrm{re}\big(\mathbb{X}\mathrm{re}_0\big(\langle \sigma, \langle \tau, t\rangle\rangle\big)\big)\,,\\ \mathrm{flat}^\times(\mathrm{re}_0(\langle \sigma, t\rangle)) &= \mathrm{flat}^\times({}^\sigma\iota(t))\\ &= {}^\sigma(\mathrm{flat}^\times \circ \iota)(t)\\ &= \mathrm{re}\big(\big\langle \sigma, (\mathrm{flat}^\times \circ \iota)(t)\big\rangle\big) = (\mathrm{re} \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota))(\langle \sigma, t\rangle)\,,\\ \mathrm{re}(\mathrm{in}(t)) &= \mathrm{re}(\langle \mathrm{id}, t\rangle) = {}^{\mathrm{id}} t = t\,.\end{aligned}$$

(g) By (c), we have

$$\mathrm{un} \circ \mathrm{re}_0 = \mathrm{un} \circ \mathrm{re} \circ \mathbb{X}\iota = \mathrm{comp} \circ \mathbb{X}\mathrm{un} \circ \mathbb{X}\iota = \mathrm{comp} \circ \mathbb{X}\mathrm{in} = \mathrm{id}\,. \qquad \square$$

We can understand point (a) of this lemma as saying that $\mathbb{T}^\times$ is a retract of $\mathbb{X}\mathbb{T}^\circ$, but only as functors, not necessarily as monads. For the latter we first have to establish that $\mathbb{X}\mathbb{T}^\circ$ forms a monad and that the operations un and re_0 are morphisms of monads.

Proposition 4.15.

(a) $\mathbb{X}\mathbb{T}^\circ$ *forms a monad with multiplication*

$$\mathrm{un} \circ \mathrm{re} \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota \circ \mathbb{T}^\circ \mathrm{re}_0) : \mathbb{X}\mathbb{T}^\circ\mathbb{X}\mathbb{T}^\circ \Rightarrow \mathbb{X}\mathbb{T}^\circ$$

and unit

$$\mathrm{in} \circ \mathrm{sing} : \mathrm{Id} \Rightarrow \mathbb{X}\mathbb{T}^\circ \,.$$

(b) $\mathrm{re}_0 : \mathbb{X}\mathbb{T}^\circ \Rightarrow \mathbb{T}^\times$ *and* $\mathrm{un} : \mathbb{T}^\times \Rightarrow \mathbb{X}\mathbb{T}^\circ$ *are isomorphisms of monads.*
(c) $\mathrm{in} : \mathbb{T}^\circ \Rightarrow \mathbb{X}\mathbb{T}^\circ$ *is an injective morphism of monads.*

Proof. (a), (b) By Lemma 4.14 (c), (e), and (a), we have

$$\begin{aligned}
\mathrm{re}_0 \circ \mathrm{in} \circ \mathrm{sing} &= \mathrm{re}_0 \circ \mathrm{un} \circ \iota \circ \mathrm{sing} = \iota \circ \mathrm{sing} = \mathrm{sing}^\times, \\
\mathrm{flat}^\times \circ \mathrm{re}_0 \circ \mathbb{X}\mathbb{T}^\circ \mathrm{re}_0 &= \mathrm{re} \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota \circ \mathbb{T}^\circ \mathrm{re}_0) \\
&= \mathrm{re}_0 \circ \mathrm{un} \circ \mathrm{re} \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota \circ \mathbb{T}^\circ \mathrm{re}_0) \,.
\end{aligned}$$

As re_0 is a surjective natural transformation, most of the claim therefore follows by Lemma 2.5. It only remains to check that un is also a morphism of monads. For this, note that by Lemma 4.14 (c), (a), and (e) we have

$$\begin{aligned}
\mathrm{in} \circ \mathrm{sing} &= \mathrm{un} \circ \iota \circ \mathrm{sing} = \mathrm{un} \circ \mathrm{sing}^\times \,, \\
\mathrm{un} \circ \mathrm{flat}^\times &= \mathrm{un} \circ \mathrm{flat}^\times \circ \mathrm{re}_0 \circ \mathrm{un} \\
&= \mathrm{un} \circ \mathrm{re} \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota) \circ \mathrm{un} \\
&= \mathrm{un} \circ \mathrm{re} \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota) \circ \mathrm{un} \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathrm{un}) \\
&= \mathrm{un} \circ \mathrm{re} \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota \circ \mathbb{T}^\circ \mathrm{re}_0) \circ \mathrm{un} \circ \mathbb{T}^\times \mathrm{un} \,.
\end{aligned}$$

(c) As un and ι are morphisms of monads, so is $\mathrm{un} \circ \iota = \mathrm{in}$. ☐

Corollary 4.16. $\mathbb{T}^\times \cong \mathbb{X}\mathbb{T}^\circ$ *(as monads)*

One could hope to construct a distributive law $\mathbb{T}^\circ\mathbb{X} \Rightarrow \mathbb{X}\mathbb{T}^\circ$ by applying the Theorem of Beck (Theorem 2.12) to the monad structure on $\mathbb{X}\mathbb{T}^\circ$. This does not work for the following reason.

Lemma 4.17. *The natural transformation* $\mathbb{X}\mathrm{sing} : \mathbb{X} \Rightarrow \mathbb{X}\mathbb{T}^\circ$ *is* not *a morphism of monads.*

Proof. The following of the two axioms fails:

$$\mathbb{X}(\mathrm{flat}^\times \circ \iota \circ \mathbb{T}^\circ \mathrm{re}_0) \circ \mathbb{X}\mathrm{sing} \circ \mathbb{X}\mathbb{X}\mathrm{sing} \neq \mathbb{X}\mathrm{sing} \circ \mathrm{comp}.$$

To see this, fix $\langle \sigma, \langle \tau, a \rangle \rangle \in \mathbb{X}\mathbb{X}A$. Then

$$\begin{aligned}
&(\mathbb{X}(\mathrm{flat}^\times \circ \iota \circ \mathbb{T}^\circ \mathrm{re}_0) \circ \mathbb{X}\mathrm{sing} \circ \mathbb{X}\mathbb{X}\mathrm{sing})(\langle \sigma, \langle \tau, a \rangle \rangle) \\
&= \mathbb{X}(\mathrm{flat}^\times \circ \iota \circ \mathbb{T}^\circ \mathrm{re}_0 \circ \mathrm{sing} \circ \mathbb{X}\mathrm{sing})(\langle \sigma, \langle \tau, a \rangle \rangle) \\
&= \mathbb{X}(\mathrm{flat}^\times \circ \iota \circ \mathrm{sing} \circ \mathrm{re}_0 \circ \mathbb{X}\mathrm{sing})(\langle \sigma, \langle \tau, a \rangle \rangle) \\
&= \mathbb{X}(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing})(\langle \sigma, \langle \tau, a \rangle \rangle) \\
&= \langle \sigma, {}^\tau\mathrm{sing}(a) \rangle \,,
\end{aligned}$$

whereas

$$\begin{aligned}&(\mathbb{X}\mathrm{sing} \circ \mathrm{comp})(\langle \sigma, \langle \tau, a\rangle\rangle)\\ &= \mathbb{X}\mathrm{sing}(\langle \sigma \circ \tau, a\rangle)\\ &= \langle \sigma \circ \tau, \mathrm{sing}(a)\rangle\,.\end{aligned}$$

For $\tau \neq \mathrm{id}$, these two values are different. □

4.3. Graphs and unravellings. The next step is to transfer the unravelling operation from $\mathbb{T}^{\times}A$ to arbitrary sets.

Definition 4.18. (a) An *unravelling structure* $\langle A, \mathrm{re}, \mathrm{un}\rangle$ consists of a set $A \in \mathsf{Pos}^{\Xi}_{+}$ equipped with two functions

$$\mathrm{re} : \mathbb{X}A \to A \quad \text{and} \quad \mathrm{un} : A \to \mathbb{X}A$$

such that $\langle A, \mathrm{re}\rangle$ forms a $\mathbb{X}$-algebra while un satisfies

$$\mathbb{X}\mathrm{un} \circ \mathrm{un} = \mathbb{X}\mathrm{in} \circ \mathrm{un} \quad \text{and} \quad \mathrm{re} \circ \mathrm{un} = \mathrm{id}\,.$$

We call $\mathrm{un}(a)$ the *unravelling* of a. To keep notation simple, we write

$${}^{\sigma}a := \mathrm{re}(\langle \sigma, a\rangle)\,.$$

(b) A *morphism of unravelling structures* is a function $\varphi : A \to B$ satisfying

$$\mathrm{un} \circ \varphi = \mathbb{X}\varphi \circ \mathrm{un} \quad \text{and} \quad \varphi \circ \mathrm{re} = \mathrm{re} \circ \mathbb{X}\varphi\,.$$ ⌟

Clearly, the operations re and un defined above for trees $t \in \mathbb{T}^{\times}A$ induce an unravelling structure on $\mathbb{T}^{\times}A$. But note that this is not the case for $\mathbb{R}A$ since we have $\mathrm{re}(\mathrm{un}(g)) \neq g$, for every $g \in \mathbb{R}A$ that is not a tree.

Example 4.19. For each $\mathbb{T}^{\times}$-algebra $\mathfrak{A} = \langle A, \pi\rangle$, we can equip the universe A with the *trivial* unravelling structure where

$$\mathrm{un} := \mathrm{in} \quad \text{and} \quad {}^{\sigma}a := \pi({}^{\sigma}\mathrm{sing}(a))\,.$$ ⌟

Remark 4.20. Note that the monad multiplication $\mathrm{flat}^{\times}$ is *not* a morphism of unravelling structures since $\mathrm{un} \circ \mathrm{flat}^{\times} \neq \mathbb{X}\mathrm{flat}^{\times} \circ \mathrm{un}$. In what follows we will therefore *not* work in the category of unravelling structures and their morphisms. Instead we will work in the weaker category of unravelling structures with arbitrary monotone maps as morphisms. ⌟

As a technical tool, we use the following generalisation of the unravelling relation for graphs where we do not only unravel the graph itself but also each label. The intuition is as follows. Suppose we are given a relation $\theta \subseteq A \times B$ and a graph $h \in \mathbb{R}B$. We construct an (unravelled) graph $g \in \mathbb{R}A$ as follows. Starting at the root v, we pick some element $c\;\theta\;h(v)$, and label $g(v)$ by the unravelling of c. Then we recursively choose labellings for the successors. Note that the shapes of g and h are different since we are unravelling g, so the labels in h might have a higher arity than the corresponding ones in g. Consequently, we simultaneously construct a graph homomorphism $\varphi : g \to h$ to keep track of which vertices of g correspond to which ones of h.

To simplify the definition, we will split the construction into two stages. In the first step we apply the unravelling operation to every label of h, resulting in a graph $\mathbb{R}\mathrm{un}(h) \in \mathbb{R}\mathbb{X}B$. What is then left for the second step is the following relation, which does the choosing of the label and the unravelling of the tree. What makes this operation complicated is the fact that the unravelling depends on the chosen label, while the label may depend on which copy (produced by previous unravelling steps) of a vertex we are at. So we cannot separate the second stage into two independent phases.

Definition 4.21. (a) Let $g \in \mathbb{R}_\xi A$ and $h \in \mathbb{R}_\zeta B$. A *graph homomorphism* is a function $\varphi : \mathrm{dom}(g) \to \mathrm{dom}(h)$ such that

- φ maps the root of g to the root of h;
- $\varphi(u)$ is a successor of $\varphi(v)$ if, and only if, u is a successor of v (not necessarily with the same edge labelling); and
- $\varphi(v)$ is labelled by a variable if, and only if, v is labelled by one.

(b) Suppose that $\varphi : g \to h$ is a surjective graph homomorphism and let $v \in \mathrm{dom}(g)$ be a vertex of sort ξ with successors $(u_x)_{x\in\xi}$ and suppose that $\varphi(v)$ has sort ζ. We denote by $\varphi_{/v} : \xi \to \zeta$ the function such that

$$\varphi(u_x) \text{ is the } \varphi_{/v}(x)\text{-successor of } \varphi(v)\,.$$

(c) Let $s \in \mathbb{R}A$, $t \in \mathbb{R}B$, and $\theta \subseteq \mathbb{X}A \times B$. We write

$$\varphi, \sigma : s \;\theta^{\mathrm{sel}}\; t$$

if the following conditions are satisfied.

- $s \in \mathbb{T}^\circ A$
- $\varphi : s \to t$ is a surjective graph homomorphism.
- $\sigma : \xi \to \zeta$ is surjective.
- $\langle \varphi_{/v}, s(v)\rangle \;\theta\; t(\varphi(v))\,,\quad$ for every $v \in \mathrm{dom}_0(g)\,.$
- $\sigma(s(v)) = t(\varphi(v))\,,\qquad$ if $s(v) = x$ is a variable. ⌟

We are mostly interested in the cases where θ is either the identity $=$ or set membership $\in$. The resulting relations are

$$\varphi, \sigma : s =^{\mathrm{sel}} t\,,\quad \text{for } s \in \mathbb{T}^\times A \text{ and } t \in \mathbb{T}^\times \mathbb{X}A\,,$$

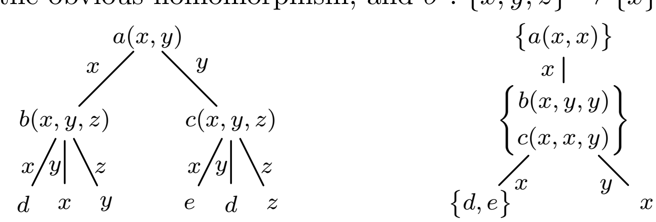
$$\varphi, \sigma : s \in^{\mathrm{sel}} t\,,\quad \text{for } s \in \mathbb{T}^\times A \text{ and } t \in \mathbb{T}^\times \mathbb{U}\mathbb{X}A\,.$$

Combining them with the unravelling operation as explained above, we obtain the relations

$$\varphi, \sigma : s =^{\mathrm{un}} t \qquad :\text{iff} \qquad \varphi, \sigma : s =^{\mathrm{sel}} \mathbb{R}\mathrm{un}(t)\,,$$

$$\varphi, \sigma : s \in^{\mathrm{un}} t\,, \qquad :\text{iff} \qquad \varphi, \sigma : s \in^{\mathrm{sel}} \mathbb{R}\mathbb{U}\mathrm{un}(t)\,.$$

Example 4.22. We have $\varphi, \sigma : g \in^{\mathrm{un}} h$ where g is the tree on the left, h the one on the right, $\varphi : g \to h$ is the obvious homomorphism, and $\sigma : \{x, y, z\} \to \{x\}$.



⌟

Remark 4.23. (a) For every graph g, there exists a canonical graph homomorphism $\varphi : \mathrm{un}_0(g) \to g$.

(b) Note that

$$\varphi, \sigma : g =^{\mathrm{sel}} k \quad \text{and} \quad k\;\theta^{\mathbb{R}}\; h \quad \text{implies} \quad \varphi, \sigma : g\;\theta^{\mathrm{sel}}\; h\,,$$

but the converse is generally not true since the function φ does not need to be injective and we can choose different values $\langle \varphi_{/u}, c_u \rangle$, $\langle \varphi_{/v}, c_v \rangle \mathrel{\theta} h(w)$ for $u, v \in \varphi^{-1}(w)$. For this reason, we cannot reduce the relation $\in^{\mathrm{sel}}$ to the much simpler $=^{\mathrm{sel}}$. ⌟

Let us derive an algebraic description of the relation $\varphi, \sigma : s =^{\mathrm{sel}} t$ that is much easier to work with. We introduce a function un^+ satisfying

$$\langle \sigma, s \rangle = \mathrm{un}^+(t) \quad \text{iff} \quad \varphi, \sigma : s =^{\mathrm{sel}} t\,, \quad \text{for some } \varphi\,,$$

and a similar function dun associated with the relation $=^{\mathrm{un}}$.

Definition 4.24. (a) For a set A, we define the *strong unravelling operation* $\mathrm{un}^+ : \mathbb{T}^\times \mathbb{X} A \to \mathbb{X}\mathbb{T}^\circ A$ by

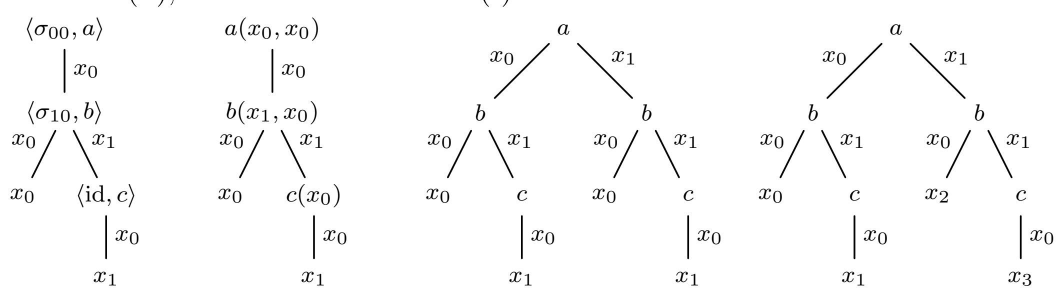
$$\mathrm{un}^+ := \mathrm{un} \circ \mathrm{flat}^\times \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing})\,.$$

(b) For an unravelling structure A, we define the *deep unravelling operation* $\mathrm{dun} : \mathbb{T}^\times A \to \mathbb{X}\mathbb{T}^\circ A$ by

$$\mathrm{dun} := \mathrm{un}^+ \circ \mathbb{T}^\times \mathrm{un}\,.$$

⌟

Example 4.25. To understand the definition of un^+, let us consider the following tree $t \in \mathbb{T}^\times \mathbb{X} A$. Below we have depicted t itself, the intermediate terms $t' := \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing})(t)$ and $t'' := \mathrm{flat}(t')$, and the end result $\mathrm{un}^+(t)$.



Here $a, b \in A_{\{x_0, x_1\}}$, $c \in A_{\{x_0\}}$, and σ_{ij} denotes the function mapping $x_0 \mapsto x_i$ and $x_1 \mapsto x_j$. ⌟

Let us check that the above definitions have the desired effect.

Lemma 4.26. *We have*

$$\langle \sigma, s \rangle = \mathrm{un}^+(t) \quad \text{iff} \quad \varphi, \sigma\colon s =^{\mathrm{sel}} t\,, \quad \textit{for some } \varphi\,,$$
$$\langle \sigma, s \rangle = \mathrm{dun}(t) \quad \text{iff} \quad \varphi, \sigma\colon s =^{\mathrm{un}} t\,, \quad \textit{for some } \varphi\,.$$

Proof. We only have to prove the first equivalence. Then the second one follows by definition of dun and $=^{\mathrm{un}}$. Hence, set

$$r := \mathbb{R}(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing})(t) \quad \text{and} \quad \langle \sigma, s \rangle := \mathrm{un}(\mathrm{flat}^\times(r))\,,$$

let $\varphi : \mathrm{dom}(\mathrm{flat}^\times(r)) \to \mathrm{dom}(t)$ be the homomorphism from above, let $\varphi : \mathrm{dom}(\mathrm{flat}^\times(r)) \to \mathrm{dom}(t)$ be the graph homomorphism induced by the canonical map

$$\mathrm{dom}_0(\mathrm{flat}^\times(r)) \to \sum_{v \in \mathrm{dom}_0(r)} \mathrm{dom}_0(r(v))\,,$$

and suppose that $\varphi', \sigma' : s' =^{\mathrm{sel}} t$. We have to show that

$$\varphi = \varphi'\,, \quad \sigma = \sigma'\,, \quad \text{and} \quad s = s'\,.$$

We start by proving that $\varphi(v) = \varphi'(v)$ and $s(v) = s'(v)$, by induction on v. For the root $v = \langle\rangle$ of $\mathrm{flat}^\times(r)$, we have $\varphi(\langle\rangle) = \langle\rangle = \varphi'(\langle\rangle)$.

For the inductive step, suppose that we have already shown that $\varphi(v) = \varphi'(v)$. We will prove that $s(v) = s'(v)$ and that $\varphi(u) = \varphi'(u)$, for every successor u of v. By definition of $=^{\mathrm{sel}}$, we have

$$t(\varphi'(v)) = \langle \varphi'_{/v}, s'(v)\rangle\,, \quad \text{for } v \in \mathrm{dom}(s')\,.$$

This implies that

$$r(\varphi'(v)) = (\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing})\big(\langle \varphi'_{/v}, s'(v)\rangle\big) = {}^{\varphi'_{/v}}\mathrm{sing}(s'(v))\,.$$

Consequently,

$$s(v) = \mathrm{flat}^\times(r)(v) = r(\varphi(v))(\langle\rangle) = r(\varphi'(v))(\langle\rangle) = s'(v)\,.$$

To complete the induction, it remains to show that $\varphi_{/v} = \varphi'_{/v}$. Let $(u_x)_x$ be the successors of v in s and let $(w_y)_y$ be the successors of $\varphi(v)$ in r. Then

$$r(\varphi(v)) = {}^{\varphi'_{/v}}\mathrm{sing}(s(v))$$

implies that the x-successor of v in s corresponds (via φ) to the $\varphi'_{/v}(x)$-successor of $\varphi(v)$ in r. Thus

$$\varphi(u_x) = w_{\varphi'_{/v}(x)}\,.$$

But, by definition of $\varphi_{/v}$, we also have $\varphi(u_x) = w_{\varphi_{/v}(x)}$. Hence,

$$\varphi_{/v}(x) = \varphi'_{/v}(x)\,.$$

This completes the induction. To finish the proof it remains to show that $\sigma = \sigma'$ and that $s(v) = s'(v)$, for all $v \in \mathrm{dom}(s) \setminus \mathrm{dom}_0(s)$. For the latter, note that the vertices of s carrying a variable are the same as those of s' carrying one. Since the variable labelling is determined by the ordering of these vertices with respect to the length-lexicographic order, it follows that the two labellings coincide.

Hence, let v be such a vertex. Then

$$\sigma(s(v)) = \mathrm{flat}^\times(r)(v) = r(\varphi(v)) = t(\varphi(v)) = \sigma'(s'(v)) = \sigma'(s(v))\,.$$

Thus, $\sigma(x) = \sigma'(x)$, for all x, which implies that $\sigma = \sigma'$ □

Let us collect a few basic properties of the operations we have just introduced.

Lemma 4.27.

(a) $\mathbb{X}(\mathrm{un} \circ \mathrm{flat}^\times \circ \iota) \circ \mathrm{dun} = \mathbb{X}(\mathrm{in} \circ \mathrm{flat}^\times \circ \iota) \circ \mathrm{dun}$
(b) $\mathrm{flat}^\times \circ \mathrm{re} \circ \mathrm{dun} = \mathrm{flat}^\times$
(c) $\mathrm{un} \circ \mathrm{flat}^\times = \mathbb{X}(\mathrm{flat}^\times \circ \iota) \circ \mathrm{dun}$
(d) $\mathrm{un}^+ \circ \mathbb{T}^\times\mathrm{in} = \mathrm{un}$
(e) $\mathrm{un}^+ \circ \mathrm{sing}^\times = \mathbb{C}\mathrm{sing}$

Proof. (a) Let $\langle \sigma, s\rangle = \mathrm{dun}(t)$. By Lemma 4.26, it follows that $\varphi, \sigma : s =^{\mathrm{un}} t$. Consequently, we have

$$\mathrm{un}(t(\varphi(v))) = \langle \varphi_{/v}, s(v)\rangle\,, \quad \text{for all } v \in \mathrm{dom}_0(s)\,.$$

In particular, $s(v) \in \mathbb{T}^\circ A$ and, therefore, $s \in \mathbb{T}^\circ\mathbb{T}^\circ A$. This implies that $\mathrm{flat}(s) \in \mathbb{T}^\circ A$. Hence, $\mathrm{un}(\mathrm{flat}(s)) = \langle \mathrm{id}, \mathrm{flat}(s)\rangle$ and we have

$$\begin{aligned}
\mathbb{X}(\mathrm{un} \circ \mathrm{flat})(\mathrm{dun}(t)) &= \langle \sigma, \mathrm{un}(\mathrm{flat}(s))\rangle \\
&= \langle \sigma, \langle \mathrm{id}, \mathrm{flat}(s)\rangle\rangle \\
&= \langle \sigma, \mathrm{in}(\mathrm{flat}(s))\rangle = \mathbb{X}(\mathrm{in} \circ \mathrm{flat})(\mathrm{dun}(t))\,.
\end{aligned}$$

(b) From Lemma 4.14 it follows that

$$\begin{aligned}
\mathrm{flat}^\times \circ \mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un} &= \mathrm{re}_0 \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota) \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un} \\
&= \mathrm{re}_0 \circ \mathbb{X}(\mathrm{flat}^\times \circ \mathrm{sing}^\times) \circ \mathrm{un} \\
&= \mathrm{re}_0 \circ \mathrm{un} \\
&= \mathrm{id}\,.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\mathrm{flat}^\times \circ \mathrm{re} \circ \mathrm{dun} &= \mathrm{flat}^\times \circ \mathrm{re} \circ \mathrm{un} \circ \mathrm{flat}^\times \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un}) \\
&= \mathrm{flat}^\times \circ \mathrm{flat}^\times \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un}) \\
&= \mathrm{flat}^\times \circ \mathbb{T}^\times\mathrm{flat}^\times \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un}) \\
&= \mathrm{flat}^\times \circ \mathbb{T}^\times\mathrm{id} \\
&= \mathrm{flat}^\times.
\end{aligned}$$

(c) By (a) and Lemma 4.14, we have

$$\begin{aligned}
\mathbb{X}(\mathrm{flat}^\times \circ \iota) \circ \mathrm{dun} &= \mathrm{comp} \circ \mathbb{X}(\mathrm{in} \circ \mathrm{flat}^\times \circ \iota) \circ \mathrm{dun} \\
&= \mathrm{comp} \circ \mathbb{X}(\mathrm{in} \circ \mathrm{flat}^\times \circ \iota) \circ \mathrm{dun} \\
&= \mathrm{comp} \circ \mathbb{X}\mathrm{un} \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota) \circ \mathrm{dun} \\
&= \mathrm{un} \circ \mathrm{re} \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota) \circ \mathrm{dun} \\
&= \mathrm{un} \circ \mathrm{flat}^\times \circ \mathrm{re}_0 \circ \mathrm{dun} \\
&= \mathrm{un} \circ \mathrm{flat}^\times.
\end{aligned}$$

(d) We have

$$\begin{aligned}
\mathrm{un}^+ \circ \mathbb{T}^\times\mathrm{in} &= \mathrm{un} \circ \mathrm{flat}^\times \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing}) \circ \mathbb{T}^\times\mathrm{in} \\
&= \mathrm{un} \circ \mathrm{flat}^\times \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathrm{in} \circ \mathrm{sing}) \\
&= \mathrm{un} \circ \mathrm{flat}^\times \circ \mathbb{T}^\times\mathrm{sing} \\
&= \mathrm{un}\,.
\end{aligned}$$

(e)

$$\begin{aligned}
\mathrm{un}^+ \circ \mathrm{sing}^\times &= \mathrm{un} \circ \mathrm{flat}^\times \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing}) \circ \mathrm{sing}^\times \\
&= \mathrm{un} \circ \mathrm{flat}^\times \circ \mathrm{sing}^\times \circ \mathrm{re}_0 \circ \mathbb{C}\mathrm{sing} \\
&= \mathrm{un} \circ \mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} \\
&= \mathbb{X}\mathrm{sing}\,. \qquad \square
\end{aligned}$$

In Lemma 4.26, we have found an algebraic characterisation of the relations $=^{\mathrm{sel}}$ and $=^{\mathrm{un}}$ in terms of the operations un^+ and dun. Unfortunately, there does not seem to be a purely algebraic definition of a similar operation characterising the relation $\in^{\mathrm{sel}}$. Instead, we have to define it directly in terms of $\in^{\mathrm{sel}}$.

Definition 4.28. We define the *selection operation* $\mathrm{sel} : \mathbb{T}^\times\mathbb{U}\mathbb{X} \Rightarrow \mathbb{U}\mathbb{X}\mathbb{T}^\circ$ by

$$\mathrm{sel}(t) := \{\, \langle \sigma, s\rangle \mid \varphi, \sigma : s \in^{\mathrm{sel}} t \,\}\,. \qquad \lrcorner$$

The properties of this operation are as follows.

Lemma 4.29.

(a) $\mathrm{sel} : \mathbb{T}^\times\mathbb{U}\mathbb{X} \Rightarrow \mathbb{U}\mathbb{X}\mathbb{T}^\circ$ *is a natural transformation on* Pos^{Ξ_+}*.*
(b) $\mathrm{sel} \circ \mathbb{T}^\times\mathrm{pt} = \mathrm{pt} \circ \mathrm{un}^+$
(c) $\mathrm{sel} \circ \mathrm{sing}^\times = \mathbb{U}\mathbb{X}\mathrm{sing}$
(d) $\mathrm{sel} \circ \mathbb{T}^\times(\mathrm{pt} \circ \mathrm{in}) = \mathrm{pt} \circ \mathrm{un}$
(e) $\mathbb{U}(\mathrm{dun} \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un} = \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un}$

Proof. (a) Let $f : A \to B$. Then

$$\begin{aligned}
&\varphi, \sigma : s \in^{\mathrm{sel}} \mathbb{T}^\times\mathbb{U}\mathbb{X}f(t) \\
\text{iff}\quad &\langle \varphi_{/v}, s(v)\rangle \in \mathbb{U}\mathbb{X}f(t(\varphi(v)))\,, && \text{for all } v\,, \\
\text{iff}\quad &s(v) \geq f(r(v)) \quad\text{and}\quad \langle \varphi_{/v}, r(v)\rangle \in t(\varphi(v))\,, && \text{for all } v\,, \\
\text{iff}\quad &s \geq \mathbb{T}^\times f(r) \quad\text{and}\quad \varphi, \sigma : r \in^{\mathrm{sel}} t\,,
\end{aligned}$$

implies that $\mathrm{sel}(\mathbb{T}^\times\mathbb{U}\mathbb{X}f(t)) = \mathbb{U}\mathbb{X}\mathbb{T}^\times f(\mathrm{sel}(t))$.

(b) To simplify notation, we will again leave the universal quantification over vertices v implicit. Let $t \in \mathbb{T}^\times\mathbb{X}A$. Then

$$\begin{aligned}
\mathrm{sel}(\mathbb{T}^\times\mathrm{pt}(t)) &= \Uparrow\{\, \langle \sigma, s\rangle \mid \varphi, \sigma : s \in^{\mathrm{sel}} \mathbb{T}^\times\mathrm{pt}(t) \,\} \\
&= \Uparrow\Big\{\, \langle \sigma, s\rangle \;\Big|\; \langle \varphi_{/v}, s(v)\rangle \in \mathrm{pt}(t(\varphi(v))) \text{ or} \\
&\qquad\qquad [s(v) = x \text{ and } \mathbb{T}^\times\mathrm{pt}(t)(\varphi(v)) = \sigma(x)] \,\Big\} \\
&= \Uparrow\Big\{\, \langle \sigma, s\rangle \;\Big|\; \langle \varphi_{/v}, s(v)\rangle \geq t(\varphi(v)) \text{ or} \\
&\qquad\qquad [s(v) = x \text{ and } t(\varphi(v)) = \sigma(x)] \,\Big\} \\
&= \Uparrow\Big\{\, \langle \sigma, s\rangle \;\Big|\; \langle \varphi_{/v}, s(v)\rangle = t(\varphi(v)) \text{ or} \\
&\qquad\qquad [s(v) = x \text{ and } t(\varphi(v)) = \sigma(x)] \,\Big\} \\
&= \Uparrow\{\, \langle \sigma, s\rangle \mid \varphi, \sigma : s =^{\mathrm{sel}} t \,\} \\
&= \Uparrow\{\mathrm{un}^+(t)\} \\
&= \mathrm{pt}(\mathrm{un}^+(t))\,.
\end{aligned}$$

(c) Let $I \in \mathbb{U}\mathbb{X}A$. Then

$$\begin{aligned}
\mathrm{sel}(\mathrm{sing}^\times(I)) &= \Uparrow\{\, \langle \sigma, s\rangle \mid \varphi, \sigma : s \in^{\mathrm{sel}} \mathrm{sing}^\times(I) \,\} \\
&= \Uparrow\{\, \langle \sigma, s\rangle \mid s = \mathrm{sing}(a)\,,\ \langle \tau, a\rangle \in I\,,\ \sigma = \tau \,\} \\
&= \Uparrow\{\, \langle \sigma, \mathrm{sing}(a)\rangle \mid \langle \sigma, a\rangle \in I \,\} \\
&= \mathbb{U}\mathbb{X}\mathrm{sing}(I)\,.
\end{aligned}$$

(d) By (b) and Lemma 4.27 (d), we have

$$\mathrm{sel} \circ \mathbb{T}^\times(\mathrm{pt} \circ \mathrm{in}) = \pi \circ \mathrm{un}^+ \circ \mathbb{T}^\times\mathrm{in} = \pi \circ \mathrm{un}\,.$$

(e) Let $\langle \sigma, s\rangle \in \mathrm{sel}(\mathbb{T}^\times\mathbb{U}\mathrm{un}(t))$. Then $\varphi, \sigma : s \in^{\mathrm{sel}} \mathbb{T}^\times\mathbb{U}\mathrm{un}(t)$, which implies that

$$\langle \varphi_{/v}, s(v)\rangle \in \mathrm{un}(t(\varphi(v)))\,.$$

Consequently, we have $\mathrm{un}(s(v)) = \langle \mathrm{id}, s(v)\rangle$, that is, $\mathbb{T}^\circ\mathrm{un}(s) = \mathbb{T}^\circ\mathrm{in}(s)$. Hence,

$$\begin{aligned}(\mathbb{T}^\circ\mathrm{un} \circ \mathrm{re}_0)(\langle \sigma, s\rangle) &= \mathbb{T}^\circ\mathrm{un}({}^\sigma s)\\ &= {}^\sigma\mathbb{T}^\circ\mathrm{un}(s)\\ &= {}^\sigma\mathbb{T}^\circ\mathrm{in}(s)\\ &= \mathbb{T}^\circ\mathrm{in}({}^\sigma s) = (\mathbb{T}^\circ\mathrm{in} \circ \mathrm{re}_0)(\langle \sigma, s\rangle)\,.\end{aligned}$$

Furthermore, $s \in \mathbb{T}^\circ A$ implies that $\mathrm{un}(s) = \langle \mathrm{id}, s\rangle$. It therefore follows by Lemma 4.27 (d) that

$$\begin{aligned}(\mathrm{dun} \circ \mathrm{re}_0)(\langle \sigma, s\rangle) &= (\mathrm{un}^+ \circ \mathbb{T}^\times\mathrm{un} \circ \mathrm{re}_0)(\langle \sigma, s\rangle)\\ &= (\mathrm{un}^+ \circ \mathbb{T}^\times\mathrm{in} \circ \mathrm{re}_0)(\langle \sigma, s\rangle)\\ &= (\mathrm{un} \circ \mathrm{re}_0)(\langle \sigma, s\rangle)\\ &= (\mathrm{comp} \circ \mathbb{X}\mathrm{un})(\langle \sigma, s\rangle)\\ &= (\mathrm{comp} \circ \mathbb{X}\mathrm{in})(\langle \sigma, s\rangle)\\ &= \langle \sigma, s\rangle\,.\end{aligned}$$

Consequently,

$$\mathbb{U}(\mathrm{dun} \circ \mathrm{re}_0) \restriction (\mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un}) = \mathbb{U}\mathrm{id} \restriction (\mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un})\,. \qquad \square$$

We need one more equation concerning the operation sel whose proof is more involved: Lemma 4.31 below contains a commutation relation between sel and $\mathrm{flat}^\times$ that is similar to one of the axioms of a distributive law. The proof makes use of the following technical lemma.

Lemma 4.30. *Let $r \in \mathbb{T}^\times\mathbb{T}^\times A$ and $t \in \mathbb{T}^\times\mathbb{T}^\times B$ be trees, set $s := \mathrm{flat}(r)$, let*

$$\begin{aligned}\chi &: \mathrm{dom}(s) \to \mathrm{dom}(\mathrm{flat}^\times(t))\,,\\ \varphi &: \mathrm{dom}(r) \to \mathrm{dom}(t)\,,\\ \psi_v &: \mathrm{dom}(r(v)) \to \mathrm{dom}(t(\varphi(v)))\end{aligned}$$

be surjective graph homomorphisms, and let

$$\begin{aligned}\lambda &: \mathrm{dom}(\mathrm{flat}^\times(t)) \to \sum_{v \in \mathrm{dom}_0(t)} \mathrm{dom}_0(t(v)) + [\mathrm{dom}(t) \setminus \mathrm{dom}_0(t)]\,,\\ \mu &: \mathrm{dom}(\mathrm{flat}(r)) \to \sum_{v \in \mathrm{dom}_0(r)} \mathrm{dom}_0(r(v)) + [\mathrm{dom}(r) \setminus \mathrm{dom}_0(r)]\end{aligned}$$

be the functions induced by the canonical maps

$$\begin{aligned}\mathrm{dom}_0(\mathrm{flat}^\times(t)) &\to \sum_{v \in \mathrm{dom}_0(t)} \mathrm{dom}_0(t(v))\\ \mathrm{dom}_0(\mathrm{flat}(r)) &\to \sum_{v \in \mathrm{dom}_0(r)} \mathrm{dom}_0(r(v))\,.\end{aligned}$$

Then

$$\lambda(\chi(w)) = \langle \varphi(v), \psi_v(u)\rangle\,, \quad \textit{for every } w \in \mathrm{dom}_0(s) \textit{ with } \mu(w) = \langle v, u\rangle\,,$$

implies that

$$\chi_{/w} = (\psi_v)_{/u}\,, \quad \textit{for } \mu(w) = \langle v, u\rangle\,.$$

Proof. Consider a vertex $w \in \mathrm{dom}_0(s)$ with $\mu(w) = \langle v, u\rangle$ and an x-successor $\tilde{u}$ of u. Suppose that $\lambda(\chi(w)) = \langle v', u'\rangle$. First, let us consider the case where $\tilde{u} \in \mathrm{dom}_0(r(v))$. Let $\tilde{w}$ be the successor of w with $\mu(\tilde{w}) = \langle v, \tilde{u}\rangle$. By assumption, we have $\lambda(\chi(\tilde{w})) = \langle \varphi(v), \psi_v(\tilde{u})\rangle$ and $\psi_v(\tilde{u})$ is the y-successor of $\psi_v(u)$ in $t(\varphi(v))$, for some y. By definition, it follows that $\chi_{/w}(x) = y$ and $(\psi_v)_{/u}(x) = y$.

It remains to consider the case where $\tilde{u} \notin \mathrm{dom}_0(r(v))$. Then $r(v)(\tilde{u}) = z$, for some variable z. Let v' be the z-successor of v, let $\langle\rangle$ be the root of $r(v')$, and let $\tilde{w}$ be the successor of w with $\mu(\tilde{w}) = \langle v', \langle\rangle\rangle$. Then $\lambda(\chi(w)) = \langle \varphi(v'), \psi_{v'}(\langle\rangle)\rangle$. Let y be the variable such that $\lambda(\varphi(v'), \psi_{v'}(\langle\rangle))$ is the y-successor of $\lambda(\varphi(v), \psi_v(u))$. Then $\chi_{/w}(x) = y$ and $(\psi_v)_{/u}(x) = y$. □

Lemma 4.31. $\mathrm{sel} \circ \mathrm{flat}^\times = \mathbb{UX}\mathrm{flat} \circ \mathrm{sel} \circ \mathbb{T}^\times \mathrm{sel}$

Proof. Note that the canonical function

$$\mathrm{dom}_0(\mathrm{flat}^\times(t)) \to \sum_{v \in \mathrm{dom}_0(t)} \mathrm{dom}_0(t(v))$$

induces a function

$$\lambda : \mathrm{dom}(\mathrm{flat}^\times(t)) \to \sum_{v \in \mathrm{dom}_0(t)} \mathrm{dom}_0(t(v)) + [\mathrm{dom}(t) \setminus \mathrm{dom}_0(t)]\,.$$

Similarly, for a tree r (which we will specify below), we obtain a function

$$\mu : \mathrm{dom}(\mathrm{flat}(r)) \to \sum_{v \in \mathrm{dom}_0(r)} \mathrm{dom}_0(r(v)) + [\mathrm{dom}(r) \setminus \mathrm{dom}_0(r)]\,.$$

To prove the lemma, we check the two inclusions separately.

($\supseteq$) Suppose that $\langle \sigma, s\rangle \in \mathbb{UX}\mathrm{flat}(\mathrm{sel}(\mathbb{T}^\times \mathrm{sel}(t)))$. Then

$$s = \mathrm{flat}(r) \quad \text{for some} \quad \varphi, \sigma : r \in^{\mathrm{sel}} \mathbb{T}^\times \mathrm{sel}(t)\,.$$

For every vertex v of r, it follows that

$$\langle \varphi_{/v}, r(v)\rangle \in \mathrm{sel}(t(\varphi(v))) \quad \text{or} \quad r(v) = x \text{ and } \mathrm{sel}(t(\varphi(v))) = \sigma(x)\,.$$

This implies that

$$\psi_v, \varphi_{/v} : r(v) \in^{\mathrm{sel}} t(\varphi(v)) \quad \text{or} \quad r(v) = x \text{ and } t(\varphi(v)) = \sigma(x)\,,$$

for some homomorphism ψ_v. Let χ be the unique graph homomorphism satisfying the equations

$$\lambda(\chi(w)) = \begin{cases} \langle \varphi(v), \psi_v(u)\rangle & \text{if } \mu(w) = \langle v, u\rangle\,, \\ \varphi(v) & \text{if } \mu(w) = v\,, \end{cases}$$

where λ and μ are the homomorphisms defined above. We claim that $\chi, \sigma : s \in^{\mathrm{sel}} \mathrm{flat}^\times(t)$, which implies that $\langle \sigma, s\rangle \in \mathrm{sel}(\mathrm{flat}^\times(t))$.

Hence, fix a vertex $w \in \mathrm{dom}(s) = \mathrm{dom}(\mathrm{flat}(r))$. First, consider the case where $w \in \mathrm{dom}_0(s)$. Suppose that $\mu(w) = \langle v, u\rangle$. Then $\psi_v, \varphi_{/v} : r(v) \in^{\mathrm{sel}} t(\varphi(v))$ implies that

$$\langle (\psi_v)_{/u}, r(v)(u)\rangle \in t(\varphi(v))(\psi_v(u))\,.$$

Consequently, we have

$$\langle (\psi_v)_{/u}, s(w) \rangle \in t(\varphi(v))(\psi_v(u)) = \mathrm{flat}^\times(t)(\chi(w))\,.$$

Furthermore, we have $(\psi_v)_{/u} = \chi_{/w}$ by Lemma 4.30.

It remains to consider the case where $s(w) = x$ is a variable. Then $\mu(w) = v$, for some $v \in \mathrm{dom}(r)$, and $r(v) = x$ implies that $t(\varphi(v)) = \sigma(x)$. Hence,

$$\mathrm{flat}^\times(t)(\chi(w)) = t(\lambda(\chi(w))) = t(\varphi(v)) = \sigma(x)\,.$$

($\subseteq$) Suppose that $\langle \sigma, s \rangle \in \mathrm{sel}(\mathrm{flat}^\times(t))$. Then

$$\chi, \sigma : s \in^{\mathrm{sel}} \mathrm{flat}^\times(t)\,, \quad \text{for some } \chi\,.$$

We define a tree r with $\mathrm{flat}(r) = s$ as follows. Intuitively, we factorise s by cutting every edge $w \to w'$ such that the corresponding vertices $\chi(w)$ and $\chi(w')$ in $\mathrm{flat}^\times(t)$ belong to different components $t(v)$ and $t(v')$, i.e., if $\lambda(\chi(w)) = \langle v, u \rangle$ and $\lambda(\chi(w')) = \langle v', u' \rangle$ with $v \neq v'$. The formal definition is as follows. Let us call a vertex $w \in \mathrm{dom}(s)$ *principal* if its image under χ corresponds to the root of some conponent $t(v)$, or to a leaf, that is, if

$$\lambda(\chi(w)) = \langle v, \langle\rangle \rangle \quad \text{or} \quad \lambda(\chi(w)) = v\,, \quad \text{for some } v\,,$$

(where $\langle\rangle$ denotes the root of $t(v)$). We define the domain of r by

$$\mathrm{dom}(r) := \{\, w \in \mathrm{dom}(s) \mid w \text{ is principal} \,\}$$

and the edge relation as follows. Given a principal vertex w, let $w_0, \dots, w_{n-1}$ be an enumeration of all minimal principal vertices w' with $w \prec w'$. We make w_i an i-successor of w. (The precise labels i are not important, only the fact that they are pairwise distinct.) Finally, the labelling of r is given by

$$r(w) := \begin{cases} r_w & \text{if } w \in \mathrm{dom}_0(s)\,, \\ s(w) & \text{if } w \notin \mathrm{dom}_0(s)\,, \end{cases}$$

where r_w is the tree with

$$\mathrm{dom}(r_w) := \big\{\, u \in \mathrm{dom}(s) \bigm| w \preceq u \text{ and there is no principal } w' \text{ with } w \prec w' \prec u \,\big\}\,,$$

$$r_w(u) := \begin{cases} s(u) & \text{if } u \notin \mathrm{dom}(r) \text{ or } u = w\,, \\ i & \text{if } u = w_i \in \mathrm{dom}(r) \text{ is the } i\text{-successor of } w \text{ in } r\,. \end{cases}$$

By definition, it follows that $\mathrm{flat}(r) = s$ and that

$$\mu(w) = \langle v, w \rangle\,, \quad \text{if } w \in \mathrm{dom}_0(s)\,, \text{ where } v \text{ is the maximal principal vertex with } v \preceq w\,,$$
$$\text{and} \quad \mu(w) = w\,, \qquad \text{if } w \notin \mathrm{dom}_0(s)\,.$$

Let φ and ψ_v be the functions defined by the equations

$$\begin{aligned} \langle \varphi(v), \psi_v(u) \rangle &= \lambda(\chi(w))\,, && \text{for } \mu(w) = \langle v, u \rangle\,, \\ \varphi(w) &= \lambda(\chi(w))\,, && \text{if } w \in \mathrm{dom}(s) \setminus \mathrm{dom}_0(s)\,, \\ \psi_v(u) &= u'' && \text{if } u \in \mathrm{dom}(r(v)) \setminus \mathrm{dom}_0(r(v))\,, \end{aligned}$$

where the vertex u'' in the last equation is chosen as follows. Given u, let u' be the predecessor of u and let x be the label of the edge $u' \to u$. Then u'' is the $(\psi_v)_{/u'}(x)$-successor of $\psi_v(u')$.

We claim that, for all v,

$$\psi_v, \varphi_{/v} : r(v) \in^{\mathrm{sel}} t(\varphi(v)) \quad \text{or} \quad r(v) = x \text{ and } t(\varphi(v)) = \sigma(x)\,.$$

Then it follows that

$$\langle \varphi_{/v}, r(v) \rangle \in \mathrm{sel}(t(\varphi(v))) \quad \text{or} \quad r(v) = x \text{ and } \mathrm{sel}(t(\varphi(v))) = \sigma(x)\,.$$

Thus,

$$\langle \sigma, r \rangle \in \mathrm{sel}(\mathbb{T}^\times \mathrm{sel}(t)) \quad \text{and} \quad \langle \sigma, s \rangle \in \mathbb{U}\mathbb{X}\mathrm{flat}(\mathrm{sel}(\mathbb{T}^\times \mathrm{sel}(t)))\,,$$

as desired. Hence, it remains to prove the above claim.

If $r(v) = x$ is a variable, we have $s(v) = r(v) = x$ and, therefore,

$$t(\varphi(v)) = t(\lambda(\chi(v))) = \mathrm{flat}^\times(t)(\chi(v)) = \sigma(x)\,,$$

as desired. Otherwise, $v \in \mathrm{dom}_0(r)$ and we have to show that

$$\psi_v, \varphi_{/v} : r(v) \in^{\mathrm{sel}} t(\varphi(v))\,.$$

Note that $\chi, \sigma : s \in^{\mathrm{sel}} \mathrm{flat}^\times(t)$ implies that

$$\langle \chi_{/w}, s(w) \rangle \in \mathrm{flat}^\times(t)(\chi(w))\,, \quad \text{for all } w\,.$$

We distinguish two cases. If $u \in \mathrm{dom}_0(r(v))$, let $w \in \mathrm{dom}_0(s)$ be the vertex with $\mu(w) = \langle v, u \rangle$. Then

$$\langle \chi_{/w}, r(v)(u) \rangle = \langle \chi_{/w}, s(w) \rangle \in \mathrm{flat}^\times(t)(\chi(w)) = t(\varphi(v))(\psi_v(u))\,.$$

By Lemma 4.30, we have $\chi_{/w} = (\psi_v)_{/u}$, which implies that

$$\langle (\psi_v)_{/u}, r(v)(u) \rangle \in t(\varphi(v))(\psi_v(u))\,.$$

If $u \in \mathrm{dom}(r(v)) \setminus \mathrm{dom}_0(r(v))$ with label $r(v)(u) = x$, let v' be the x-successor of v. By definition of $\varphi_{/v}$, it follows that $\varphi(v')$ is the $\varphi_{/v}(x)$-successor of $\varphi(v)$ in t. This implies that $t(v)(\psi_v(u)) = \varphi_{/v}(x)$. □

4.4. A partial distributive law. The idea to find our partial distributive law is to work in the category of unravelling structures, although this does not solve our problems entirely. First of all, there is no obvious way to lift the functor $\mathbb{U}$ to unravelling structures. Given an unravelling structure A, we can define an 'unravelling map' $\mathbb{U}\mathrm{un} : \mathbb{U}A \to \mathbb{U}\mathbb{X}A$, but we would need one of the form $\mathbb{U}A \to \mathbb{X}\mathbb{U}A$, and there is no natural transformation $\mathbb{U}\mathbb{X} \Rightarrow \mathbb{X}\mathbb{U}$. The functor $\mathbb{T}^\times$ on the other hand *can* be lifted to the category of unravelling structures, but only in a trivial way: given A we can forget its unravelling structure, construct $\mathbb{T}^\times A$, and equip it with the canonical unravelling structure defined above (which does not depend on that of A). In particular, with this definition the monad multiplication $\mathrm{flat}^\times$ would not be a morphism of the resulting unravelling structure. What would be more useful would be a lift that uses deep unravelling dun as the unravelling operation on $\mathbb{T}^\times A$. But there is no corresponding reconstitution operation re satisfying $\mathrm{re} \circ \mathrm{dun} = \mathrm{id}$.

What we will do instead is to use an ad-hoc argument showing how to define a lift of $\mathbb{U}$ to sufficiently well-behaved $\mathbb{T}^\times$-algebras. We are mainly interested in free $\mathbb{T}^\times$-algebras, but a slightly more abstract definition helps to make the proof more modular. We extract the needed properties of the algebras in question in the following technical definition.

Definition 4.32. We say that a $\mathbb{T}^\times$-algebra $\mathfrak{A} = \langle A, \pi \rangle$ *supports unravelling* if its universe A can be equipped with an unravelling structure that satisfies the following conditions.

$$\begin{aligned} \pi \circ \mathrm{re} \circ \mathbb{X}\mathrm{sing}^\times &= \mathrm{re}\,, \\ \mathrm{un} \circ \mathrm{re} &= \mathrm{comp} \circ \mathbb{X}\mathrm{un}\,, \\ \mathbb{X}(\mathrm{un} \circ \pi \circ \iota) \circ \mathrm{dun} &= \mathbb{X}(\mathrm{in} \circ \pi \circ \iota) \circ \mathrm{dun}\,. \end{aligned}$$ ⌟

The intended target for this definition are the free algebras. We start by noting that these satisfy the above conditions.

Proposition 4.33. *The free $\mathbb{T}^\times$-algebra $\langle \mathbb{T}^\times A, \mathrm{flat}^\times \rangle$ supports unravelling.*

Proof. Using the operations un and re from Definitions 4.8 and 4.13, it follows by Lemma 4.14 (e) and (b), that

$$\begin{aligned} \mathrm{flat}^\times \circ \mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} &= \mathrm{re}_0 \circ \mathbb{X}(\mathrm{flat}^\times \circ \iota) \circ \mathbb{X}\mathrm{sing} \\ &= \mathrm{re}_0 \circ \mathbb{X}(\mathrm{flat}^\times \circ \mathrm{sing}^\times) = \mathrm{re}_0\,, \\ \mathrm{un} \circ \mathrm{re}_0 &= \mathrm{comp} \circ \mathbb{X}\mathrm{un}\,, \end{aligned}$$

while the third condition follows by Lemma 4.27 (a). □

For the proof below, let us collect a few basic properties of algebras that support unravelling.

Lemma 4.34. *Let $\mathfrak{A}$ be a $\mathbb{T}^\times$-algebra that supports unravelling.*

(a) $\pi \circ \mathrm{re} = \mathrm{re} \circ \mathbb{X}\pi$
(b) $\pi \circ \mathrm{re} \circ \mathrm{dun} = \pi$
(c) $\mathrm{un} \circ \pi \circ \mathrm{re} \circ \mathrm{dun} = \mathbb{X}\pi \circ \mathrm{dun}$
(d) $\mathrm{un} \circ \pi = \mathbb{X}\pi \circ \mathrm{dun}$
(e) $\mathbb{U}(\mathrm{un} \circ \pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un} = \mathbb{U}\mathbb{X}\pi \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un}$

Proof. Below we will make freely use of the equations from Lemma 4.14.

(a) We have

$$\begin{aligned} \pi \circ \mathrm{re} &= \pi \circ \mathrm{re} \circ \mathbb{X}(\mathrm{flat}^\times \circ \mathrm{sing}^\times) \\ &= \pi \circ \mathrm{flat}^\times \circ \mathrm{re} \circ \mathbb{X}\mathrm{sing}^\times \\ &= \pi \circ \mathbb{T}^\times\pi \circ \mathrm{re} \circ \mathbb{X}\mathrm{sing}^\times \\ &= \pi \circ \mathrm{re} \circ \mathbb{X}\mathbb{T}^\circ\pi \circ \mathbb{X}\mathrm{sing}^\times \\ &= \pi \circ \mathrm{re} \circ \mathbb{X}\mathrm{sing}^\times \circ \mathbb{X}\pi \\ &= \mathrm{re} \circ \mathbb{X}\pi\,, \end{aligned}$$

where the last step follows from the fact that $\mathfrak{A}$ supports unravelling.

(b) Since

$$\pi \circ \mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un} = \mathrm{re}_0 \circ \mathbb{X}\pi \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un} = \mathrm{re}_0 \circ \mathrm{un} = \mathrm{id}\,,$$

we have

$$\begin{aligned}\pi \circ \mathrm{re} \circ \mathrm{dun} &= \pi \circ \mathrm{re} \circ \mathrm{un} \circ \mathrm{flat}^\times \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un}) \\ &= \pi \circ \mathrm{flat}^\times \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un}) \\ &= \pi \circ \mathbb{T}^\times\pi \circ \mathbb{T}^\times(\mathrm{re}_0 \circ \mathbb{X}\mathrm{sing} \circ \mathrm{un}) \\ &= \pi \circ \mathbb{T}^\times\mathrm{id} \\ &= \pi\,.\end{aligned}$$

(c) By (a) and the fact that $\mathfrak{A}$ supports unravelling, we have

$$\begin{aligned}\mathrm{un} \circ \pi \circ \mathrm{re} \circ \mathrm{dun} &= \mathrm{un} \circ \mathrm{re} \circ \mathbb{X}\pi \circ \mathrm{dun} \\ &= \mathrm{comp} \circ \mathbb{X}\mathrm{un} \circ \mathbb{X}\pi \circ \mathrm{dun} \\ &= \mathrm{comp} \circ \mathbb{X}(\mathrm{in} \circ \pi) \circ \mathrm{dun} = \mathbb{X}\pi \circ \mathrm{dun}\,.\end{aligned}$$

(d) By (c) and (b), we have

$$\mathbb{X}\pi \circ \mathrm{dun} = \mathrm{un} \circ \pi \circ \mathrm{re} \circ \mathrm{dun} = \mathrm{un} \circ \pi\,.$$

(e) By (a), Lemma 4.29 (e), and the fact that $\mathfrak{A}$ supports unravelling, we have

$$\begin{aligned}&\mathbb{U}(\mathrm{un} \circ \pi \circ \mathrm{re}) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un} \\ =\ &\mathbb{U}(\mathrm{un} \circ \mathrm{re} \circ \mathbb{X}\pi) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un} \\ =\ &\mathbb{U}(\mathrm{comp} \circ \mathbb{X}\mathrm{un} \circ \mathbb{X}\pi) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un} \\ =\ &\mathbb{U}(\mathrm{comp} \circ \mathbb{X}(\mathrm{un} \circ \pi) \circ \mathrm{dun} \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un} \\ =\ &\mathbb{U}(\mathrm{comp} \circ \mathbb{X}(\mathrm{in} \circ \pi) \circ \mathrm{dun} \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un} \\ =\ &\mathbb{U}(\mathrm{comp} \circ \mathbb{X}(\mathrm{in} \circ \pi)) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un} \\ =\ &\mathbb{U}\mathbb{X}\pi \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un}\,.\end{aligned}$$

□

Finally we can state our partial distributive law for $\mathbb{U}$ and $\mathbb{T}^\times$ for algebras that support unravelling.

Proposition 4.35. *If $\mathfrak{A} = \langle A, \pi\rangle$ is a $\mathbb{T}^\times$-algebra supporting unravelling, we can form a $\mathbb{T}^\times$-algebra $\mathbb{U}\mathfrak{A} := \langle \mathbb{U}A, \hat\pi\rangle$ with product*

$$\hat\pi := \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un}\,.$$

Furthermore, the function $\mathrm{pt} : A \to \mathbb{U}A$ induces an embedding $\mathfrak{A} \to \mathbb{U}\mathfrak{A}$.

Proof. We have to check three equations. To see that pt is an embedding, note that

$$\begin{aligned}\hat\pi \circ \mathbb{T}^\times\mathrm{pt} &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un} \circ \mathbb{T}^\times\mathrm{pt} \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathrm{pt} \circ \mathbb{T}^\times\mathrm{un} \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{pt} \circ \mathrm{un}^+ \circ \mathbb{T}^\times\mathrm{un} \\ &= \mathrm{pt} \circ \pi \circ \mathrm{re}_0 \circ \mathrm{un}^+ \circ \mathbb{T}^\times\mathrm{un} \\ &= \mathrm{pt} \circ \pi \circ \mathrm{re}_0 \circ \mathrm{dun} \\ &= \mathrm{pt} \circ \pi \\ &= \mathbb{U}\pi \circ \mathrm{pt}\,.\end{aligned}$$

where the third step follows by Lemma 4.29 (b) and the sixth one by Lemma 4.34 (b). For the unit law, we have

$$\begin{aligned}\hat{\pi} \circ \mathrm{sing}^{\times} &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un} \circ \mathrm{sing}^{\times} \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathrm{sing}^{\times} \circ \mathbb{U}\mathrm{un} \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathbb{U}\mathbb{X}\mathrm{sing} \circ \mathbb{U}\mathrm{un} \\ &= \mathbb{U}(\pi \circ \mathrm{sing} \circ \mathrm{re}_0 \circ \mathrm{un}) \\ &= \mathbb{U}(\mathrm{id} \circ \mathrm{id}) \\ &= \mathrm{id}\,,\end{aligned}$$

where the third step follows by Lemma 4.29 (c). Finally, for the associative law,

$$\begin{aligned}\hat{\pi} \circ \mathbb{T}^{\times}\hat{\pi} &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un} \circ \mathbb{T}^{\times}(\mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un}) \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^{\times}(\mathbb{U}(\mathrm{un} \circ \pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un}) \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^{\times}(\mathbb{U}\mathbb{X}\pi \circ \mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un}) \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathbb{U}\mathbb{X}\mathbb{T}^{\circ}\pi \circ \mathrm{sel} \circ \mathbb{T}^{\times}(\mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un}) \\ &= \mathbb{U}(\pi \circ \mathbb{T}^{\circ}\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^{\times}(\mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un}) \\ &= \mathbb{U}(\pi \circ \mathrm{flat}^{\times} \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^{\times}(\mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un}) \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0 \circ \mathbb{X}\mathrm{flat}^{\times}) \circ \mathrm{sel} \circ \mathbb{T}^{\times}(\mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un}) \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathrm{flat}^{\times} \circ \mathbb{T}^{\times}\mathbb{T}^{\times}\mathbb{U}\mathrm{un} \\ &= \mathbb{U}(\pi \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^{\times}\mathbb{U}\mathrm{un} \circ \mathrm{flat}^{\times} \\ &= \hat{\pi} \circ \mathrm{flat}^{\times}\,.\end{aligned}$$

where the third step follows by Lemma 4.34 (e) and the eighth one by Lemma 4.31. □

For technical reasons, we have worked so far in the category $\mathsf{Pos}^{\Xi+}$. But the category we are actually interested in is Pos^{Ξ}. The following consequence can be considered the main result of this section.

Theorem 4.36. *In Pos^{Ξ}, the set $\mathbb{U}\mathbb{T}^{\times}A$ forms a $\mathbb{T}^{\times}$-algebra with product*

$$\hat{\pi}(t) := \Uparrow\{\, \mathrm{flat}^{\times}({}^{\sigma}s) \mid \varphi, \sigma : s \in^{\mathrm{un}} t \,\}\,.$$

Proof. We know by Proposition 4.35 that $\mathbb{U}\mathbb{T}^{\times}A^{\uparrow}$ forms a $\mathbb{T}^{\times}$-algebra in $\mathsf{Pos}^{\Xi+}$. Since $\mathbb{U}\mathbb{T}^{\times}A = (\mathbb{U}\mathbb{T}^{\times}A^{\uparrow})|_{\Xi}$, the claim follows by Lemma 4.7. □

In order to strengthen this theorem to obtain a $\mathbb{U}\mathbb{T}^{\times}$-algebra, we would need to prove that $\mathbb{U}\mathbb{T}^{\times}$ forms a monad. The next result shows that the canonical choice for the corresponding monad multiplication does not work. (Note that this is not a simple consequence of Theorem 3.23 since it might be the case that, instead of condition (M1) of Theorem 2.12 (4), it is (M2) or (M3) that is violated.)

Proposition 4.37. *The function $\kappa : \mathbb{U}\mathbb{T}^{\times}\mathbb{U}\mathbb{T}^{\times}A \to \mathbb{U}\mathbb{T}^{\times}A$ with*

$$\kappa(T) := \Uparrow\{\, \mathrm{flat}^{\times}({}^{\sigma}s) \mid \varphi, \sigma : s \in^{\mathrm{un}} t\,,\ t \in T \,\}$$

does *not* satisfy the associative law*

$$\kappa \circ \kappa = \kappa \circ \mathbb{U}\mathbb{T}^{\times}\kappa\,.$$

Proof. We use term notation $a(c), b(c,d), \dots$ for trees. Note that, for two sets

$$X = \{\, a_i(x_0, x_0) \mid i < m \,\} \quad \text{and} \quad Y = \{\, \mathrm{sing}^\times(c_i) \mid i < n \,\}$$

(where $a_i \in A_2$ and $c_i \in A_0$) we have

$$\begin{aligned}
\kappa(\{X(Y)\}) &= \{\, \mathrm{flat}^\times({}^\sigma s) \mid \varphi, \sigma : s \in^{\mathrm{un}} X(Y) \,\} \\
&= \big\{\, \mathrm{flat}^\times({}^\sigma s) \bigm| s = u(v,w)\,,\ u = \mathrm{sing}^\times(a_i)\,, \\
&\qquad\qquad v = \mathrm{sing}^\times(c_k)\,,\ w = \mathrm{sing}^\times(c_l)\,,\ i < m\,,\ k, l < n \,\big\} \\
&= \{\, a_i(c_k, c_l) \mid i < m\,,\ k, l < n \,\}\,.
\end{aligned}$$

Similarly, if the $a_i \in A_1$ are unary, we obtain

$$\kappa(\{X(Y)\}) = \{\, a_i(c_k) \mid i < m\,,\ k < n \,\}\,.$$

Setting

$$\begin{aligned}
I &:= \{a(x_0, x_0)\}\,, & C &:= \{c\}\,, \\
J &:= \{b(x_0, x_0)\}\,, & D &:= \{d\}\,, \\
K &:= \{\mathrm{sing}^\times(I), \mathrm{sing}^\times(J)\}\,, & E &:= \{\mathrm{sing}^\times(C), \mathrm{sing}^\times(D)\}\,,
\end{aligned}$$

we obtain

$$\begin{aligned}
\kappa(\{K(E)\}) &= \{I(C), I(D), J(C), J(D)\}\,, \\
\kappa(\{I(C)\}) &= \{a(c,c)\}\,, \qquad \kappa(\{I(D)\}) = \{a(d,d)\}\,, \\
\kappa(\{J(C)\}) &= \{b(c,c)\}\,, \qquad \kappa(\{J(D)\}) = \{b(d,d)\}\,, \\
(\kappa \circ \kappa)(\{K(E)\}) &= \{a(c,c), a(d,d), b(c,c), b(d,d)\}\,, \\
\kappa(K) &= I \cup J \ \ =: X\,, \\
\kappa(E) &= C \cup D =: Y\,, \\
\mathbb{UT}^\times \kappa(\{K(E)\}) &= \{X(Y)\}\,, \\
(\kappa \circ \mathbb{UT}^\times \kappa)(\{K(E)\}) &= \{\, u(v,w) \mid u \in \{a,b\}\,,\ v, w \in \{c,d\} \,\}\,.
\end{aligned}$$

Hence,

$$(\kappa \circ \kappa)(\{K(D)\}) \neq (\kappa \circ \mathbb{UT}^\times \kappa)(\{K(D)\})\,.$$

(For instance, the tree $a(c,d)$ does belong to the right-hand side, but not to the left-hand one.) □

5. Substitutions

As a first application of the tools we have developed above, let us take a look at substitutions for tree languages. We present a simplified account of a recent result by Camino et al. [CDD+22] about finding solutions to inequalities of the form $\sigma[L] \subseteq R$ for regular tree languages L and R. This simplification stems mainly from the terminology and notation introduced above. It does not rely on the results we have proved, except for Lemma 5.3, which depends on Theorem 4.36. In the next section we will give a second, more involved application that makes use of Theorem 4.36 in a more substantial way.

Definition 5.1. Let Σ be an alphabet.

(a) A *substitution* is a function $\sigma : X \to \mathbb{U}\mathbb{T}^\times\Sigma$. We call σ *regular* if every $\sigma(x) \subseteq \mathbb{T}^\times\Sigma$ is a regular tree language.

(b) A substitution $\sigma : X \to \mathbb{U}\mathbb{T}^\times\Sigma$ induces a function $\mathbb{T}^\times X \to \mathbb{U}\mathbb{T}^\times\Sigma$ in two different ways. The *inside-out* morphism σ_{io} is defined by

$$\sigma_{\mathrm{io}}(t) := \{\, \mathrm{flat}^\times(s) \mid s \in^{\mathbb{R}} \mathbb{R}\sigma(t) \,\},$$

while the *outside-in* morphism σ_{oi} is defined by

$$\sigma_{\mathrm{oi}}(t) := \{\, \mathrm{flat}^\times({}^{\sigma}s) \mid \varphi, \sigma : s \in^{\mathrm{un}} \mathbb{R}\sigma(t) \,\}.$$

⌟

Remark 5.2. (a) The reader should compare the simple definition above with the much more involved one given in [CDD+22]. As it turns out such simplifications are not uncommon when using the monadic framework.

(b) Intuitively, the difference between these two variants is that, with the inside-out version σ_{io}, we have to choose the same image $s(u) \in \sigma(t(v))$ for every vertex u of s corresponding to $v \in \mathrm{dom}(t)$, while the outside-in σ_{oi} version allows us to choose a different tree for each of them. The former has the advantage of simplicity, but the latter turns out to be more natural from an algebraic perspective: we will show below that it forms a morphism of $\mathbb{T}^\times$-algebras.

(c) In the notation of Section 4, we can rewrite the above definitions as

$$\sigma_{\mathrm{io}} = \mathbb{U}\mathrm{flat}^\times \circ \mathrm{dist} \circ \mathbb{T}^\times\sigma\,,$$

$$\sigma_{\mathrm{oi}} = \mathbb{U}(\mathrm{flat}^\times \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^\times(\mathbb{U}\mathrm{un} \circ \sigma)\,.$$

Hence, σ_{io} is based on the failed distributive law dist, while σ_{oi} is based on the more successful attempt using the relation $\in^{\mathrm{un}}$. ⌟

For the next lemma, let us recall from Theorem 4.36 that $\mathbb{U}\mathbb{T}^\times\Sigma$ indeed forms a $\mathbb{T}^\times$-algebra.

Lemma 5.3. *$\sigma_{\mathrm{oi}} : \mathbb{T}^\times X \to \mathbb{U}\mathbb{T}^\times\Sigma$ is a morphism of $\mathbb{T}^\times$-algebras.*

Proof. According to Theorem 4.36, the product of the algebra $\mathbb{U}\mathbb{T}^\times\Sigma$ is given by

$$\hat{\pi} := \mathbb{U}(\mathrm{flat}^\times \circ \mathrm{re}_0) \circ \mathrm{sel} \circ \mathbb{T}^\times\mathbb{U}\mathrm{un}\,.$$

Hence, $\sigma_{\mathrm{oi}} = \hat{\pi} \circ \mathbb{T}^\times\sigma$ and it follows that

$$\begin{aligned}\sigma_{\mathrm{oi}} \circ \mathrm{flat}^\times &= \hat{\pi} \circ \mathbb{T}^\times\sigma \circ \mathrm{flat}^\times \\ &= \hat{\pi} \circ \mathrm{flat}^\times \circ \mathbb{T}^\times\mathbb{T}^\times\sigma \\ &= \hat{\pi} \circ \mathbb{T}^\times\hat{\pi} \circ \mathbb{T}^\times\mathbb{T}^\times\sigma = \hat{\pi} \circ \mathbb{T}^\times\sigma_{\mathrm{oi}}\,.\end{aligned}$$

□

Remark 5.4. Note that the function $\sigma_{\mathrm{io}} : \mathbb{T}^\times X \to \mathbb{U}\mathbb{T}^\times\Sigma$ is *not* a morphism of $\mathbb{T}^\times$-algebras. ⌟

For the simpler inside-out substitutions, we can solve inequalities $\rho_{\mathrm{io}}[L] \subseteq R$ as follows.

Theorem 5.5 [CDD+22]. *Let $L \subseteq \mathbb{T}^\times X$ and $R \subseteq \mathbb{T}^\times\Sigma$ be regular tree languages, $\sigma, \tau : X \to \mathbb{U}\mathbb{T}^\times\Sigma$ regular substitutions, and let S be the set of all substitutions ρ such that*

$$\sigma \subseteq \rho \subseteq \tau \quad \text{and} \quad \rho_{\mathrm{io}}[L] \subseteq R\,.$$

Then

(a) *S has finitely many maximal elements.*

(b) *Every maximal element of S is regular.*
(c) *We can effectively compute the maximal elements of S.*

Proof. Since R is regular, it is recognised by some morphism $\eta : \mathbb{T}^\times \Sigma \to \mathfrak{A}$ into a finitary $\mathbb{T}^\times$-algebra $\mathfrak{A} = \langle A, \pi \rangle$ (for a proof see [Blu20, Blu21]). We define the *saturation* $\hat{\rho} : X \to \mathbb{UT}^\times \Sigma$ of a given substitution $\rho : X \to \mathbb{UT}^\times \Sigma$ by

$$\hat{\rho}(x) := \big\{\, s \in \mathbb{T}^\times \Sigma \;\big|\; \eta(s) \in \mathbb{U}\eta(\rho(x)) \,\big\}\,.$$

Then we have $\mathbb{U}\eta \circ \hat{\rho} = \mathbb{U}\eta \circ \rho$. Note that we can rewrite the definition of ρ_{io} as

$$\rho_{\mathrm{io}} = \mathbb{U}\mathrm{flat}^\times \circ \mathrm{dist} \circ \mathbb{T}^\times \rho\,.$$

It follows that

$$\begin{aligned}\mathbb{U}\eta \circ \rho_{\mathrm{io}} &= \mathbb{U}(\eta \circ \mathrm{flat}^\times) \circ \mathrm{dist} \circ \mathbb{T}^\times \rho \\ &= \mathbb{U}(\pi \circ \mathbb{T}^\times \eta) \circ \mathrm{dist} \circ \mathbb{T}^\times \rho = \mathbb{U}\pi \circ \mathrm{dist} \circ \mathbb{T}^\times \mathbb{U}\eta \circ \mathbb{T}^\times \rho\,.\end{aligned}$$

Consequently, we have

$$\begin{aligned}\mathbb{U}\eta \circ \rho_{\mathrm{io}} &= \mathbb{U}\pi \circ \mathrm{dist} \circ \mathbb{T}^\times (\mathbb{U}\eta \circ \rho) \\ &= \mathbb{U}\pi \circ \mathrm{dist} \circ \mathbb{T}^\times (\mathbb{U}\eta \circ \hat{\rho}) = \mathbb{U}\eta \circ \hat{\rho}_{\mathrm{io}}\,.\end{aligned}$$

As $\eta(s) = \eta(t)$ implies $s \in R \Leftrightarrow t \in R$, it therefore follows that

$$\rho_{\mathrm{io}}(t) \subseteq R \quad \text{implies} \quad \hat{\rho}_{\mathrm{io}}(t) \subseteq R\,.$$

Since $\rho \subseteq \hat{\rho}$ this implies that the maximal elements of S satisfy $\rho = \hat{\rho} \cap \tau$. In particular, a substitution of this form is regular. This proves (b).

For (a), note that the number of substitutions of the form $\hat{\rho}$ is bounded by the number of functions $X \to \mathbb{U}A$. As X is finite and A is sort-wise finite, there are only finitely many such functions.

It remains to establish (c). We can enumerate all functions $X \to \mathbb{U}A$. This gives an enumeration of all substitutions of the form $\hat{\rho}$. For each of them, we can check whether $\sigma \leq \hat{\rho} \cap \tau$. If so, $\hat{\rho} \cap \tau$ is a maximal element of S. Otherwise, it is not. □

The more complicated case of outside-in substitutions is still open.

Remark 5.6. There is one technical detail worth mentioning: the way we have defined substitutions, every tree in $\sigma(x)$, for $x \in X_\xi$, contains *all* variables in ξ. But usually one uses a more general notion of a substitution where the trees in $\sigma(x)$ can omit some or all of these variables. We can formalise this generalisation in our setting as follows.

We consider a substitution as a function $\sigma : X \to \mathbb{UT}^< \Sigma$, where $\mathbb{T}^<$ is the functor with

$$\mathbb{T}^<_\xi X := \sum_{\zeta \subseteq \xi} \mathbb{T}^\times_\zeta X\,.$$

We can extend the monad operation to $\mathbb{T}^<$ in the obvious way. As above we define two induced operations $\sigma_{\mathrm{io}}, \sigma_{\mathrm{oi}} : \mathbb{T}^< X \to \mathbb{UT}^< \Sigma$. The definition of the outside-in version is the same as above

$$\sigma_{\mathrm{oi}}(t) := \{\, \mathrm{flat}^<({}^\sigma s) \mid \varphi, \sigma : s \in^{\mathrm{un}} \mathbb{R}^< \sigma(t) \,\}$$

(where $\mathbb{R}^<$ is the corresponding variant of $\mathbb{R}$).

But the inside-out version is more complicated. The problem is that some sets $\sigma(x)$ might be empty, but a tree t might still have a non-empty image $\sigma_{\mathrm{io}}(t)$ because, for every vertex v with $\sigma(t(v)) = \emptyset$, there might be some vertex u higher up in the tree where we have chosen an element $s \in t(u)$ which omits the variable corresponding to the subtree containing v. The easiest way to formalise this process is to make the problem disappear by adding dummy elements to all sets $\sigma(x)$. Hence, fix some element $\perp \notin \Sigma$ and let $\mu : \mathbb{U}\Sigma \to \mathbb{U}(\Sigma + \{\perp\})$ be the function with

$$\mu(I) := I \cup \{\perp\}\,.$$

Then we set

$$\sigma_{\mathrm{io}}(t) := \bigl\{\, \mathrm{flat}^{<}(s) \bigm| s \in^{\mathbb{R}^<} \mathbb{R}^{<}(\mu \circ \sigma)(t)\,,\ \mathrm{flat}^{<}(s) \in \mathbb{T}^{<}\Sigma \,\bigr\}\,.$$

The proof of Theorem 5.5 can now straightforwardly be adapted to these new definitions. ⌟

6. Regular expressions for infinite trees

As a second, more involved application of our results let us define regular expressions for languages of infinite trees. Such expressions seem to be folklore, but we have not found them anywhere in the literature (except for a few remarks in [Tho90]).

We consider tree languages of the form $L \subseteq \mathbb{T}^{\times}_{\xi}\Sigma$, for some alphabet Σ and some fixed sort $\xi \in \Xi$. Alphabets will always be assumed to be finite and unordered. Note that, if Σ is unordered, so is $\mathbb{T}^{\times}_{\xi}\Sigma$ and $\mathbb{U}\mathbb{T}^{\times}_{\xi}\Sigma$ is just the power set. Hence, we can regard every language $L \subseteq \mathbb{T}^{\times}_{\xi}\Sigma$ as an element of $\mathbb{U}\mathbb{T}^{\times}_{\xi}\Sigma$.

We aim for a characterisation of which elements of this set are regular languages. Towards this goal we introduce a few operations on $\mathbb{U}\mathbb{T}^{\times}\Sigma$. They are based on the well-known version for finite trees (see, e.g,. Section 2.4 of [LT21]), suitably modified to work in the sorted setting and to generate infinite trees.

Before presenting the definition we need to deal with the problem that $\mathbb{U} \circ \mathbb{T}^{\times}$ does not form a monad and that $\mathbb{U}\mathbb{T}^{\times}\Sigma$ not a $\mathbb{U}\mathbb{T}^{\times}$-algebra. For this reason we will work with what we call *bialgebras:* a set A equipped both with a $\mathbb{T}^{\times}$-algebra product $\pi : \mathbb{T}^{\times}A \to A$ and a $\mathbb{U}$-algebra product $\rho : \mathbb{U}A \to A$ (without any compatibility condition between them). (Note that this is not the usual use of the word 'bialgebra'.) By Theorem 4.36, $\mathbb{U}\mathbb{T}^{\times}\Sigma$ forms a bialgebra with respect to the monads $\mathbb{T}^{\times}$ and $\mathbb{U}$.

We use the following operations for our version of regular expressions:

- variables $x \in X$,
- letters of the alphabet $a \in \Sigma$,
- substitution $\cdot_x$, iteration $-^{+x}$, and ω-power $-^{\omega x}$ with respect to a single variable x,
- relabelling ${}^{\sigma}-$ of the variables,
- union $+$ and the empty language $\emptyset$.

The formal definition is as follows.

Definition 6.1. Given a bialgebra $\mathfrak{A} = \langle A, \pi, \rho\rangle$ we define the following operations.

(a) Each $a \in A_\xi$, induces an operation $a : A^{\xi} \to A$ by

$$a(\bar b) := \pi(s)\,,$$

where $s \in \mathbb{T}^{\times}_{\xi}A$ is the tree obtained form $\mathrm{sing}(a)$ by replacing each leaf with label $x \in \xi$ by the tree $\mathrm{sing}(b_x)$.

(b) For sorts $\xi, \zeta \in \Xi$ and a variable $x \in \xi$, we define a binary *substitution operation*

$$\cdot_x\ : A_\xi \times A_\zeta \to A_{(\xi\setminus\{x\})\cup\zeta} \quad \text{by} \quad a \cdot_x b := \pi(s)\,,$$

where s is the tree obtained from $\mathrm{sing}(a)$ by replacing the leaf labelled x by the tree $\mathrm{sing}(b)$.

(c) For $a \in A_\xi$ and a surjective map $\sigma : \xi \to \zeta$, we set

$$^{\sigma}a := \pi(s)\,,$$

where s is the tree obtained from $\mathrm{sing}(a)$ by replacing each label $x \in \xi$ by $\sigma(x)$.

(d) We define $+ : A_\xi \times A_\xi \to A_\xi$ and $\emptyset \in A_\xi$ by

$$a + b := \rho(\Uparrow\{a,b\}) \quad \text{and} \quad \emptyset := \rho(\emptyset)\,.$$

(e) Let $\zeta \in \Xi$. We call a tree s *ζ-trivial* if, for all $v \in \mathrm{dom}(s)$ and $z \in \zeta$, we have

$$s(v) = z \quad \text{iff} \quad v \text{ is an } z\text{-successor.}$$

(I.e., all z-successors are labelled by z and there are no other occurrences of z.) For a finite sequence of elements $a_i \in A_{\xi_i}$, $i < n$, and a variable $x \in \zeta := \xi_0 \cup \cdots \cup \xi_{n-1}$, we define the *$\omega$-power* and the *iteration* by

$$(a_0 + \cdots + a_{n-1})^{\omega x} := \rho\big(\big\{\ \pi(s) \bigm| s \in \mathbb{T}^\times_{\zeta\setminus\{x\}}\{a_0,\dots,a_{n-1}\} \text{ is } (\zeta \setminus \{x\})\text{-trivial }\big\}\big)\,,$$

$$(a_0 + \cdots + a_{n-1})^{+x} := \rho\big(\big\{\ \pi(s) \bigm| s \in \mathbb{T}^\times_{\zeta}\{a_0,\dots,a_{n-1}\} \text{ has finite height and it is } (\zeta \setminus \{x\})\text{-trivial }\big\}\big)\,.$$

(f) For a sort $\xi \in \Xi$ and a set Σ, the set $\mathbb{E}_\xi\Sigma$ of *regular expression* over Σ consists of all finite terms R that can be built up from variables and the operations (a)–(e) (for the bialgebra $\mathbb{UT}^\times\Sigma$), where

- we restrict the operations from (a) to those where $a = \Uparrow\mathrm{sing}(c)$, for some $c \in \Sigma$, and
- the free variables are exactly those in ξ.

We write $[\![R]\!] \subseteq \mathbb{UT}^\times\Sigma$ for the value of $R \in \mathbb{E}\Sigma$ in $\mathbb{UT}^\times\Sigma$. ⌟

Remark 6.2. The iteration and the ω-power in (e) have a built-in sum operation in order to support choices between terms of different sorts, which is not possible using the normal sum operation from (d). ⌟

Example 6.3. We consider the alphabet $\Sigma = \{a, b, c\}$ where a and b have sort $\{x, y\}$ and c has sort $\emptyset$.

(a) A regular expression for the language $\mathbb{T}^\times\Sigma$ is

$$E := \big(\big(a(x,y) + b(x,y) + c\big)^{\omega x}\big)^{\omega y}.$$

(b) An expression for the language of all trees with an infinite branch labelled by a is given by

$$R := \big(a(x,z) + a(z,y)\big)^{\omega z} \cdot_x E \cdot_y E\,.$$

(c) Finally, the following expression describes all trees containing the letter a.

$$S := a(x,y) \cdot_x E \cdot_y E + \big(b(x,z) + b(z,y)\big)^{+z} \cdot_z a(x,y) \cdot_x E \cdot_y E\,.$$ ⌟

We still have to show that regular expressions capture the class of regular languages. For the proof, let us quickly recall the notion of a tree automaton (see, e.g., [Tho97, GTW02, Löd21] for details). A *parity automaton* $\mathcal{A} = \langle Q, \Sigma, \zeta, \Delta, q_\mathrm{I}, \Omega\rangle$ consists of a finite set Q

of states, an input alphabet Σ, an input sort $\zeta \in \Xi$, an initial state $q_{\mathrm{I}} \in Q$, a priority function Ω, and a transition relation

$$\Delta \subseteq \sum_{\xi \in \Xi} (Q \times \Sigma_\xi \times Q^{|\xi|}) + (Q \times \zeta)\,.$$

A *run* ρ of such an automaton on an input tree $t \in \mathbb{T}^\times_\zeta \Sigma$ is a labelling of t by states such that

- the root is labelled by q_{I},
- $\big\langle \rho(v), t(v), \rho(u_0), \dots, \rho(u_{n-1}) \big\rangle \in \Delta$, for every vertex v with successors $u_0, \dots, u_{n-1}$,
- every infinite branch $v_0, v_1, \dots$ of t satisfies the parity condition:

$$\liminf_{n\to\infty} \Omega(\rho(v_n)) \quad \text{is even.}$$

A *partial run* is defined exactly like a run, except that the state at the root can be arbitrary and that we do not require the transition relation to hold at vertices v labelled by a variable. Let ρ be a partial run on the tree $t \in \mathbb{T}^\times_\zeta \Sigma$. The *profile* of ρ is the pair $\langle p, (U_z)_{z\in\zeta} \rangle$ where p is the state at the root and, for each variable $z \in \zeta$, U_z is the set of all pairs $\langle k, q \rangle$ such that there is a vertex v labelled z with state q and such that k is the least priority seen along the path from the root to v. We define an ordering on profiles by

$$\langle p, \bar U \rangle \leq \langle p', \bar U' \rangle \qquad :\text{iff} \qquad p = p' \text{ and } U_z \subseteq U'_z \text{ for all } z \in \zeta\,.$$

If $\sigma \leq \tau$, we say that the profile σ is *bounded* by τ.

Theorem 6.4. *Let Σ be an alphabet. A language $L \subseteq \mathbb{T}^\times_\zeta \Sigma$ is regular if, and only if, $L = [\![R]\!]$, for some regular expression $R \in \mathbb{E}_\zeta \Sigma$.*

Proof. ($\Leftarrow$) The class of all regular tree languages is closed under all operations that can appear in a regular expression.

($\Rightarrow$) Let $\mathcal{A} = \langle Q, \Sigma, \zeta, \Delta, q_{\mathrm{I}}, \Omega \rangle$ be an automaton recognising L and fix an enumeration $q_0, \dots, q_{n-1}$ of Q such that $\Omega(q_0) \geq \dots \geq \Omega(q_{n-1})$. For every profile τ of $\mathcal{A}$ and every number $k \leq n$, we will construct a regular expressions R^k_τ defining the language

$$[\![R^k_\tau]\!] = \big\{\, t \in \mathbb{T}^\times \Sigma \;\big|\; \text{there is a partial run on } t \text{ whose profile is bounded by } \tau \text{ and whose internal states are among } q_0, \dots, q_{k-1} \,\big\}\,.$$

Then we obtain the desired expression for L by setting

$$R := \sum_{\tau \in H} R^n_\tau\,,$$

where H is the set of all profiles $\tau = \langle q_{\mathrm{I}}, \bar U \rangle$ such that, for all $z \in \zeta$,

$$\langle k, p \rangle \in U_z \quad \text{implies} \quad \langle p, z \rangle \in \Delta\,.$$

We define the expressions R^k_τ by induction on k. For $k = 0$, we only need to consider runs without internal states. Hence, we can set

$$R^0_\tau := \sum \big\{\, a(\bar x) \mid a \in \Sigma\,,\ \text{there is a partial run on } \mathrm{sing}(a) \text{ whose profile is bounded by } \tau \,\big\}\,.$$

For the inductive step, suppose that $\tau = \langle p, \bar U \rangle$, let ξ be the sort of τ, and let $D := \operatorname{rng} \Omega$ be the set of priorities used by $\mathcal{A}$. We start with an expression describing runs starting with the state q_k and with only finitely many occurrences of q_k on each branch. For a set $\eta \subseteq \xi$ of variables, we write $\bar{U}|_\eta$ for the subtuple $(U_x)_{x\in\eta}$. Let $V := D \times \{q_k\}$, let $y_0, y_1, \dots$ be new variables not in ξ, and set

$$T^k_{\bar{U}} := R^k_{q_k,\bar{U}} + \sum_{\zeta\cup\eta_0\cup\dots\cup\eta_{n-1}=\xi} (S^{\zeta,n}_0 + \dots + S^{\zeta,n}_{m-1})^{+y_0} \cdot_{y_0} R^k_{q_k,\bar{U}|_{\eta_0}} \cdot_{y_1} \dots \cdot_{y_{n-1}} R^k_{q_k,\bar{U}|_{\eta_{n-1}}},$$

where

- the sum ranges over all sequences $\zeta, \eta_0, \dots, \eta_{n-1}$ of subsets of ξ whose union is equal to ξ and such that $\eta_i \neq \eta_j$, for $i \neq j$, and
- $S^{\zeta,n}_0, \dots, S^{\zeta,n}_{m-1}$ is an enumeration of all expressions of the form $R^k_{q_k,\bar{U}|_\eta V\cdots V}$ where $\eta \subseteq \zeta$, $\upsilon \subseteq \{y_0, \dots, y_{n-1}\}$, and with $|\upsilon|$ copies of V that correspond to the variables $y \in \upsilon$.

Then $T^k_{\bar{U}}$ describes all trees that have a run with profile bounded by $\langle q_k, \bar{U}\rangle$ and such that every branch contains only finitely many occurrences of the state q_k.

Similarly, we obtain an expression for all such trees with possibly infinitely many occurrences of q_k by setting

$$\hat{T}^k_{\bar{U}} := T^k_{\bar{U}} + \sum_{\zeta\cup\eta_0\cup\dots\cup\eta_{n-1}=\xi} (S^{\zeta,n}_0 + \dots + S^{\zeta,n}_{m-1})^{\omega z} \cdot_{y_0} T^k_{\bar{U}|_{\eta_0}} \cdot_{y_1} \dots \cdot_{y_{n-1}} T^k_{\bar{U}|_{\eta_{n-1}}},$$

where the $S^{\zeta,n}_i$ are defined as above, except that there is an additional copy of V corresponding to the variable z.

If $\Omega(q_k)$ is odd, we can now set

$$R^{k+1}_\tau := R^k_\tau + \sum_{\zeta\cup\eta_0\cup\dots\cup\eta_{n-1}=\xi} R^k_{p,\bar{U}|_\zeta V\dots V} \cdot_{y_0} T^k_{\bar{U}|_{\eta_0}} \cdot_{y_1} \dots \cdot_{y_{n-1}} T^k_{\bar{U}|_{\eta_{n-1}}}.$$

where the variables $y_0, \dots, y_{n-1}$ are the ones corresponding to the n copies of the set V. If $\Omega(q_k)$ is even, we instead use

$$R^{k+1}_\tau := R^k_\tau + \sum_{\zeta\cup\eta_0\cup\dots\cup\eta_{n-1}=\xi} R^k_{p,\bar{U}|_\zeta V\dots V} \cdot_{y_0} \hat{T}^k_{\bar{U}|_{\eta_0}} \cdot_{y_1} \dots \cdot_{y_{n-1}} \hat{T}^k_{\bar{U}|_{\eta_{n-1}}}.$$

□

7. Conclusion

We have introduced the upwards-closed power-set monad $\mathbb{U}$ on Pos^{Ξ} and studied possible distributive laws between it and two monads of infinite trees: linear trees $\mathbb{T}$ and non-linear ones $\mathbb{T}^\times$. For the monad $\mathbb{T}$, we have shown in Theorems 3.19 and 3.22 that there exists a unique distributive law dist : $\mathbb{T}\mathbb{U} \Rightarrow \mathbb{U}\mathbb{T}$. For the monad $\mathbb{T}^\times$ on the other hand, we have proved in Theorem 3.23 that there is no distributive law $\mathbb{T}^\times\mathbb{U} \Rightarrow \mathbb{U}\mathbb{T}^\times$. Our main result (Theorem 4.36) states that, nevertheless, every set of the form $\mathbb{U}\mathbb{T}^\times A$ forms a $\mathbb{T}^\times$-algebra when equipped with a suitable product. The two examples in Section 5 and 6 show that this partial result is frequently sufficient for applications.

There are several possible directions where one can go from here. Of interest to language theorists would be to consider other functors similar to the power-set one, for instance the functor producing linear combinations over a given semiring, or similar analogues of the power-set functor for weighted languages.

More category-theoretically inspired considerations would include a more systematic study of when a distributive law with the power-set monad exists. In particular, it would be interesting to transfer the results in Section 3 from polynomial monads to quotients of such monads. Another avenue to pursue would be to generalise Theorem 3.23 to other monads than the power-set one by extracting the abstract properties of the power-set monad needed for the proof.

References

[Bec69] J. Beck. Distributive Laws. In B. Eckmann, editor, *Seminar on triples and categorical homology theory*, Lecture Notes in Mathematics 80, pages 119–140. Springer, 1969.

[BKS] M. Bojanczyk, B. Klin, and J. Salamanca. Monadic monadic second order logic. arXiv:2201.09969, unpublished.

[Blu20] A. Blumensath. Regular Tree Algebras. *Logical Methods in Computer Science*, 16:16:1–16:25, 2020.

[Blu21] A. Blumensath. Algebraic Language Theory for Eilenberg–Moore Algebras. *Logical Methods in Computer Science*, 17:6:1–6:60, 2021.

[Boj] M. Bojańczyk. Recognisable languages over monads. unpublished note, arXiv:1502.04898v1.

[Boj20] M. Bojańczyk. Languages Recognises by Finite Semigroups and their generalisations to objects such as Trees and Graphs with an emphasis on definability in Monadic Second-Order Logic. lecture notes, arXiv:2008.11635, 2020.

[CDD+22] C. Camino, V. Diekert, B. Dundua, M. Marin, and G. Sénizergues. Regular matching problems for infinite trees. *Logical Methods in Computer Science*, 18:25:1–25:38, 2022.

[Gar20] R. Garner. The Vietoris Monad and Weak Distributive Laws. *Appl. Categorical Struct.*, 28:339–354, 2020.

[GPA21] A. Goy, D. Petrisan, and M. Aiguier. Powerset-like monads weakly distribute over themselves in toposes and compact hausdorff spaces. In N. Bansal, E. Merelli, and J. Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12–16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPIcs*, pages 132:1–132:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.

[GTW02] E. Grädel, W. Thomas, and T. Wilke. *Automata, Logic, and Infinite Games.* LNCS 2500. Springer-Verlag, 2002.

[Jac04] B. Jacobs. Trace Semantics for Coalgebras. In J. A. and S. Milius, editors, *Proceedings of the Workshop on Coalgebraic Methods in Computer Science, CMCS 2004, Barcelona, Spain, March 27–29, 2004*, volume 106 of *Electronic Notes in Theoretical Computer Science*, pages 167–184, 2004.

[Löd21] C. Löding. Automata on infinite trees. In J.-É. Pin, editor, *Handbook of Automata Theory*, pages 265–302. European Mathematical Society, 2021.

[LT21] C. Löding and W. Thomas. Automata on finite trees. In J.-É. Pin, editor, *Handbook of Automata Theory*, pages 235–264. European Mathematical Society, 2021.

[MM07] E. Manes and P. S. Mulry. Monad compositions I: general constructions and recursive distributive laws. *Theory and Applications of Categories*, 18:172–208, 2007.

[MM08] E. Manes and P. S. Mulry. Monad compositions II: Kleisli strength. *Math. Struct. Comput. Sci.*, 18:613–643, 2008.

[Sch65] M. P. Schützenberger. On Finite Monoids Having Only Trivial Subgroups. *Information and Control*, 8:190–194, 1965.

[SN] D. I. Spivak and N. Niu. Polynomial Functors: A General Theory of Interaction. lecture notes.

[Tho90] W. Thomas. Automata on Infinite Objects. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 135–191. Elsevier, Amsterdam, 1990.

[Tho97] W. Thomas. Languages, Automata, and Logic. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3, pages 389–455. Springer, New York, 1997.

[UACM17] H. Urbat, J. Adámek, L.-T. Chen, and S. Milius. Eilenberg theorems for free. In *42nd International Symposium on Mathematical Foundations of Computer Science, MFCS 2017, August 21–25, 2017 – Aalborg, Denmark*, volume 83, pages 43:1–43:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017.

[ZM22] M. Zwart and D. Marsden. No-Go Theorems for Distributive Laws. *Log. Methods Comput. Sci.*, 18:13:1–13:61, 2022.