THE EXPANSION PROBLEM FOR INFINITE TREES

Achim Blumensath

1st July 2024

We study Ramsey like theorems for infinite trees and similar combinatorial tools. As an application we consider the expansion problem for tree algebras.

1 INTRODUCTION

While the theory of languages of infinite trees is well-established by now, it is far less developed than other formal language theories. Further progress in this direction is currently hampered by our lack of understanding of the combinatorial properties of infinite trees. In particular, for many purposes the currently known Ramsey type theorems for trees are simply not strong enough. What would be needed instead are, for instance, analogues of Simon factorisation trees for infinite trees. Such Ramsey arguments are ubiquitous in the study of languages of infinite objects. For instance, in the theory of ω -words they appear in the original complementation proof for Büchi automata (for a modern account see, e.g., [27]), when expanding a Wilke algebra to an ω -semigroup (see, e.g., [22]), or in the more recent work on distance automata and boundedness problems (see, e.g., [24, 25, 1, 18]).

There are at least two persistent problems when trying to extend our repertoire of combinatorial tools to infinite trees. The first one concerns the step from arbitrary trees to regular ones: while many arguments only work if the considered trees are regular, the only known method of reducing a given tree to an equivalent regular one is based on automata, which are not applicable in all contexts. For instance, in [12] all proofs work exclusively with regular trees, and only at the very end the authors transfer their results to arbitrary trees (which was possible in this particular case since the languages under consideration were regular and therefore uniquely determined by which regular trees they contain).

The second problem concerns trees that are highly-branching. Many of the known tools from the theory of ω -semigroups can be generalised to trees that are thin, i.e., that have only countably many infinite branches. But all attempts to extend them to (at least some) non-thin trees have failed so far. For example, in [13] the authors only consider languages of thin trees since their methods do not apply to non-thin ones. Later in [2], they then successfully adapted their approach to study unambiguous languages of non-thin trees, utilising the fact that trees in unambiguous languages are in a certain sense governed by their thin prefixes.

The only combinatorial methods known so far that work well even in light of the above issues are those based on automata and games since, in a certain sense, games provide a way to reduce a problem concerning the whole tree into one only involving a single branch. (In fact, one of the motivations for this paper stems from a wish to deeper understand how exactly this is achieved, in particular during the translation of a formula into an automaton. Theorem 5.24 below might be considered to give a partial answer.) Unfortunately, there are many questions that resist to being phrased in automata-theoretic or game-theoretic terms.

In this article we start by quickly reviewing the existing techniques to study combinatorial properties of infinite trees, presenting them in the unifying language of tree algebras from [7]. We then take a look at several new approaches. We determine how far they carry and what the problems are that prevent us from continuing. Our contributions are mainly conceptual. We raise many open questions, but provide few answers. None of the results below are very deep and several remain partial. The main purpose of the article is to draw attention to a problem I consider central for further progress.

It seems that such progress will likely not come from abstract considerations but by working on concrete problems. The recently developed algebraic approaches to languages of infinite trees [12, 5, 6, 13, 2, 14, 11, 7, 8, 9] seems to provide many opportunities to test our combinatorial tools. Our focus will therefore be on a particular application from this domain, one that we call the *Expansion Problem*: the problem of whether a given algebra whose product is defined only for some trees can be expanded to one whose product is defined everywhere, analogously to the expansion of a Wilke algebra (where the product is defined only for ultimately periodic words) to an ω -semigroup (where we can multiply arbitrary ω -words). This problem turns out to be a good test bed for the various approaches we consider. We solve it in some special cases, but none of our approaches is strong enough to solve the general case.

The overview of the article is as follows. We start in Section 2 with setting up the algebraic framework we will be working in. Section 3 contains a brief survey of the existing Ramsey Theorems for trees. The Expansion Problem is defined in Section 4, where we also recall some tools from [8] to prove uniqueness of expansions. The main technical part of the article are Sections 5 and 6, which contain two tools to study expansions. The first one are so-called *evaluations*, which are a weak form of a Simon tree, the second one are *consistent labellings*, which are somewhat similar to automata. The final two sections (7 and 8) contain two applications. The first one recalls results of [2] about a characterisation of unambiguous languages in terms of consistent labellings, while the second one uses consistent labellings to define classes of tree algebras with unique expansions.

Finally, let us highlight the concrete contributes of this article. (All terminology will be defined in the respective section below.)

- We streamline and generalise the definitions of two combinatorial tools from the literature: evaluations [23, 16] and consistent labellings [2].
- We prove the existence of expansions for MSO-definable \mathbb{T}^{reg} -algebras in Theorem 4.6.
- We prove the existence of certain evaluations in Theorems 5.23 and 5.24.
- We solve the expansion problem for thin trees in Section 5.1, and the expansion problems for deterministic and branch-continuous tree algebras in Section 8.

2 TREE ALGEBRAS

We start with a brief introduction to the algebraic framework we will be working in. A more detailed account can be found in [7, 8, 3] (in increasing order of abstractness). Let us fix notation and conventions. For $n < \omega$, we set [n] := $\{0, ..., n-1\}$. We denote tuples $\bar{a} = \langle a_0, ..., a_{n-1} \rangle$ with a bar. The empty tuple is $\langle \rangle$. The *range* of a function $f : A \to B$ is the set rng f := f[A]. We denote the disjoint union of two sets by A + B, and we denote the union of two functions $f : A \to B$ and $f' : A' \to B$ by $f + f' : A + A' \to B$.

Let us quickly recall some material from the theory of ω -semigroups (see, e.g., [22] for an introduction). An ω -semigroup is a two-sorted structure $\langle S_1, S_{\omega} \rangle$ with

three products

$$:: S_1 \times S_1 \to S_1, \quad :: S_1 \times S_\omega \to S_\omega, \quad \pi: (S_1)^\omega \to S_\omega$$

satisfying several associative laws. A *Wilke algebra* is a two-sorted structure (S_1, S_{ω}) with two products and an ω -power operation

$$\cdot : S_1 \times S_1 \to S_1, \quad \cdot : S_1 \times S_\omega \to S_\omega, \quad -^\omega : S_1 \to S_\omega$$

again satisfying several associative laws. The laws for the ω -power are

$$(ab)^{\omega} = a(ba)^{\omega}$$
 and $(a^n)^{\omega} = a^{\omega}$, for all $a, b \in S_1$ and $o < n < \omega$

One can show via a Ramsey argument that every finite Wilke algebra has a unique expansion to an ω -semigroup.

To model ranked trees, we work in a many-sorted setting where the sort of a tree represents the set of *variables* or *holes* appearing in it. Hence, we fix a countably infinite set X of variables and use the set $\Xi := \mathcal{P}_{fin}(X)$ of finite subsets of X as sorts. In addition, for technical reasons we equip the labels of our trees with an ordering. Hence, we will work with *partially ordered* Ξ -sorted sets which are families $A = (A_{\xi})_{\xi \in \Xi}$ where each component A_{ξ} is partial ordered. A function $f : A \to B$ between two such sets is then a family $f = (f_{\xi})_{\xi \in \Xi}$ of monotone functions $f_{\xi} : A_{\xi} \to B_{\xi}$. In the following we will for simplicity use the term *sorted set* for 'partially ordered Ξ -sorted set' and the term *function* for a function between such sets. Sometimes it is convenient to identify a sorted set $A = (A_{\xi})_{\xi \in \Xi}$ with its disjoint union $A = \sum_{\xi \in \Xi} A_{\xi}$. Then a function $f : A \to B$ corresponds to a sort-preserving and order-preserving function between the corresponding disjoint unions.

To make this article accessible to a wider audience, we have tried to keep the category-theoretic prerequisites at a minimum. Let us briefly recall some standard notions. A *functor* \mathbb{F} is an operation that maps every sorted set *A* to some sorted set $\mathbb{F}A$, and every function $f : A \to B$ between such sets to a function $\mathbb{F}f : \mathbb{F}A \to \mathbb{F}B$ such that \mathbb{F} preserves identity maps and composition of functions. A *natural transformation* $\tau : \mathbb{F} \Rightarrow \mathbb{G}$ between two functors \mathbb{F} and \mathbb{G} is a family $\tau = (\tau_A)_A$ (indexed by all sorted sets *A*) of functions $\tau_A : \mathbb{F}A \to \mathbb{G}A$ satisfying

$$\tau_A \circ \mathbb{F}f = \mathbb{G}f \circ \tau_B$$
, for all $f : A \to B$.

Usually, we omit the index A and simply write τ instead of τ_A .

Given a sorted set *A* an *A*-labelled tree is a possibly infinite tree *t* where the vertices are labelled by elements of *A* and the edges by variables from *X* in such a way that a vertex with a label $a \in A_{\xi}$ of sort ξ has exactly one outgoing edge labelled by *x*, for every $x \in \xi$ (and no other edges). We identify such a tree with a function $t : \text{dom}(t) \rightarrow A$, where dom(t) is the set of vertices of *t*. (We consider dom(t) to be a sorted set where $v \in \text{dom}(t)$ has the same sort as its label t(v).) As usual, we identify the vertices of *t* with finite sequences of directions. Since, in our case, we can take the variables for directions, this turns dom(t) into a prefix-closed subset of X^* . Using this identification, we can write the root of *t* as $\langle \rangle$. If there is an *x*-labelled edge from a vertex *u* to *v*, we call *v* the *x*-successor of *u*. We denote it by $\text{suc}_x(v)$. A branch of *t* is a maximal path starting at the root.

Definition 2.1. Let *A* be a sorted set.

(a) We set $\mathbb{T}^{\times}A := (\mathbb{T}_{\xi}^{\times}A)_{\xi \in \Xi}$ where $\mathbb{T}_{\xi}^{\times}A$ denotes the set of all $(A + \xi)$ -labelled trees *t* (where the elements in ξ are assumed to have sort \emptyset) satisfying the following conditions.

- Every variable $x \in \xi$ appears at least once in *t*.
- The root of *t* is not labelled by a variable.

(b) For a function $f : A \to B$, we denote by $\mathbb{T}^{\times} f : \mathbb{T}^{\times} A \to \mathbb{T}^{\times} B$ the function applying *f* to every label of the given tree (leaving the variables unchanged).

We need the following two operations on trees.

Definition 2.2. Let *A* be a sorted set.

(a) The *singleton operation* sing : $A \to \mathbb{T}^{\times}A$ maps every letter $a \in A_{\xi}$, to the tree sing(*a*) consisting of the root with label *a* attached to which is one leaf with label *x*, for every $x \in \xi$.

(b) The *flattening operation* flat : $\mathbb{T}^{\times}\mathbb{T}^{\times}A \to \mathbb{T}^{\times}A$ is a generalisation of term substitution. It takes a tree *t* labelled by trees $t(v) \in \mathbb{T}^{\times}A$ and combines them into a single tree as follows (see Figure 1).

- We take the disjoint union of all trees t(v), for v ∈ dom(t) (where, if t(v) = x ∈ X is a variable and not a tree, we treat t(v) as a 1-vertex tree whose root is labelled x);
- from each component t(v) such that t(v) is a proper tree and not just a variable, we delete every vertex labelled by a variable x ∈ X;
- we redirect every edge of t(v) leading to such a deleted vertex to the root of t(ux), where ux is the x-successor of v in t; and





• we unravel the resulting graph into a tree.

Remark. The triple $\langle \mathbb{T}^{\times}, \text{flat}, \text{sing} \rangle$ forms what is called a *monad* in category-theoretical language, which means that flat : $\mathbb{TT} \Rightarrow \mathbb{T}$ and sing : Id $\Rightarrow \mathbb{T}$ are natural transformations satisfying the following three equations.

flat \circ sing = id, flat $\circ \mathbb{T}^{\times}$ sing = id, flat $\circ \mathbb{T}^{\times}$ flat = flat \circ flat.

Definition 2.3. Let $t \in \mathbb{T}^{\times}A$ be a tree.

(a) The *tree order* \leq is the ordering on dom(*t*) defined by

 $u \le v$: iff *u* lies on the path from the root to *v*.

(b) A *factorisation* of *t* is a tree $T \in \mathbb{T}^{\times}\mathbb{T}^{\times}A$ with flat(T) = t. We call each tree T(v), for $v \in \text{dom}(T)$, a *factor* of *t*.

(c) Given vertices u and $\bar{v} = (v_x)_{x \in \xi}$ of t such that \bar{v} forms an antichain (with respect to \leq) and $u < v_x$, for all x, we define

 $[u, \bar{v}) \coloneqq \{ w \in \operatorname{dom}(t) \mid u \leq w \text{ and } v_x \nleq w, \text{ for all } x \}.$

We call ξ the *sort* of $[u, \bar{v})$.

We denote by $t[u, \bar{v})$ the restriction of $t : \text{dom}(t) \to A$ to the set $[u, \bar{v}) \cup \bar{v}$ (where v_x is labelled by the variable x while all other vertices have the same label as in t). We call $t[u, \bar{v})$ the *factor* of t between u and \bar{v} . In the special case where $\xi = \emptyset$, we obtain the *subtree* of t rooted at u, which we usually denote by $t|_u$.

Remark. (a) When identifying the vertices of a tree with words in X^* , the tree order \leq is just the prefix ordering. This also explains our notation () for the root.

(b) A factor $t[u, \bar{v})$ may contain additional variables besides those at the vertices \bar{v} . More precisely, we have $t[u, \bar{v}) \in \mathbb{T}_{\xi \cup \zeta}^{\times} A$ where ξ is the sort of $[u, \bar{v})$ and ζ is the set of those variables of t that appear at some vertex $w \in [u, \bar{v})$.

There are several special classes of trees we are interested in below.

Definition 2.4. (a) A *submonad* of \mathbb{T}^{\times} is a functor \mathbb{T}° such that

- $\mathbb{T}^{\circ}A \subseteq \mathbb{T}^{\times}A$, for every sorted set *A*,
- $\mathbb{T}^{\circ} f = \mathbb{T}^{\times} f$, for every function *f*, and
- $\mathbb{T}^{\circ}A$ is closed under flat and sing, that is,

flat
$$(t) \in \mathbb{T}^{\circ}A$$
, for all $t \in \mathbb{T}^{\circ}\mathbb{T}^{\circ}A$,
sing $(a) \in \mathbb{T}^{\circ}A$, for all $a \in A$.

We write $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{\times}$ to denote this fact.

(b) We are particularly interested in the following submonads. \mathbb{T} denotes the subset of all *linear trees*, i.e., trees where each variable appears exactly once. \mathbb{T}^{fin} denotes the subset of *finite linear trees*, \mathbb{T}^{reg} the subset of *regular linear trees*, \mathbb{T}^{thin} the subset of *thin linear trees*, i.e., trees with only countably many infinite branches, and $\mathbb{T}^{\text{wilke}} := \mathbb{T}^{\text{thin}} \cap \mathbb{T}^{\text{reg}}$ the subset of all trees that are thin and regular. The corresponding classes of non-linear trees are denoted by $\mathbb{T}^{\times \text{fin}}$, $\mathbb{T}^{\times \text{reg}}$, etc.

Remark. Note that $\mathbb{T}^{\times \text{thin}}$ and $\mathbb{T}^{\times \text{wilke}}$ do not form submonads of \mathbb{T}^{\times} , since they are not closed under flat. This is different for \mathbb{T}^{thin} and $\mathbb{T}^{\text{wilke}}$, which are in fact submonads of \mathbb{T} .

In algebraic language theory one uses algebras (usually finite ones) to recognise languages. In our setting these algebras take the following form.

Definition 2.5. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{\times}$.

(a) A \mathbb{T}° -algebra $\mathfrak{A} = \langle A, \pi \rangle$ consists of a sorted set A and a product $\pi : \mathbb{T}^{\circ}A \to A$ satisfying

 $\pi \circ \operatorname{sing} = \operatorname{id} \quad \operatorname{and} \quad \pi \circ \mathbb{T}^{\circ} \pi = \pi \circ \operatorname{flat}.$

The first equation is called the *unit law*, the second one the *associative law*.

(b) A \mathbb{T}° -algebra \mathfrak{A} is *finitary* if it is finitely generated and every domain A_{ξ} is finite.

(c) A *morphism* between two \mathbb{T}° -algebras $\mathfrak{A} = \langle A, \pi \rangle$ and $\mathfrak{B} = \langle B, \pi \rangle$ is a function $\varphi : A \to B$ commuting with the respective products in the sense that

$$\varphi \circ \pi = \pi \circ \mathbb{T}^{\mathsf{o}} \varphi \,.$$

(d) A \mathbb{T}° -algebra \mathfrak{A} recognises a language $K \subseteq \mathbb{T}^{\circ}_{\xi}\Sigma$ if there exists a morphism $\eta : \mathbb{T}^{\circ}\Sigma \to \mathfrak{A}$ with $K = \eta^{-1}[P]$, for some $P \subseteq A_{\xi}$.

Example. The following algebra $\mathfrak{A} = \langle A, \pi \rangle$ recognises the language *K* of all trees $t \in \mathbb{T}_{\emptyset}^{\times} \{a, b\}$ that contain at least one letter *a*. For each sort ξ , we use two elements $o_{\xi, 1\xi}$. Hence,

$$A_{\xi} \coloneqq \{\mathbf{o}_{\xi}, \mathbf{1}_{\xi}\}.$$

The product is defined by

 $\pi(t) \coloneqq \begin{cases} 1 & \text{if } t \text{ contains the label 1,} \\ 0 & \text{otherwise.} \end{cases}$

Then $K = \varphi^{-1}(1_{\emptyset})$, where $\varphi : \mathbb{T}^{\times}\{a, b\} \to A$ is the morphism mapping *a* to 1 and *b* to 0.

We will often use the usual term notation for trees and elements in an algebra. That is, for $s \in \mathbb{T}_{\xi}^{\times} A$ and a ξ -tuple \bar{r} of trees and/or variables, we denote by $s(\bar{r})$ the tree obtained from s by replacing every variable $x \in \xi$ by the tree r_x . Similarly, for an algebra \mathfrak{A} , an element $a \in A_{\xi}$, and a ξ -tuple \bar{b} of elements and/or variables, we denote by $a(\bar{b})$ the product of the tree $s(\bar{r})$ where $s := \operatorname{sing}(a)$ and $r_x := \operatorname{sing}(b_x)$ (or $r_x := b_x$, if b_x is a variable).

A complication of the theory of infinite trees is the fact that some finitary \mathbb{T}^{\times} -algebras recognise non-regular languages [14]. For this reason we have to consider a smaller class of algebras.

Definition 2.6. (a) We denote *first-order logic* by FO and *monadic second-order logic* by MSO.

(b) Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{\times}$. A \mathbb{T}° -algebra \mathfrak{A} is MSO-*definable* if it is finitary and there exists a finite set $C \subseteq A$ of generators of \mathfrak{A} with the following property: for every $a \in A$, there exists an MSO-formula φ_a such that

$$t \models \varphi_a$$
 : iff $\pi(t) \ge a$, for all $t \in \mathbb{T}^{\circ}C$.

If all formulae φ_a belong to FO, we call \mathfrak{A} FO-*definable*.

Example. The algebra from the previous example is FO-definable. (The formulae φ_0 and φ_1 only have to check whether or not the given tree contains the label 1.)

Using this notion we obtain the following characterisation (for proofs see [7, 8]).

Theorem 2.7. A finitary \mathbb{T}^{\times} -algebra \mathfrak{A} is MSO-definable if, and only if, every language recognised by \mathfrak{A} is regular.

The definition above is not very enlightening as it is basically just a restatement of the preceding theorem. Although a more algebraic characterisation has been found in [7], a simpler one would be appreciated. In particular, it would be nice to find a system of inequalities axiomatising the class of MSO-definable algebras.

Open Question. Find a concrete description of a system of inequalities that axiomatises the class of MSO-definable \mathbb{T}^{\times} -algebras.

By general arguments we know that such a system of inequalities exists, although it might be infinite and the terms in the inequalities are in general profinite (see [8] for the details).

3 PARTITION THEOREMS FOR TREES

Let us start with a brief overview of the existing partition theorems for trees, followed by some remarks on how they might be extended and how they might not. The seminal partition theorem for trees is the one by Milliken.

Definition 3.1. Let $t \in \mathbb{T}A$ be a tree.

(a) A factor $[u, \bar{v})$ of *t* is *properly embedded* if $\bar{v} = (v_x)_{x \in \xi}$ and $\operatorname{suc}_x(u) \leq v_x$, for all $x \in \xi$, where ξ is the sort of *u*. Similarly, we say that a sorted set $P \subseteq \operatorname{dom}(t)$ is *properly embedded* if, for every vertex $u \in P$ that is not maximal in *P* and for every *x*-successor v_x of *u*, the subtree $t|_{v_x}$ has a unique minimal vertex that belongs to *P*. (Thus every properly embedded set is comprised of one or several properly embedded factors.)

(b) A *proper labelling* of *t* is a function λ that assigns a colour to every properly embedded factor of *t*.

(c) Let λ be a proper labelling of t. A sorted set $H \subseteq \text{dom}(t)$ is *homogeneous* (with respect to λ) if

$$\lambda([u,\bar{v})) = \lambda([u',\bar{v}')), \text{ for all } u, u', v_x, v'_x \in H \text{ such that } [u,\bar{v}) \text{ and} \\ [u',\bar{v}') \text{ are properly embedded.}$$

L

Theorem 3.2 (Milliken [21]). Let $t \in \mathbb{T}A$ be an infinite tree without leaves and *C* a finitary set of colours. For every proper labelling λ of *t* with colours from *C*, there exists an infinite homogeneous set *H* that is properly embedded in *t*.

Remark. The actual theorem by Milliken is stronger than the above version, but the extra strength is not relevant in our context.

The limitation of this theorem is that it does not give us any information about factors $[u, \bar{v})$ whose end-points do not belong to the set *H*. For a stronger statement we need additional assumptions on the labelling. For instance, for labellings of finite words, there is the Factorisation Tree Theorem of Simon [24] which states that, if the labelling is additive (i.e., the colours form a semigroup), we can recursively factorise the given word into homogeneous parts. This theorem has been adapted to trees by Colcombet [18] as follows.

Definition 3.3. Let $t \in \mathbb{T}A$ be a tree and $\mathfrak{S} = \langle S_1, S_\omega \rangle$ an ω -semigroup.

(a) An \mathfrak{S} -*labelling* of *t* is a function λ that maps every edge *e* of *t* to a semigroup element $\lambda(e) \in S_1$. Each such function can be extended to all non-empty (finite or infinite) paths $p = (v_i)_i$ by setting

$$\lambda(p) \coloneqq \prod_i \lambda(v_i, v_{i+1}).$$

For u < v, we will also use the notation

 $\lambda(u, v) \coloneqq \lambda(p)$, where *p* is the path from *u* to *v*.

(b) Given a function $\sigma : \operatorname{dom}(t) \to [k]$, we define a binary relation \sqsubset_{σ} on $\operatorname{dom}(t)$ (see Figure 2) by

$$x \sqsubset_{\sigma} y$$
 : iff $x \prec y$, $\sigma(x) = \sigma(y)$, and
 $\sigma(z) \le \sigma(x)$, for all $x \le z \le y$.

As usual, \sqsubseteq_{σ} denotes the reflexive version of \sqsubset_{σ} .

(c) A weak Ramseyan split of an \mathfrak{S} -labelling λ is a function $\sigma : \operatorname{dom}(t) \to [k]$ such that

$$\lambda(x, y) = \lambda(x, y) \cdot \lambda(x', y'), \quad \text{for all } x \sqsubset_{\sigma} y \text{ and } x' \sqsubset_{\sigma} y'$$

such that $y \sqsubseteq_{\sigma} y' \text{ or } y' \sqsubseteq_{\sigma} y.$

ы

Figure 2: A function σ (along a single branch of *t*) and the corresponding relation \equiv_{σ} , indicated via the grey bars.

Theorem 3.4 (Colcombet). Let t be a tree and \mathfrak{S} a finite ω -semigroup. Every \mathfrak{S} -labelling λ of t has a weak Ramseyan split. Furthermore, this split is MSO-definable.

As an example of how to apply this theorem, let us mention the following result from [6] that can be used to turn arbitrary trees into regular ones while preserving an edge labelling.

Definition 3.5. Let \mathfrak{S} be an ω -semigroup and λ an \mathfrak{S} -labelling of some tree *t*. We write

 $\lim \lambda := \left\{ \lambda(\beta) \mid \beta \text{ a branch of } t \right\}.$

Theorem 3.6. Let \mathfrak{S} be a finite ω -semigroup. For every \mathfrak{S} -labelling λ of a tree t, there exists a regular tree t_0 and a regular \mathfrak{S} -labelling λ_0 of t_0 such that $\lim \lambda_0 = \lim \lambda$.

Ц

The proof consists in fixing a weak Ramseyan split of λ and using it to replace certain subtrees of *t* by back-edges. The unravelling of the resulting graph is the desired regular tree t_0 .

As a second application let us see how to use Ramseyan splits to evaluate products in a \mathbb{T} -algebra.

Definition 3.7. Let \mathfrak{A} be a \mathbb{T} -algebra.

(a) The *canonical edge labelling* λ of a tree $t \in \mathbb{T}A$ is defined by

 $\lambda(u, v) \coloneqq \pi(t[u, v)), \quad \text{for } u < v.$

(b) Let *L* be either FO or MSO, and let λ be a given type of labelling, i.e., a function assigning to each tree *t* some labelling λ_t (which may be an edge labelling or a vertex labelling). We say that \mathfrak{A} is *L*-definable *with respect to* λ if \mathfrak{A} is finitary and, for every finite $C \subseteq A$ and every $a \in A$, there exists an *L*-formula φ_a such that

 $\langle t, \lambda_t \rangle \vDash \varphi_a$ iff $\pi(t) \ge a$, for all $t \in \mathbb{T}C$,

where $\langle t, \lambda_t \rangle := \langle \operatorname{dom}(t), \leq, (\operatorname{suc}_x)_x, (P_c)_{c \in C}, (R_s)_{s \in S} \rangle$ denotes the usual encoding of *t* as a relational structure with additional relations *R_s* (either unary or binary) for the labelling λ_t . In case λ_t is an edge labelling, we assume that the relations *R_s* contain the labels for all vertices u < v, not only those where *v* is a successor of *u*.

Remark. Note that the canonical edge labelling of a tree $t \in \mathbb{T}_{\xi}A$ is an \mathfrak{S} -labelling for the ω -semigroup $\mathfrak{S} = \langle S_1, S_{\omega} \rangle$ with elements

$$S_1 \coloneqq \sum_{\eta \subseteq \xi} A_{\eta + \{z\}}$$
 and $S_\omega \coloneqq \sum_{\eta \subseteq \xi} A_{\eta}$,

where $z \in X$ is some fixed variable with $z \notin \xi$ that we use to define the semigroup product.

Proposition 3.8. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{\times}$. Every finitary \mathbb{T}° -algebra is FO-definable with respect to the canonical edge labelling.

Proof. Fix $a \in A_{\xi}$ and $C \subseteq A$. The formula φ_a checks whether the given tree *t* has a leaf. If this is the case it picks one, say *v*, and then checks whether

 $\lambda(\langle \rangle, v) \cdot t(v) \geq a$.

Otherwise, for every sort ζ used by some element of *C*, we fix some variable $z \in \zeta$. If we start at the root of *t* and follow the successors labelled by one of these chosen variables, we obtain an infinite branch of *t*. This branch is FO-definable. The formula φ_a checks that the branch contains an infinite sequence $v_0 < v_1 < \ldots$ of vertices such that

$$\lambda(v_i, v_k) = \lambda(v_0, v_1), \quad \text{for all } i < k < \omega,$$

$$\lambda(\langle \rangle, v_0) \cdot \lambda(v_0, v_1)^{\omega} \ge a.$$

This fact can be expressed in first-order logic using a trick of Thomas [26, Lemma 1.4].

We can improve this result by encoding the corresponding edge labelling by a vertex labelling.

Definition 3.9. Let *L* be either FO or MSO, and let $\tau : \text{dom}(t) \to C$ be a vertex labelling of some tree $t \in \mathbb{T}^{\times}A$.

(a) We say that τ is *L*-definable if, for every $c \in \operatorname{rng} \tau$, there exists an *L*-formula $\varphi_c(x)$ such that

 $\tau(v) = c$ iff $t \models \varphi_c(v)$, for every $v \in \text{dom}(t)$.

The definition for labellings of vertices is analogous.

(b) We say that the product π of an algebra \mathfrak{A} is *L*-definable on *t* in terms of a vertex labelling $\tau : \operatorname{dom}(t) \to C$ if, for every $a \in A$, there exists an *L*-formula $\psi_a(x, \bar{y})$ such that

 $\pi(t[u, \bar{v})) = a$ iff $\langle t, \tau \rangle \vDash \psi_a(u, \bar{v})$, for every factor $[u, \bar{v})$.

The following result is based on a similar theorem by Colcombet [17].

Theorem 3.10. Let \mathfrak{A} be a finitary \mathbb{T} -algebra. For every finite set $C \subseteq A$, there exists a finite (unsorted) set S such that every $t \in \mathbb{T}C$ has a labelling $\sigma : \operatorname{dom}(t) \to S$ such that the product on t is FO-definable in terms of σ . Furthermore, if \mathfrak{A} is MSO-definable, then so is σ .

Proof. We have shown in Proposition 3.8 that \mathfrak{A} is FO-definable with respect to the canonical edge labelling λ . Consequently, it is sufficient to find a vertex labelling σ such that λ is FO-definable in terms of σ .

By Theorem 3.4, there exists an MSO-definable weak Ramseyan split σ_0 : dom $(t) \rightarrow [k]$ for λ . Unfortunately, λ does not need to be FO-definable in terms of σ_0 . Therefore, we define an extended labelling

 $\sigma: \operatorname{dom}(t) \to [k] \times A_{\{z\}} \times A_{\{z\}}$

(for some fixed but arbitrary variable *z*) as follows. For a vertex $v \in \text{dom}(t)$, let p(v) be the predecessor of *v* (if it exists) and let $q(v) \prec v$ be the maximal vertex (if it exists) such that $q(v) \sqsubset_{\sigma_0} v$. Fixing an arbitrary element $a_0 \in A_{\{z\}}$, we set

$$\sigma(v) \coloneqq \begin{cases} \left\langle \sigma_{\circ}(v), \lambda(p(v), v), \lambda(q(v), v) \right\rangle & \text{if } q(v) \text{ is defined,} \\ \left\langle \sigma_{\circ}(v), \lambda(p(v), v), a_{\circ} \right\rangle & \text{if } p(v) \text{ is defined, but } q(v) \\ & \text{is not,} \\ \left\langle \sigma_{\circ}(v), a_{\circ}, a_{\circ} \right\rangle & \text{otherwise.} \end{cases}$$

Note that, if \mathfrak{A} is MSO-definable, then so is σ . Hence, it remains to prove that λ is FO-definable in $\langle t, \sigma \rangle$.

Consider two vertices u < v. We can compute $\lambda(u, v)$ from σ as follows. We start by defining the sequence $u_0 < \cdots < u_m \le v_n < \cdots < v_o$ where $u_0 = u, v_0 = v$, u_{i+1} is the minimal vertex bettween u_i and v such that $\sigma_0(u_{i+1}) > \sigma_0(u_i)$, and v_{i+1} is the maximal vertex bettween u and v_i such that $\sigma_0(v_{i+1}) > \sigma_0(v_i)$. Note that this sequence if FO-definable and its length is bounded in terms of k. Since

$$\lambda(u,v) = \lambda(u_0, u_1) \cdot \dots \cdot \lambda(u_{m-1}, u_m)$$
$$\cdot \lambda(u_m, v_n)$$
$$\cdot \lambda(v_n, v_{n-1}) \cdot \dots \cdot \lambda(v_1, v_0)$$

(where we drop the factor $\lambda(u_m, v_n)$ in case $u_m = v_n$), we can evaluate $\lambda(u, v)$ in first-order logic, provided that we can define the values $\lambda(u_i, u_{i+1}), \lambda(u_m, v_n)$, and $\lambda(v_{i+1}, v_i)$.

Let us explain how to compute $\lambda(u_i, u_{i+1})$. The other two cases are analogous. To do so, let m < k be a number and u' < v' two vertices such that

$$\sigma_{o}(w) \leq m$$
, for all $u' \leq w < v'$.

By induction on *m*, we will construct an FO-formula computing $\lambda(u', v')$. Let w_0 be the minimal vertex $u' \le w_0 < v'$ with $\sigma_0(w_0) = m$, let w_2 be the maximal one, and let w_1 be the minimal vertex $w_0 < w_1 \le w_2$ with $\sigma_0(w_1) = m$. We distinguish several cases. If v' is the successor of u', we can read off $\lambda(u', v')$ from $\sigma(v')$. If w_0 does not exist, we can use the formula from the inductive hypothesis. If $w_0 = w_2$, we have

$$\lambda(u',v') = \lambda(u',w_{o}) \cdot \lambda(w_{o},v'),$$

where both factors can be computed by inductive hypothesis. Finally, consider the case where $w_0 < w_2$. Then $w_0 \sqsubset_{\sigma_0} w_1 \sqsubseteq_{\sigma_0} w_2$ implies that

$$\lambda(u',v') = \lambda(u',w_0) \cdot \lambda(w_0,w_2) \cdot \lambda(w_2,v')$$
$$= \lambda(u',w_0) \cdot \lambda(w_0,w_1) \cdot \lambda(w_2,v'),$$

where the middle factor can be read off from $\sigma(w_1)$ and the other two factors can be computed by inductive hypothesis.

The problem with the above theorems is that they require access to the canonical edge-labelling λ and, in order to obtain this labelling, we need to be able to compute products $\pi([u, v))$ where the factors [u, v) are usually infinite.

Unfortunately, in many applications we only know the products of *finite* factors. For instance, we cannot combine Theorems 3.6 and 3.10 to conclude that every tree $t \in \mathbb{T}A$ over a finitary \mathbb{T} -algebra \mathfrak{A} can be replaced by a regular one with the same product since we do not know whether or not the regular labelling λ_0 constructed in Theorem 3.6 is a canonical edge-labelling. It seems that a good first step towards progress would be to prove a finitary version of Theorem 3.10 like the following one.

Conjecture. Let \mathfrak{A} be a finitary \mathbb{T}^{fin} -algebra. For every finite set $C \subseteq A$, there exists a finite (unsorted) set S such that every $t \in \mathbb{T}C$ has a labelling $\sigma : \operatorname{dom}(t) \to S$ such that, for every finite factor [u, v] of t,

 $\pi(t[u, v))$ can be computed from σ by an FO-formula.

We conclude this section with a counterexample showing that some natural ways to approach this conjecture (or a similar one) do not work. It turns out that in general, if we want to compute $\pi(t[u, \bar{v}))$, we need to know how the factor $[u, \bar{v})$ is embedded in the tree. Just looking at the values $\sigma(u)$ and $\sigma(v_i)$ provided by some labelling σ is not enough.

Definition 3.11. Let $t \in \mathbb{T}^{\times}A$ be a tree.

(a) For $u, v \in \text{dom}(t)$, we denote by $u \sqcap v$ their infimum in the tree order \leq .

(b) The *branching pattern* of a tuple \bar{v} of vertices is the partial order consisting of the root of *t* and all vertices of the form $v_i \sqcap v_j$ where each edge $u \prec w$ is labelled by the variable *x* such that $suc_x(u) \leq w$.

(c) The *branching type* of a factor $[u, \bar{v})$ of *t* is the isomorphism type of the branching patterns of \bar{v} in the subtree $t|_u$. Alternatively, we can define the branching type as the atomic type of \bar{v} in the structure $\langle t|_u, (\prec_x)_x, \neg, u \rangle$ where

$$u <_x v$$
 : iff $\operatorname{suc}_x(u) \le v$.

Example. We construct a finitary \mathbb{T} -algebra \mathfrak{A} such that, for every $t \in \mathbb{T}A$ with factors $[u, \tilde{v})$ and $[u', \tilde{v}')$,

$$\pi(t[u,\bar{v})) = \pi(t[u',\bar{v}')) \quad \text{implies} \quad [u,\bar{v}) \text{ and } [u',\bar{v}') \text{ have the same}$$

branching type.

For $\xi \in \Xi$, let A_{ξ} be the set of all branching types of trees of sort ξ . (Up to the labelling of the edges, there are only finitely many of such types. We may assume that every sort ξ appearing in *t* is a finite initial segment of ω .) Note that, given a

tree $s \in \mathbb{TT}\Sigma$, we can compute the branching type of flat(*s*) from the types of s(v), for $v \in \text{dom}(s)$. Here the branching type of a tree is defined as the branching type of the factor $[\langle \rangle, \bar{v} \rangle$ where \bar{v} are the vertices labelled by a variable. It follows that there exists a function $\pi : \mathbb{T}A \to A$ satisfying $\pi \circ \mathbb{T}\tau = \tau \circ \pi$, where $\tau : \mathbb{T}\Sigma \to A$ is the function mapping each tree to its branching type. It follows that $\mathfrak{A} := \langle A, \pi \rangle$ is a finitary \mathbb{T} -algebra with the desired property.

Let $t \in \mathbb{T}\Sigma$ be an infinite binary tree and let λ be the labelling mapping a factor $[u, \bar{v})$ to its branching type. We claim that there is no function $\sigma : \operatorname{dom}(t) \to C$ with a finite codomain *C* such that the label $\lambda([u, \bar{v}))$ only depends on the values $\sigma(u)$ and $\sigma(v_x)$. We fix a vertex $u \in \operatorname{dom}(t)$ such that the set

$$C_{o} \coloneqq \{ \sigma(v) \mid v \geq u \}$$

is minimal. Set $c := \sigma(u)$. By choice of u there are vertices $v_0, v_1 \in \sigma^{-1}(c)$ with $suc_0(u) \le v_0$ and $suc_1(u) \le v_1$. Similarly, we can find $v'_1 \in \sigma^{-1}(c)$ with $suc_1(v_0) \le v'_1$. Then $suc_0(u) \le v'_1$ implies that $\lambda([u, v_0v_1)) \ne \lambda([u, v_0v'_1))$, but all four vertices have the same colour c.

We conclude this section by briefly mentioning two alternative approaches.

Remark. When talking about partition theorems for trees, we also have to mention automata. Every automaton can be seen as a prescription producing labellings (runs) of trees. The advantage of automata is that they can be used even if we know very little about the underlying algebra. In particular, they can be used in cases where we can only evaluate finite trees. Their disadvantage is that runs are usually not unique and that every run only contains a limited amount of information. For instance, in general there is no automaton that allows us to evaluate every factor $\pi(t[u, v))$ of a given tree t, only factors of a fixed arity.

Remark. The proof of Theorem 3.4 is based on semigroup-theoretic methods, in particular, it makes heavy use of Green's relations. It looks plausible that, in order to prove a stronger partition theorem for trees, we have to develop a similar theory of Green's relations for tree algebras. As it turns out, it is rather straightforward to generalise these relations to the setting of monoidal categories where all homsets are finite. (We omit the details since the statements and proofs are virtually identical to those for semigroups.) Furthermore, every \mathbb{T}^{fin} -algebra \mathfrak{A} can be seen as such a category \mathcal{A} where the objects are the sorts $\xi \in \Xi$ and the hom-sets $\mathcal{A}(\xi, \zeta)$ are given by $(A_{\xi})^{\zeta}$. The question is how to apply the resulting theory to the case at hand. The main problem with doing so seems to be that, in general, a finitary \mathbb{T}^{fin} -algebra can have infinitely many J-classes.

4 EXPANSIONS AND DENSE SUBMONADS

When looking for a strengthening of the results in the previous section it is always useful to have an application in mind that can serve as a test case and reality check. The following problem on expansions seems to be a good candidate for this purpose.

Definition 4.1. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{1} \subseteq \mathbb{T}^{\times}$ be submonads. We say that a \mathbb{T}° -algebra \mathfrak{A}_{\circ} is a *reduct* of a \mathbb{T}^{1} -algebra $\mathfrak{A}_{1} = \langle A, \pi \rangle$ if $\mathfrak{A}_{\circ} = \langle A, \pi_{\circ} \rangle$ where $\pi_{\circ} : \mathbb{T}^{\circ}A \to A$ is the restriction of $\pi_{1} : \mathbb{T}^{1}A \to A$. In this case, we call \mathfrak{A}_{1} a \mathbb{T}^{1} -expansion of \mathfrak{A}_{\circ} .

Expansion Problem. Given monads $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{1} \subseteq \mathbb{T}^{\times}$, which \mathbb{T}° -algebras have \mathbb{T}^{1} -expansions? And for which algebras are these expansions unique?

The motivating example of an expansion problem is the result that every finite Wilke algebra has a unique expansion to an ω -semigroup. For trees, problems of this kind turn often out to be quite hard and seem to require advanced techniques from combinatorics. In this article, we develop tools that help answering such questions, with a focus on the tree monads we have defined above. Of particular interest are the cases $\mathbb{T}^{\text{reg}} \subseteq \mathbb{T}$ and $\mathbb{T}^{\text{thin}} \subseteq \mathbb{T}$. Unfortunately, these two cases are also the hardest ones and seem to be out of reach with our current techniques.

We start our investigation with recalling some results from [8] that can be used to prove the uniqueness of expansions, if not their existence.

Definition 4.2. A submonad $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{1}$ is *dense* in \mathbb{T}^{1} over a class C of \mathbb{T}^{1} -algebras if, for all algebras $\mathfrak{A} \in C$, subsets $C \subseteq A$, and trees $s \in \mathbb{T}^{1}C$, there exists a tree $s^{\circ} \in \mathbb{T}^{\circ}C$ with $\pi(s^{\circ}) = \pi(s)$.

Example. The notion of denseness is a generalisation of the fact from the theory of ω -semigroups that every infinite product has a factorisation of the form ab^{ω} . This translates to the fact that, over the class of all finite ω -semigroups, the monad for Wilke algebras is dense in the monad for ω -semigroups: for every infinite word u in a finite ω -semigroup there is an ultimately periodic word u° with the same product. Details can be found in [8].

We have shown in Lemma 4.13 (a) of [8] that denseness implies the uniqueness of expansions (if they exist). One technical requirement of the proof is that we need to assume that the class in question is closed under binary products. Such products are defined in the usual way: the product $\mathfrak{A} \times \mathfrak{B}$ of two algebras has the universe $A \times B := (A_{\xi} \times B_{\xi})_{\xi \in \Xi}$ and the product is defined component-wise, i.e.,

the product of a tree $t \in \mathbb{T}(A \times B)$ is defined by taking the projections of *t* to the two components, multiplying them separately, and returning the pair of values obtained in this way.

Proposition 4.3. Let \mathbb{T}° be dense in \mathbb{T}^{1} over some class *C* that is closed under binary products. Then every \mathbb{T}° -algebra \mathfrak{A} has at most one \mathbb{T}^{1} -expansion that belongs to *C*.

Remark. The class of MSO-definable \mathbb{T}^{\times} -algebras is closed under binary products [7]. The same holds for MSO-definable \mathbb{T} -algebras.

The fact that a regular language is uniquely determined by which regular trees it contains can be generalised to the following theorem (which is a consequence of Theorem 10.1 of [8]).

Theorem 4.4. \mathbb{T}^{reg} is dense in \mathbb{T} over the class of all MSO-definable \mathbb{T} -algebras, and $\mathbb{T}^{\times \text{reg}}$ is dense in \mathbb{T}^{\times} over the class of all MSO-definable \mathbb{T}^{\times} -algebras.

It is currently unknown whether this property characterises the class of MSOdefinable algebras.

Open Question. Let \mathfrak{A} be a finitary \mathbb{T}^{\times} -algebra such that, for every $C \subseteq A$ and every $t \in \mathbb{T}C$, there is some $t^{\circ} \in \mathbb{T}^{\times \operatorname{reg}}C$ with $\pi(t^{\circ}) = \pi(t)$. Is \mathfrak{A} MSO-definable?

For linear trees we can strengthen Theorem 4.4 to include existence. The proof requires some tools from automata theory (for details we refer the reader to [28, 20]). We only consider automata over trees of sort Ø. When we need an automaton to read a tree $t \in \mathbb{T}_{\xi}^{\times} \Sigma$, we will encode it as an element of $\mathbb{T}_{\emptyset}^{\times} (\Sigma + \xi)$. Therefore, we will use (alternating) tree automata of the form $\mathcal{A} = (Q, \Sigma + \xi, \Delta, q_o, \Omega)$ where Q is the set of states, $\Sigma + \xi$ the input alphabet, $\Omega : Q \to \omega$ a priority function, and Δ the transition relation. The latter consists of triples of the form $\langle p, a, (P_z)_{z \in \mathcal{I}} \rangle$ for a state $p \in Q$, a letter $a \in (\Sigma + \xi)_{\zeta}$, and sets of states $P_z \subseteq Q$. The behaviour of such an automaton \mathcal{A} on a given input tree $t \in \mathbb{T}_{k}^{\times} \Sigma$ is defined via a certain parity game \mathcal{G} , called the Automaton-Pathfinder game of \mathcal{A} on t. The two players in this game are called Automaton and Pathfinder. The positions for the former are the pairs in dom(*t*) × *Q*, while the positions for the latter are those in dom(*t*) × Δ . In a position $\langle v, q \rangle$, Automaton chooses a transition $\langle p, a, \bar{P} \rangle \in \Delta$ with p = q and a = t(v); Pathfinder replies with some z-successor u_z of v and some state $r \in P_z$; and the game continues in the position (u_z, r) . Automaton wins a play in this game if he either manoevres Pathfinder into a position where the latter cannot

make a move, or if the play is infinite and the corresponding sequence $(q_n)_{n < \omega}$ of states from the Automaton positions satisfies the parity condition:

 $\liminf_{n<\omega}\Omega(q_n) \quad \text{is even.}$

We say that the automaton \mathcal{A} accepts the tree $t \in \mathbb{T}_{\xi}^{\times} \Sigma$ if Player Automaton has a winning strategy for \mathcal{G} when starting in the position $\langle \langle \rangle, q_{\circ} \rangle$ that consists of the root $\langle \rangle$ of t and the initial state q_{\circ} of \mathcal{A} . Finally, we call the automaton *non-deterministic* if, in every transition $\langle p, a, (P_z)_{z \in \zeta} \rangle \in \Delta$, all the sets P_z are singletons.

It is frequently useful to consider the restriction of an Automaton-Pathfinder game to some factor $[u, \bar{v})$ of the given input tree *t*. In such cases we can collect information about the possible strategies of Automaton on this restriction in a set of *profiles*. Such a profile is a pair $\langle p, \bar{U} \rangle$ consisting of a state $p \in Q$ and a tuple $\tilde{U} = (U_z)_{z \in \zeta}$ of sets $U_z \subseteq \operatorname{rng} \Omega \times Q$, where ζ is the sort of the factor $[u, \bar{v})$. We say that $\langle p, \bar{U} \rangle$ is a profile of the factor $[u, \bar{v})$ in *t* if there exists a strategy for Automaton such that, when starting the game in position $\langle u, p \rangle$, Pathfinder has, for each $z \in \zeta$ and every $\langle k, q \rangle \in U_z$, a strategy to reach the position $\langle v_z, q \rangle$ such that the least priority seen in between is equal to *k*.

Finally, given a tree $s \in \mathbb{T}^{\times}_{\zeta}\Sigma$ with an arbitrary sort ζ , we say that $\langle p, \overline{U} \rangle$ is a profile of *s* if there exists some tree $t \in \mathbb{T}^{\times}_{\zeta}\Sigma$ containing a factor $[u, \overline{v})$ such that $s = t[u, \overline{v})$ and $\langle p, \overline{U} \rangle$ is a profile of $[u, \overline{v})$ in *t*. (Note that this fact does not depend on the choice of *t*.)

By definition, every MSO-definable \mathbb{T}^{\times} -algebra \mathfrak{A} has a finite set of generators $C \subseteq A$ such that the preimages $\pi^{-1}(a) \cap \mathbb{T}^{\times}C$ are regular, that is, recognised by an automaton. For each given sort $\xi \in \Xi$, we can combine all these automata into a single automaton \mathcal{A} that has, for every element $a \in A_{\xi}$, some state q_a such that, when using q_a as the initial state, the automaton accepts all trees in $\mathbb{T}_{\xi}^{\times}C$ with product a.

The fact that this combined automaton \mathcal{A} does not work for all trees in $\mathbb{T}_{\xi}^{*}A$, but only for those in $\mathbb{T}_{\xi}^{*}C$, is frequently inconvenient. The problem is that the set A of possible labels is infinite. Given an automaton $\mathcal{A} = \langle Q, C + \xi, \Delta, q_o, \Omega \rangle$ over the alphabet $C + \xi$, we can extend it to the infinite alphabet $A + \xi$ as follows. Since C is a set of generators, there exists some function $\vartheta : A \to \mathbb{T}^{\times}C$ such that $\pi \circ \vartheta = \text{id}$. As new state of states we sue $D \times Q$, where $D := \operatorname{rng} \Omega$ is the set of priorities. The priority function is $\Omega^+(k, p) := k$ and the initial state $q_o^+\langle o, q_o \rangle$. Finally, the extended transition relation Δ_+ consists of all triples $\langle \langle l, p \rangle, a, (U_z)_{z \in \zeta} \rangle$ such that, in the original automaton, the tree $\vartheta(a)$ has the profile $\langle p, (U_z)_z \rangle$. Let $\mathcal{A}_+ := \langle D \times Q, A + \xi, \Delta_+, q_o^+, \Omega \rangle$ be the automaton with this transition relation. (Note that \mathcal{A}_+ is not a finite automaton anymore, although it still have only finitely many states.) By construction, it follows that \mathcal{A}_+ accepts a tree $t \in \mathbb{T}_{\xi}^{\times} A$ if, and only if, \mathcal{A} accepts the tree flat($\mathbb{T}^{\times} \vartheta(t)$). Since

$$(\pi \circ \operatorname{flat} \circ \mathbb{T}^{\times} \vartheta)(t) = (\pi \circ \mathbb{T}^{\times} \pi \circ \mathbb{T}^{\times} \vartheta)(t) = (\pi \circ \mathbb{T}^{\times} (\pi \circ \vartheta))(t) = \pi(t),$$

it follows that A_+ recognises the language $\pi^{-1}(a)$ (with no restriction on the labels allowed in the trees).

To prove the converse of Theorem 4.4, we start with a technical lemma which is based on an argument originally due to Bojańczyk and Klin (the first published version of which appears in [7]).

Lemma 4.5. Let Σ be a finite sorted set and let $R \subseteq \mathbb{T}\Sigma \times \mathbb{T}\Sigma$ be a binary relation between trees. For every tree $U \in \mathbb{T}_{\xi}R$, there exists a tree $U^{\circ} \in \mathbb{T}_{\xi}^{\operatorname{reg}}R$ such that, for each i < 2, the trees

$$\operatorname{flat}(\mathbb{T}p_i(U))$$
 and $\operatorname{flat}(\mathbb{T}p_i(U^\circ))$

have identical MSO-theories of quantifier-rank *m*, where $p_0, p_1 : \mathbb{T}\Sigma \times \mathbb{T}\Sigma \to \mathbb{T}\Sigma$ are the two projections.

Proof. To construct U° we use a variant of the Automaton-Pathfinder Game from above. Given two tree automata \mathcal{A} and \mathcal{B} , we define a game $\mathcal{G}_R(\mathcal{A}, \mathcal{B})$ where the first player wins if, and only if, there exist two trees $S, T \in \mathbb{TT}\Sigma$ with the same domain such that

- $\langle S(v), T(v) \rangle \in R$, for all vertices v,
- ⋆ A accepts flat(S),
- \mathcal{B} accepts flat(T).

The difference to the usual Automaton-Pathfinder Game is that we simulate two automata at the same time and that, instead of playing single letters, we play larger trees in each step. The game has two players *Automaton* and *Pathfinder*. Each round starts in a position of the form $\langle p, q \rangle$, where *p* is a state of \mathcal{A} and *q* one of \mathcal{B} . In the first round of the game, *p* and *q* are the initial states of the respective automata. Given such a position $\langle p, q \rangle$, Automaton chooses

- a pair of trees $(s, t) \in R_{\xi}$, for some $\xi \in \Xi$,
- a profile $\langle p, \bar{V} \rangle$ for A on *s* that starts in the state *p*, and
- a profile $\langle q, \bar{W} \rangle$ for \mathcal{B} on t that starts in the state q.

Pathfinder responds by selecting a variable $x \in \xi$ and pairs $\langle k, p' \rangle \in V_x$ and $\langle l, q' \rangle \in W_x$. The *outcome* of this round is the pair $\langle \langle p, k, p' \rangle, \langle q, l, q' \rangle \rangle$ and the game will continue in the position $\langle p', q' \rangle$.

If at some point in the game one of the players cannot make his choice, that player loses the game. Otherwise, the players produce an infinite sequence $\langle \delta_0, \varepsilon_0 \rangle$, $\langle \delta_1, \varepsilon_1 \rangle$, ... of outcomes. Let k_i be the priority in δ_i and l_i the priority in ε_i . Player Automaton wins the game if each of the sequences k_0, k_1, \ldots and l_0, l_1, \ldots satisfies the parity condition. Otherwise, Pathfinder wins.

Clearly, if there are two trees *S*, $T \in \mathbb{TT}\Sigma$ with the same domain such that

- $\langle S(v), T(v) \rangle \in R$, for all vertices v,
- *A* accepts flat(*S*), and
- \mathcal{B} accepts flat(T),

then Automaton has the following winning strategy in $\mathcal{G}_R(\mathcal{A}, \mathcal{B})$. He fixes two winning strategies ρ and σ for the games on, respectively, flat(*S*) and flat(*T*). During the game he descends through the trees *S* and *T*. When the game reaches a vertex ν , Automaton chooses the trees $S(\nu)$ and $T(\nu)$ and the profiles on the corresponding factors of flat(*S*) and flat(*T*) that are associated with ρ and σ , respectively.

Conversely, if Automaton has a winning strategy in the game, we can use it to construct

- two trees $S, T \in \mathbb{TT}\Sigma$ such that $(S(v), T(v)) \in R$, for all v, and
- winning strategies for Automaton in the games for flat(*S*) and flat(*T*), respectively.

Having defined the game $\mathcal{G}_R(\mathcal{A}, \mathcal{B})$, we prove the statement of the lemma as follows. Fix $U \in \mathbb{T}R$, let θ_i be the MSO-theory of quantifier-rank *m* of the tree

$\operatorname{flat}(\mathbb{T}p_i(U))$,

and let \mathcal{A}_i be an automaton recognising the class of all trees with theory θ_i . It follows that Player Automaton has a winning strategy in the game $\mathcal{G}_R(\mathcal{A}_0, \mathcal{A}_1)$. As the winning condition of this game is regular, we can apply the Büchi-Landweber Theorem [15], which tells us that Automaton even has a finite-memory winning strategy. Since the choice of S(v) and T(v) by Automaton in the game only depends on the current position $\langle p, q \rangle$ and on the contents of the memory, it follows that the resulting trees S and T are regular, i.e., they belong to $\mathbb{T}^{\operatorname{reg}}\mathbb{T}\Sigma$. The tree $U^o \in \mathbb{T}^{\operatorname{reg}}R$ with labels

$$U^{\circ}(v) \coloneqq \langle S(v), T(v) \rangle$$

has the desired properties.

Theorem 4.6. Every MSO-definable \mathbb{T}^{reg} -algebra can be expanded to a unique MSO-definable \mathbb{T} -algebra.

Proof. Uniqueness follows by Theorem 4.4 and Proposition 4.3. For existence, consider an MSO-definable \mathbb{T}^{reg} -algebra \mathfrak{A} and let $C \subseteq A$ be a finite set of generators. For each $a \in A$, we fix an MSO-formula φ_a defining the set $\pi^{-1}(a) \cap \mathbb{T}^{\text{reg}}C$. Intuitively, we obtain the desired expansion $\mathfrak{A}_+ = \langle A, \pi_+ \rangle$ by taking for π_+ the function on all trees defined by these formulae. To make this work we have to (I) find a way to use the formulae φ_a on trees containing labels from $A \setminus C$, (II) show that the resulting function is well-defined, (III) show that it extends π , and (IV) show that it satisfies the axioms of a \mathbb{T} -algebra.

(1) We define π_+ as follows. As *C* generates \mathfrak{A} , there exists a function $\sigma : A \to \mathbb{T}^{\operatorname{reg}}C$ such that $\pi(\sigma(a)) = a$, for all $a \in A$. We choose σ such that $\sigma(c) = \operatorname{sing}(c)$, for $c \in C$. Let $\hat{\sigma} := \operatorname{flat} \circ \mathbb{T}\sigma : \mathbb{T}A \to \mathbb{T}C$ be its extension to $\mathbb{T}A$. We set

$$\pi_+ \coloneqq \pi_\circ \circ \hat{\sigma},$$

where

$$\pi_{o}(t) \coloneqq a \quad : \text{iff} \quad t \vDash \varphi_{a}, \quad \text{for } t \in \mathbb{T}C.$$

(II) To see that π_+ is well-defined, we have to check that, for every tree $t \in \mathbb{T}_{\xi}C$, there is exactly one element $a \in A_{\xi}$ with $t \models \varphi_a$. For a contradiction, suppose otherwise. Then we can find a tree $t \in \mathbb{T}_{\xi}C$ such that

$$t \vDash \bigvee_{a \neq b} (\varphi_a \land \varphi_b) \lor \neg \bigwedge_{a \in A_{\mathcal{E}}} \varphi_a \, .$$

Since every non-empty MSO-definable tree language contains a regular tree, it follows that we can choose $t \in \mathbb{T}^{\text{reg}}C$. By choice of the formulae φ_a this means that $\pi(t)$ has either no value, or more than one. A contradiction.

(III) π_+ extends π since, for $t \in \mathbb{T}^{\text{reg}}A$, we have

$$\pi_+(t) = \pi_{\mathrm{o}}(\hat{\sigma}(t)) = \pi(\hat{\sigma}(t)) = \pi(t),$$

where the second step follows since $\hat{\sigma}(t) \in \mathbb{T}^{\text{reg}}C$, for $t \in \mathbb{T}^{\text{reg}}C$.

(IV) It remains to check the axioms of a \mathbb{T} -algebra. First, note that

$$\pi_{+} \circ \hat{\sigma} = \pi_{\circ} \circ \hat{\sigma} \circ \hat{\sigma}$$

$$= \pi_{\circ} \circ \text{flat} \circ \mathbb{T}\sigma \circ \text{flat} \circ \mathbb{T}\sigma$$

$$= \pi_{\circ} \circ \text{flat} \circ \text{flat} \circ \mathbb{T}(\mathbb{T}\sigma \circ \sigma) \qquad \text{[flat nat. trans.]}$$

$$= \pi_{\circ} \circ \text{flat} \circ \text{flat} \circ \mathbb{T}(\mathbb{T}\text{sing} \circ \sigma)$$

$$= \pi_{\circ} \circ \text{flat} \circ \mathbb{T}\text{sing} \circ \text{flat} \circ \mathbb{T}\sigma \qquad \text{[flat nat. trans.]}$$

$$= \pi_{\circ} \circ \text{flat} \circ \mathbb{T}\sigma \qquad \text{[flat nat. trans.]}$$

$$= \pi_{\circ} \circ \text{flat} \circ \mathbb{T}\sigma \qquad \text{[\mathbb{T} monad]}$$

$$= \pi_{\circ} \circ \hat{\sigma}$$

$$= \pi_{+} .$$

For the unit law, it therefore follows that

 $\pi_+ \circ \operatorname{sing} = \pi_+ \circ \hat{\sigma} \circ \operatorname{sing} = \pi_+ \circ \sigma = \pi \circ \sigma = \operatorname{id}$.

It remains to check the associative law. For a set of sorts $\Delta \subseteq \Xi$ and a sorted set *A*, we denote by $A|_{\Delta}$ the sorted set obtained from *A* by removing all elements whose sort does not belong to Δ . Below we will establish the following two claims.

(a) The associative law holds for all trees $t \in \mathbb{TT}^{reg}C$.

(b) The associative law holds for all trees $t \in \mathbb{T}(\mathbb{T}C)|_{\Delta}$ with finite Δ .

Then we can prove the general case as follows. Given $t \in \mathbb{TT}A$, we consider the tree $s := \operatorname{flat}(\mathbb{T}\hat{\sigma}(t))$. Each tree $r \in \mathbb{T}C$ can be written as $r = p(\bar{u}, \bar{x})$ for a finite tree p, variables \bar{x} , and infinite trees \bar{u} that do not contain variables. This implies that each tree $r \in \mathbb{T}C$ can be written as $r = \operatorname{flat}(R)$, for some $R \in \mathbb{T}^{\operatorname{fin}}(\mathbb{T}C)|_{\Delta}$, where $\Delta \subseteq \Xi$ is the finite set consisting of the sort \emptyset and the sorts of the elements of C. (Note that this argument does not work for $\mathbb{T}^{\times}C$.) Consequently, there exist a finite set $\Delta \subseteq \Xi$ and a tree $S \in \mathbb{TT}^{\operatorname{fin}}(\mathbb{T}C)|_{\Delta}$ with $s = \mathbb{T}\operatorname{flat}(S)$. By the two claims above, it follows that

$$(\pi_{+} \circ \operatorname{flat})(s) = (\pi_{+} \circ \operatorname{flat} \circ \operatorname{Tflat})(S)$$

$$= (\pi_{+} \circ \operatorname{flat} \circ \operatorname{flat})(S) \qquad [\mathbb{T} \text{ monad}]$$

$$= (\pi_{+} \circ \mathbb{T}\pi_{+} \circ \operatorname{flat})(S) \qquad [by (a)]$$

$$= (\pi_{+} \circ \mathbb{T}\pi_{+} \circ \mathbb{T}\mathbb{T}\pi_{+})(S) \qquad [flat nat. trans.]$$

$$= (\pi_{+} \circ \mathbb{T}\pi_{+} \circ \mathbb{T}\mathbb{T}\pi_{+})(S) \qquad [by (a)]$$

$$= (\pi_{+} \circ \mathbb{T}\pi_{+} \circ \mathbb{T}\operatorname{flat})(S) \qquad [by (b)]$$

$$= (\pi_{+} \circ \mathbb{T}\pi_{+})(s).$$

Hence, it remains to prove the two claims.

(a) Fix $t \in \mathbb{TT}^{\text{reg}}A$. We consider the relation $R \subseteq \mathbb{T}C \times \mathbb{T}C$ given by

$$\langle s, t \rangle \in R$$
 : iff $s, t \in \mathbb{T}^{\operatorname{reg}}C$ and $\pi_+(s) = \pi_+(t)$,

and the tree $U \in \mathbb{T}R$ with labels

$$U(v) \coloneqq \left\langle \hat{\sigma}(t(v)), \, \sigma(\pi_+(t(v))) \right\rangle, \quad \text{for } v \in \text{dom}(t) \, .$$

Let $p_0, p_1 : \mathbb{T}C \times \mathbb{T}C \to \mathbb{T}C$ be the two projections. Then

$$a := (\pi_{+} \circ \operatorname{flat} \circ \mathbb{T}p_{\circ})(U)$$

= $(\pi_{+} \circ \operatorname{flat} \circ \mathbb{T}\hat{\sigma})(t)$
= $(\pi_{+} \circ \operatorname{flat} \circ \mathbb{T}(\operatorname{flat} \circ \mathbb{T}\sigma))(t)$
= $(\pi_{+} \circ \operatorname{flat} \circ \operatorname{flat} \circ \mathbb{T}\mathbb{T}\sigma)(t)$ [T monad]
= $(\pi_{+} \circ \operatorname{flat} \circ \mathbb{T}\sigma \circ \operatorname{flat})(t)$ [flat nat. trans.]
= $(\pi_{+} \circ \hat{\sigma} \circ \operatorname{flat})(t)$
= $\pi_{+}(\operatorname{flat}(t)),$
 $b := (\pi_{+} \circ \operatorname{flat} \circ \mathbb{T}p_{1})(U)$
= $(\pi_{+} \circ \operatorname{flat} \circ \mathbb{T}(\sigma \circ \pi_{+}))(t)$
= $(\pi_{+} \circ \hat{\sigma} \circ \mathbb{T}\pi_{+})(t)$
= $\pi_{+}(\mathbb{T}\pi_{+})(t).$

Hence, we have to prove that a = b. For a contradiction, suppose otherwise. By Lemma 4.5, there exists a regular tree $U^{\circ} \in \mathbb{T}^{\operatorname{reg}} R$ with

$$\pi_{o}(\operatorname{flat}(\mathbb{T}p_{o}(U^{\circ}))) = \pi_{o}(\operatorname{flat}(\mathbb{T}p_{o}(U))) = a,$$

$$\pi_{o}(\operatorname{flat}(\mathbb{T}p_{1}(U^{\circ}))) = \pi_{o}(\operatorname{flat}(\mathbb{T}p_{1}(U))) = b.$$

Setting $s_i := \mathbb{T}p_i(U^\circ) \in \mathbb{T}^{\operatorname{reg}}\mathbb{T}^{\operatorname{reg}}C$, it follows by associativity of π and the fact that $(s_o(\nu), s_1(\nu)) \in R$ that

$$a = \pi(\text{flat}(s_0)) = \pi(\mathbb{T}\pi(s_0)) = \pi(\mathbb{T}\pi(s_1)) = \pi(\text{flat}(s_1)) = b.$$

A contradiction.

(b) Fix $t \in \mathbb{T}(\mathbb{T}C)|_{\Delta}$ where $\Delta \subseteq \Xi$ is finite. For $a \in A$, let ψ_a be a formula stating that

"For every factorisation of the given tree, there exists a labelling of the factors by elements of *A* such that

- every factor has a sort in Δ ,
- each factor with label $b \in A|_{\Delta}$ satisfies the formula φ_b , and
- the tree consisting of the guessed labels satisfies φ_a ."

Since π is associative, we have

$$s \models \psi_a \rightarrow \varphi_a$$
, for every $s \in \mathbb{T}^{\operatorname{reg}} C$.

It follows that the same is true for every $s \in \mathbb{T}C$. Consequently, we have

 $(\pi_+ \circ \operatorname{flat})(t) = (\pi_+ \circ \mathbb{T}\pi_+)(t), \qquad \Box$

Note that our uniqueness result in Proposition 4.3 only concerns expansions in the given class C. It is possible that there exist additional expansions outside of C.

Example. In [14] Bojańczyk and Klin have presented an example of a finitary \mathbb{T} -algebra that is not MSO-definable. This algebra can be used to find an MSO-definable \mathbb{T}^{reg} -algebra with several \mathbb{T} -expansions, one of them MSO-definable. (By the preceding theorem, there can only be one of the latter.) The construction of this \mathbb{T}^{reg} -algebra $\mathfrak{A}^{\circ} = \langle A, \pi \rangle$ and two of its \mathbb{T} -expansions $\mathfrak{A}^{\text{reg}} = \langle A, \pi_{\text{reg}} \rangle$ (MSO-definable) and $\mathfrak{A}^{\text{non}} = \langle A, \pi_{\text{non}} \rangle$ (not MSO-definable) is as follows.

For \mathfrak{A}^{non} , we take (a simplified version) of the algebra from [14]. Let $\Sigma := \{a, b\}$ where both elements have arity 2 and set $\Delta_{\xi} := \mathbb{T}_{\xi}^{fin} \Sigma$. As Σ contains only binary elements, every leaf of a tree $t \in \Delta_{\xi}$ must be labelled by a variable. Hence, *t* has at most $|\xi|$ leaves and, therefore, at most $|\xi| - 1$ internal vertices. This implies that Δ_{ξ} is a finite set.

We call a tree $t \in \mathbb{T}\Sigma$ antiregular if no two subtrees of t are isomorphic. It is *densely antiregular* if every subtree of t has an antiregular subtree.

The domains of all three algebras are

$$A_{\xi} \coloneqq \Delta_{\xi} \cup \{\bot, *\}, \quad \text{for } \xi \in \Xi,$$

which we order such that \perp is the least element and all other elements are incomparable. For $t \in \mathbb{T}_{\xi}A$, we define the product $\pi_{non}(t)$ of \mathfrak{A}^{non} by the following case distinction.

- $\pi_{\text{non}}(t) = \bot$ if *t* contains the label \bot .
- $\pi_{\text{non}}(t) = \text{flat}(t)$ if t is finite and $t \in \mathbb{T}_{\xi} \Delta$.

- $\pi_{\text{non}}(t) = *$ if *t* is finite and *t* contains the label *.
- $\pi_{non}(t) = *$ if t is infinite and every infinite subtree of t contains the label *.
- $\pi_{non}(t) = *$ if t is infinite and, for every infinite subtree s of t with $s \in \mathbb{T}\Delta$, the tree flat(s) has a subtree without variables that is not antiregular.
- $\pi_{non}(t) = \bot$ if t is infinite and there exists an infinite subtree s of t such that $s \in \mathbb{T}\Delta$ and every subtree of flat(s) without variables is densely antiregular.

The product $\pi_{reg}(t)$ of \mathfrak{A}^{reg} is defined as follows.

- $\pi_{reg}(t) = \bot$ if *t* contains the label \bot .
- $\pi_{\text{reg}}(t) = \text{flat}(t)$ if t is finite and $t \in \mathbb{T}_{\xi} \Delta$.
- $\pi_{reg}(t) = *$ if t is finite and t contains the label *.
- $\pi_{reg}(t) = *$ if t is infinite.

Note that this product is MSO-definable and that the restrictions of π_{reg} and π_{non} to $\mathbb{T}^{reg}A$ coincide. Thus \mathfrak{A}^{reg} and \mathfrak{A}^{non} are \mathbb{T} -expansions of the same \mathbb{T}^{reg} -algebra.

.

To conclude this section, let me mention one of the main open problems concerning the relation between regular trees and arbitrary ones.

Open Question. Does there exist a system of equations modulo which every tree is equivalent to a regular one? If so, does it have an explicit description?

Having such an equational characterisation would be invaluable for applications, where proofs frequently require a reduction to regular trees. For instance, there exists an equational characterisation of bisimulation-invariant languages of *regular* trees [12], but so far nobody was able to generalise it to languages of arbitrary trees.

5 EVALUATIONS

The notion of denseness seems to be only useful if we already know that the algebra in question has an expansion. To actually prove existence we need different techniques. We start with the following simple idea. Given a submonad $\mathbb{T}^{\circ} \subseteq \mathbb{T}$ and a \mathbb{T}° -algebra \mathfrak{A} , we try to compute the product of a tree $t \in \mathbb{T}A$ inductively bottom-up using the given \mathbb{T}° -product. That is, we factorise t into pieces that belong to \mathbb{T}° , evaluate them, and then recursively evaluate the remaining tree. If we can show that,

- after a finite number of such steps, the tree *t* is reduced to a single vertex and
- that the final result does not depend on the choice of factorisation used in each step,

it follows that we can uniquely evaluate every tree in $\mathbb{T}A$ using only the \mathbb{T}° -product. In particular, the product of \mathfrak{A} can be uniquely extended to the set of all trees.

A well-known use of this technique is given by Simon's Factorisation Tree Theorem (see, e.g., [18]). Such a factorisation tree is just a hierarchical decomposition of a given semigroup-product using binary products and products of idempotents only. A second example of this approach was used in [16] to prove, in our terminology, that a certain inclusion between monads of countable linear orders is dense. The aim of the current section is to make this idea of using an inductive approach work for trees. Below we will use recursive factorisations to settle the expansion problem for thin trees and we will derive partial results for general ones.

The definition below is a bit more general than the above intuitive description. We will need this added generality for the more powerful decompositions used further below. Suppose that we are given a \mathbb{T}° -algebra, which we want to expand to a \mathbb{T}^{1} -algebra, for some $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{1} \subseteq \mathbb{T}$, and suppose that we have already found some sorted set $S \supseteq \mathbb{T}^{\circ}A$ such that we can extend the product $\pi : \mathbb{T}^{\circ}A \to A$ to $\rho : S \to A$. To extend ρ further to a function $\mathbb{T}^{1}A \to A$, consider a tree $t \in \mathbb{T}^{1}A$. We choose a factorisation T of t where we have already inductively assigned some value val(T(v)) to each factor. If the reduced tree \mathbb{T}^{1} val(T) belongs to S, then we can set val $(t) := \rho(\mathbb{T}^{1}$ val(T)).

In the following formal definition, $\mathbb{E}_{\alpha}(\rho, \mathbb{T}^1)$ is the set of all recursive factorisations with α levels of recursion, term_{α} : $\mathbb{E}_{\alpha}(\rho, \mathbb{T}^1) \rightarrow \mathbb{T}^1 A$ maps each such factorisation to the tree it factorises, and val_{α} : $\mathbb{E}_{\alpha}(\rho, \mathbb{T}^1) \rightarrow A$ maps each factorisation to the value obtained by recursively evaluating every factor.

Definition 5.1. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{1} \subseteq \mathbb{T}$ be submonads, \mathfrak{A} a \mathbb{T}° -algebra, and $\rho : S \to A$ a function with domain $S \supseteq \mathbb{T}^{\circ}A$.

(a) For each ordinal α , we inductively define a sorted set $\mathbb{E}_{\alpha}(\rho, \mathbb{T}^1)$ of ρ -evaluations and two functions

$$\operatorname{term}_{\alpha} : \mathbb{E}_{\alpha}(\rho, \mathbb{T}^{1}) \to \mathbb{T}^{1}A \quad \operatorname{and} \quad \operatorname{val}_{\alpha} : \mathbb{E}_{\alpha}(\rho, \mathbb{T}^{1}) \to A$$

$$\begin{split} & \mathbb{E}_{\alpha}(\rho, \mathbb{T}^{1}) \coloneqq A, \\ & \mathbb{E}_{\alpha+1}(\rho, \mathbb{T}^{1}) \coloneqq \mathbb{E}_{\alpha}(\rho, \mathbb{T}^{1}) + \left\{ \gamma \in \mathbb{T}^{1}\mathbb{E}_{\alpha}(\rho, \mathbb{T}^{1}) \mid \mathbb{T}^{1}\mathrm{val}_{\alpha}(\gamma) \in \mathrm{dom}(\rho) \right\}, \\ & \mathbb{E}_{\delta}(\rho, \mathbb{T}^{1}) \coloneqq \bigcup_{\alpha < \delta} \mathbb{E}_{\alpha}(\rho, \mathbb{T}^{1}), \qquad \text{for limit ordinals } \delta, \end{split}$$

and

$$\begin{split} & \operatorname{term}_{\circ} \coloneqq \operatorname{sing}, & \operatorname{val}_{\circ} \coloneqq \operatorname{id}, \\ & \operatorname{term}_{\alpha+1} \coloneqq \operatorname{term}_{\alpha} + \operatorname{flat} \circ \mathbb{T}^{1} \operatorname{term}_{\alpha}, & \operatorname{val}_{\alpha+1} \coloneqq \operatorname{val}_{\alpha} + \rho \circ \mathbb{T}^{1} \operatorname{val}_{\alpha}, \\ & \operatorname{term}_{\delta} \coloneqq \bigcup_{\alpha < \delta} \operatorname{term}_{\alpha}. & \operatorname{val}_{\delta} \coloneqq \bigcup_{\alpha < \delta} \operatorname{val}_{\alpha}. \end{split}$$

Finally, we set

$$\mathbb{E}(\rho, \mathbb{T}^1) := \bigcup_{\alpha} \mathbb{E}_{\alpha}(\rho, \mathbb{T}^1), \quad \text{term} := \bigcup_{\alpha} \text{term}_{\alpha}, \quad \text{and} \quad \text{val} := \bigcup_{\alpha} \text{val}_{\alpha}.$$

(b) We call term(γ) the *underlying term* of $\gamma \in \mathbb{E}(\rho, \mathbb{T}^1)$ and val(γ) its *value*. If $t = \text{term}(\gamma)$, we say that γ is a ρ -evaluation of t.

(c) We say that the algebra \mathfrak{A} has ρ -evaluations if term : $\mathbb{E}(\rho, \mathbb{T}^1) \to \mathbb{T}^1 A$ is surjective, and we say that it has essentially unique ρ -evaluations if furthermore

 $\operatorname{term}(\gamma) = \operatorname{term}(\gamma')$ implies $\operatorname{val}(\gamma) = \operatorname{val}(\gamma')$.

(d) In the special case where $\rho = \pi$, we also write $\mathbb{E}(\mathfrak{A}, \mathbb{T}^1) := \mathbb{E}(\pi, \mathbb{T}^1)$ and we call the elements of $\mathbb{E}(\mathfrak{A}, \mathbb{T}^1)$ simple \mathbb{T}° -evaluations.

Example. Let \mathfrak{A} be a \mathbb{T}^{reg} -algebra, $a \in A_{\{x,y\}}$, $b \in A_{\{x\}}$, $c \in A_{\emptyset}$ elements, and t the tree consisting of an infinite branch labelled a attached to which are trees of the form $s_n := b^n(c)$, for every $n < \omega$ (see Figure 3). We construct a simple \mathbb{T}^{reg} -evaluation of t as follows. In the first step, we evaluate each of the subtrees $s_n \in \mathbb{T}^{\text{reg}}A$. The resulting tree has the form

$$t' \coloneqq a(d_0, a(d_1, \dots)), \text{ where } d_n \coloneqq \pi(s_n),$$

i.e., t' consists of an infinite branch labelled a to each vertex of which we have attached a leaf with label d_n . Since $A_{\{z\}}$ forms a finite semigroup and $d_{n+1} =$

by



Figure 3: The trees t and t'.

 $b(d_n)$, the sequence d_0, d_1, \ldots is ultimately periodic. Fixing indices k < l with $d_k = d_l$, we can write $t' = uv^{\omega}$ where

$$u := a(d_0, a(d_1, \dots a(d_{k-1}, x))),$$

$$v := a(d_k, a(d_{k+1}, \dots a(d_{l-1}, x))).$$

Consequently, t' is regular and we can take is as the second and final level of our evaluation.

As another example, let us show how to encode an evaluation by a function $\tau : \text{dom}(t) \to [n]$ similar to a Ramseyan split. By definition, an evaluation of *t* is nothing but a nested factorisation $g \in \mathbb{T}^{\times} \mathbb{T}^{\times} \dots \mathbb{T}^{\times} A$. We can encode such a factorisation by assigning to every vertex $w \in \text{dom}(t)$ the depth at which it appears in this nesting (see Figure 4 for an example).

Lemma 5.2. Let $o < n < \omega$ and let $t \in \mathbb{T}^{\times}A$. There exists a bijection between simple \mathbb{T}^{\times} -evaluations $\gamma \in \mathbb{E}_n(\mathfrak{A}, \mathbb{T}^{\times})$ of t and functions $\tau : \operatorname{dom}(t) \to [n]$ with $\tau(\langle \rangle) = n - 1$.

Proof. Given $\gamma \in \mathbb{E}_n(\mathfrak{A}, \mathbb{T}^{\times})$ we construct τ by induction on n. If n = 1, we set $\tau(w) := 0$, for all w. Otherwise, let $\tau_v : \operatorname{dom}(\operatorname{term}_{n-1}(\gamma(v))) \to [n-1]$, for $w \in \operatorname{dom}(\gamma)$, be the functions obtained from the inductive hypothesis for $\gamma(v) \in \mathbb{E}_{n-1}(\mathfrak{A}, \mathbb{T}^{\times})$. Note that $t = \operatorname{term}_n(\gamma) = \operatorname{flat}(\mathbb{T}^{\times}\operatorname{term}_{n-1}(\gamma))$. Hence, every



Figure 4: An evaluation (on the right) encoded by a function (on the left). For simplicity, we have omitted the labels from A and drawn just the τ -labelling.

vertex $w \in \text{dom}(t)$ that is not labelled by a variable corresponds to a pair $\langle v, u \rangle$ of vertices with $v \in \text{dom}(\gamma)$ and $u \in \text{dom}(\text{term}_{n-1}(\gamma(v)))$. We set

$$\tau(w) \coloneqq \begin{cases} n-1 & \text{if } u \text{ is the root of } \operatorname{term}_{n-1}(\gamma(v)), \\ \tau_{\nu}(u) & \text{otherwise.} \end{cases}$$

Conversely, let τ : dom $(t) \rightarrow [n]$ be a function with $\tau(\langle \rangle) = n - 1$. If n = 1, we set $\gamma := t$. Otherwise, let *T* be the factorisation of *t* induced by the set

$$H \coloneqq \{ w \in \operatorname{dom}(t) \mid \tau(w) = n - 1 \}.$$

That is *H* is the set of vertices corresponding to the roots of the factors $T(\nu)$. For each $\nu \in \text{dom}(T)$, we can use the inductive hypothesis to find a evaluation $\gamma_{\nu} \in \mathbb{E}_{n-1}(\mathfrak{A}, \mathbb{T}^{\times})$ of $T(\nu)$. We set

$$y(v) \coloneqq y_v$$
.

Let us next explain how to use ρ -evaluations to construct \mathbb{T}^1 -expansions. The proof makes use of the following glueing operation for evaluations.

Lemma 5.3. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{1} \subseteq \mathbb{T}^{\times}$, $\gamma \in \mathbb{T}^{1}\mathbb{E}(\rho, \mathbb{T}^{1})$, and let β be a ρ -evaluation of the tree $t := \mathbb{T}^{1}$ val (γ) . There exists a ρ -evaluation $\beta | \gamma$ of the tree $(\text{flat} \circ \mathbb{T}^{1}\text{term})(\gamma)$ such that $\text{val}(\beta|\gamma) = \text{val}(\beta)$.

Proof. We define $\beta | \gamma$ by induction on the ordinal α with $\beta \in \mathbb{E}_{\alpha}(\rho, \mathbb{T}^{1})$. If $\alpha = 0$, then $\beta = a \in A$ and $\gamma = \operatorname{sing}(\gamma_{\circ})$, for some $\gamma_{\circ} \in \mathbb{E}(\rho, \mathbb{T}^{1})$ with $\operatorname{val}(\gamma_{\circ}) = a$. Setting $\beta | \gamma := \gamma_{\circ}$ we obtain $\operatorname{val}(\beta | \gamma) = \operatorname{val}(\gamma_{\circ}) = a = \operatorname{val}(\beta)$.

Suppose that $\beta \in \mathbb{E}_{\alpha+1}(\rho, \mathbb{T}^1) \setminus \mathbb{E}_{\alpha}(\rho, \mathbb{T}^1)$. Then $\beta \in \mathbb{T}^1\mathbb{E}_{\alpha}(\rho, \mathbb{T}^1)$. Let $\gamma' \in \mathbb{T}^1\mathbb{T}^1\mathbb{E}(\rho, \mathbb{T}^1)$ be the tree with the same domain as β where $\gamma'(\nu)$ is the restriction of γ to dom(term($\beta(\nu)$)) \subseteq dom(γ). We choose for $\beta|\gamma$ the tree with the same domain as β that is given by

$$(\beta|\gamma)(\nu) \coloneqq \beta(\nu)|\gamma'(\nu), \text{ for } \nu \in \operatorname{dom}(s).$$

By inductive hypothesis, it follows that

$$\operatorname{term}((\beta|\gamma)(\nu)) = \operatorname{term}(\beta(\nu)|\gamma'(\nu)) = (\operatorname{flat} \circ \mathbb{T}^{1}\operatorname{term})(\gamma'(\nu)),$$
$$\operatorname{val}((\beta|\gamma)(\nu)) = \operatorname{val}(\beta(\nu)|\gamma'(\nu)) = \operatorname{val}(\beta(\nu)).$$

Consequently, we have

$$\operatorname{term}(\beta|\gamma) = (\operatorname{flat} \circ \mathbb{T}^{1}\operatorname{term})(\beta|\gamma)$$

= (flat \circ \mathbf{T}^{1}(flat \circ \mathbf{T}^{1}\operatorname{term}))(\gamma')
= (flat \circ flat \circ \mathbf{T}^{1}\operatorname{term})(\gamma') [\mathbf{T}^{1} monad]
= (flat \circ \mathbf{T}^{1}\operatorname{term} \circ flat)(\gamma') [flat nat. trans.]
= (flat \circ \mathbf{T}^{1}\operatorname{term})(\gamma),
$$\operatorname{val}(\beta|\gamma) = (\rho \circ \mathbb{T}^{1}\operatorname{val})(\beta|\gamma)$$

= (\rho \circ \mathbf{T}^{1}\operatorname{val})(\beta) = \operatorname{val}(\beta).

Furthermore, note that $\beta | \gamma$ really is an evaluation since

$$\mathbb{T}^{1}$$
val $(\beta|\gamma) = \mathbb{T}^{1}$ val $(\beta) \in dom(\rho)$.

Finally, if $\beta \in \mathbb{E}_{\delta}(\rho, \mathbb{T}^1)$, for some limit ordinal δ , then there is some $\alpha < \delta$ with $\beta \in \mathbb{E}_{\alpha}(\rho, \mathbb{T}^1)$. Hence, the claim follows by inductive hypothesis.

Theorem 5.4. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{1} \subseteq \mathbb{T}$ be submonads. Every \mathbb{T}° -algebra \mathfrak{A} with essentially unique ρ -evaluations has a \mathbb{T}^{1} -expansion $\langle A, \pi_{+} \rangle$ such that

 $\pi_+ \circ \text{term} = \text{val}.$

Proof. For $t \in \mathbb{T}^1 A$, we define

 $\pi_+(t) \coloneqq \operatorname{val}(\gamma)$, for some $\gamma \in \operatorname{term}^{-1}(t)$.

As ρ -evaluations are essentially unique, it does not matter which evaluation γ we choose. Hence, this function is well-defined. We claim that $\mathfrak{A}_+ := \langle A, \pi_+ \rangle$ is the desired \mathbb{T}^1 -expansion of \mathfrak{A} .

The equation $\pi_+ \circ \text{term} = \text{val follows immediately from the definition of } \pi_+$. Since every tree $t \in \mathbb{T}^\circ A = \mathbb{E}_1(\mathfrak{A}, \mathbb{T}^1)$ is its own evaluation, we further have

$$\pi_+(t) = \operatorname{val}_1(t) = \rho(t) = \pi(t), \quad \text{for } t \in \mathbb{T}^\circ A.$$

Consequently, π_+ is an extension of π and it remains to check the axioms of a \mathbb{T}^1 -algebras.

For the unit law, we have $\pi_+(\operatorname{sing}(a)) = \pi(\operatorname{sing}(a)) = a \operatorname{since} \operatorname{sing}(a) \in \mathbb{T}^{\circ}A$. For associativity, let $t \in \mathbb{T}^1\mathbb{T}^1A$. Then there exists a tree of ρ -evaluations $\gamma \in \mathbb{T}^1\mathbb{E}(\mathfrak{A}, \mathbb{T}^1)$ such that $t = \mathbb{T}^1\operatorname{term}(\gamma)$. Furthermore, we can fix an evaluation $\beta \in \mathbb{E}(\mathfrak{A}, \mathbb{T}^1)$ of the tree $\mathbb{T}^1\operatorname{val}(\gamma)$. Let $\beta|\gamma$ be the ρ -evaluation from Lemma 5.3. Then

$$(\pi_{+} \circ \operatorname{flat})(t) = (\pi_{+} \circ \operatorname{flat} \circ \mathbb{T}^{1}\operatorname{term}_{n})(\gamma)$$

= $(\pi_{+} \circ \operatorname{term})(\beta|\gamma)$
= $\operatorname{val}(\beta|\gamma)$
= $\operatorname{val}(\beta)$
= $(\pi_{+} \circ \operatorname{term})(\beta)$
= $(\pi_{+} \circ \mathbb{T}^{1}\operatorname{val}_{n})(\gamma)$
= $(\pi_{+} \circ \mathbb{T}^{1}\pi_{+} \circ \mathbb{T}^{1}\operatorname{term}_{n})(\gamma) = (\pi_{+} \circ \mathbb{T}^{1}\pi_{+})(t)$.

The theorem tells us how to use evaluations to construct expansions. Let us see next where the limits of this method are.

Proposition 5.5. Let $\mathfrak{A} = \langle A, \pi \rangle$ be a \mathbb{T}° -algebra with a \mathbb{T}^{1} -expansion $\mathfrak{A}_{+} = \langle A, \pi_{+} \rangle$, and suppose that $\rho : S \to A$ is the restriction of π_{+} to some set $S \supseteq \mathbb{T}^{\circ}A$. Then

$$val(\gamma) = (\pi_+ \circ term)(\gamma)$$
, for every ρ -evaluation γ .

Proof. We prove that $val_{\alpha} = \pi_+ \circ term_{\alpha}$ by induction on α . For $\alpha = 0$, we have

$$\operatorname{val}_{o}(\gamma) = \gamma = \pi_{+}(\operatorname{sing}(\gamma)) = \pi_{+}(\operatorname{term}_{o}(\gamma)), \text{ for } \gamma \in \mathbb{E}_{o}(\rho, \mathbb{T}^{1}) = A$$

For the successor step, suppose that the equation holds for α and consider $\gamma \in \mathbb{E}_{\alpha+1}(\rho, \mathbb{T}^1)$. Then

$$\operatorname{val}_{\alpha+1}(\gamma) = \rho(\mathbb{T}^{1}\operatorname{val}_{\alpha}(\gamma))$$

= $\rho(\mathbb{T}^{1}(\pi_{+} \circ \operatorname{term}_{\alpha})(\gamma))$
= $\pi_{+}(\mathbb{T}^{1}\pi_{+}(\mathbb{T}^{1}\operatorname{term}_{\alpha}(\gamma)))$
= $\pi_{+}(\operatorname{flat}(\mathbb{T}^{1}\operatorname{term}_{\alpha}(\gamma)))$
= $\pi_{+}(\operatorname{term}_{\alpha+1}(\gamma)).$

Finally, for a limit ordinal α , the claim follows immediately from the inductive hypothesis.

Corollary 5.6. Let \mathfrak{A} be a \mathbb{T}° -algebra.

- (a) If \mathfrak{A} has several different \mathbb{T}^1 -expansions, there exist trees $t \in \mathbb{T}^1 A$ without a simple \mathbb{T}° -evaluation.
- (b) If \mathfrak{A} has simple \mathbb{T}° -evaluations that are not essentially unique, it has no \mathbb{T}^{1} -expansion.

Example. Before using these results to study thin trees, let us quickly recall the results of [16] about countable chains. We denote by $\mathbb{C}A$ the (unsorted) set of all countable *A*-labelled linear orders and by $\mathbb{C}^{\text{reg}}A \subseteq \mathbb{C}A$ the subset of all *regular* linear orders. By definition, a linear order is regular if it can be denoted by a finite term using the following operations: (1) constants for singletons, (11) binary ordered sums, (11) multiplication by ω and ω^{op} (ω with the reverse ordering), and (1V) dense shuffles. In [16] it is shown that every finite \mathbb{C}^{reg} -algebra has essentially unique \mathbb{C}^{reg} -evaluations. This fact can be used to prove the following results (for the proofs, see [16, 10, 8]).

- ◆ $\mathbb{C}^{reg} \subseteq \mathbb{C}$ is dense over the class of all finite \mathbb{C} -algebras.
- Every finite \mathbb{C}^{reg} -algebra has a unique \mathbb{C} -expansion.
- Every finite C-algebra is MSO-definable.
- A language K ⊆ CΣ of countable chains is MSO-definable if, and only if, it is recognised by some finite C-algebra.
- Every MSO-definable language $K \subseteq \mathbb{C}\Sigma$ has a finite syntactic algebra.

5.1 THIN TREES

As an application of simple evaluations we consider thin trees, where we can use the Theorem of Ramsey and other tools from semigroup theory. Note that, with every \mathbb{T}^{fin} -algebra \mathfrak{A} , we can associate the semigroup with universe $A_{\{z\}}$ (for some fixed variable *z* whose choice does not matter) and where the product is defined by

$$a \cdot b \coloneqq a(b(z))$$
, for $a, b \in A_{\{z\}}$.

If \mathfrak{A} is a $\mathbb{T}^{\text{wilke}}$ -algebra, this semigroup can be expanded to a Wilke algebra $\langle A_{\{z\}}, A_{\varnothing} \rangle$ by setting

$$a \cdot c \coloneqq a(c)$$
 and $a^{\omega} \coloneqq \pi(t_a)$, for $a \in A_{\{z\}}$ and $c \in A_{\emptyset}$,

where t_a is an infinite path each vertex of which is labelled by a. Finally, if \mathfrak{A} is a \mathbb{T}^{thin} -algebra, we obtain an ω -semigroup with infinite product

$$\pi(a_0, a_1, \dots) \coloneqq a_0(a_1(\dots)), \quad \text{for } a_i \in A_{\{z\}}.$$

We start by generalising the fact that every finite Wilke algebra has a unique expansion to an ω -semigroup to the monads $\mathbb{T}^{\text{wilke}} \subseteq \mathbb{T}^{\text{thin}}$.

Proposition 5.7. Every finitary $\mathbb{T}^{\text{wilke}}$ -algebra has essentially unique simple $\mathbb{T}^{\text{wilke}}$ -evaluations for trees in \mathbb{T}^{thin} .

Proof. Let \mathfrak{A} be a finitary $\mathbb{T}^{\text{wilke}}$ -algebra and $t \in \mathbb{T}^{\text{thin}}A$ a thin tree. We construct the desired simple evaluation of t by induction on the Cantor-Bendixson rank α of t. (Recall that the Cantor-Bendixson rank of a tree t is the least ordinal α such that $\partial^{\alpha+1}(t)$ is empty, where $\partial(t)$ denotes the tree obtained from t by removing every subtree with only finitely many infinite branches, and ∂^{α} is the α -th iteration of ∂ . One can show that such an ordinal α exists if, and only if, the given tree t is thin.)

By inductive hypothesis, every subtree $t|_{\nu}$ of rank less than α has a simple evaluation $\gamma_{\nu} \in \mathbb{E}(\mathfrak{A}, \mathbb{T}^{\text{thin}})$. Let *s* be tree obtained from *t* by replacing every such subtree $t|_{\nu}$ by val (γ_{ν}) . It is sufficient to find a simple evaluation of *s*. Then we can use the glueing operation from Lemma 5.3 to construct the desired evaluation of *t*.

By construction, *s* has only finitely many infinite branches. We distinguish three cases.

(1) If *s* is finite, it is its own evaluation.

(II) Next, suppose that *s* has a single infinite branch. By the Theorem of Ramsey, we can find a factorisation $s = p_0 p_1 p_2 \dots$ such that $\pi(p_i) = \pi(p_j)$, for all *i*, *j* > 0. As each factor p_i is finite, we obtain simple evaluations β_i of p_i by (I). The path

 $\rho := \pi(p_0), \pi(p_1), \pi(p_2), \dots$ is of the form ae^{ω} for $a := \pi(p_0)$ and $e := \pi(p_1)$. In particular, it is regular. Let β_* be the path $\beta_0, \beta_1, \beta_2, \dots$ Then

 $\mathbb{T}^{\text{thin}} \text{val}(\beta_*) = aeee \cdots \in \mathbb{T}^{\text{reg}} A \text{ and } \text{term}(\beta_*) = p_0 p_1 p_2 \cdots = s.$

Hence, β_* is the desired simple evaluation of *s*.

(III) Finally, suppose that *s* has at least two infinite branches. Then we can factorise *s* into a finite prefix and finitely many trees with a single infinite branch. By (I) and (II), each of these factors has a simple evaluation. Let β be the finite tree consisting of these evaluations. Then β is an simple evaluation of *s*.

Corollary 5.8. $\mathbb{T}^{\text{wilke}} \subseteq \mathbb{T}^{\text{thin}}$ *is dense over the class of all finitary* \mathbb{T}^{thin} *-algebras.*

Corollary 5.9. *Every finitary* $\mathbb{T}^{\text{wilke}}$ *-algebra has a unique* \mathbb{T}^{thin} *-expansion.*

It follows that the step from a $\mathbb{T}^{\text{wilke}}$ -algebra to a \mathbb{T}^{thin} -expansion is fairly well understood. The inclusion $\mathbb{T}^{\text{fin}} \subseteq \mathbb{T}^{\text{wilke}}$ is slightly more complicated since expansions are no longer unique.

Proposition 5.10. Let $\mathfrak{A} = \langle A, \pi \rangle$ be a finitary \mathbb{T}^{fin} -algebra. There exists a bijection between all $\mathbb{T}^{\text{wilke}}$ -expansions of \mathfrak{A} and all functions $-^{\omega} : A_{\{z\}} \to A_{\emptyset}$ satisfying the equations

$$(ab)^{\omega} = a(ba)^{\omega}$$
 and $(a^n)^{\omega} = a^{\omega}$, for all $a, b \in A_{\{z\}}$.

Proof. Clearly, every $\mathbb{T}^{\text{wilke}}$ -expansion $\mathfrak{A}^+ = \langle A, \pi_+ \rangle$ of \mathfrak{A} induces an ω -power operation by

 $a^{\omega} \coloneqq \pi_+(aaa...)$ (an infinite path labelled *a*).

This operation satisfies the axioms of a Wilke algebra since π_+ is associative. It therefore remains to show that this correspondence is bijective.

Note that every tree $t \in \mathbb{T}^{\text{wilke}}A$ is the unravelling of a finite graph all of which strongly connected components are either singletons or induced cycles (cycles without any additional edges).

For injectivity, suppose that there are two expansions $\mathfrak{A}_{o} = \langle A, \pi_{o} \rangle$ and $\mathfrak{A}_{1} = \langle A, \pi_{1} \rangle$ of \mathfrak{A} with the same associated ω -power. Let $t \in \mathbb{T}^{\text{wilke}}A$ be the unravelling of a graph \mathfrak{G} with *n* strongly connected components. By induction on *n*, we prove that

 $\pi_{\rm o}(t)=\pi_{\rm i}(t)\,.$

Let *C* be the strongly connected component of \mathfrak{G} containing the root of *t*. For every vertex $v \notin C$, it follows by inductive hypothesis that

$$\pi_{\mathrm{o}}(t|_{v}) = \pi_{\mathrm{i}}(t|_{v}).$$

Hence, replacing these subtrees by their product we may assume that \mathfrak{G} has a single strongly connected component *C*. If *C* is a single vertex, we have $t = \operatorname{sing}(a)$ and

$$\pi_{\rm o}(t)=a=\pi_{\rm i}(t)\,.$$

Otherwise, *C* is a cycle and there exists a finite path *p* such that $t = p^{\omega}$. This implies that

$$\pi_{\rm o}(t)=\pi(p)^{\omega}=\pi_{\rm i}(t)\,.$$

For surjectivity, suppose that $-^{\omega} : A_{\{z\}} \to A_{\emptyset}$ is an ω -power operation. We construct an expansion $\mathfrak{A}^+ = \langle A, \pi_+ \rangle$ of \mathfrak{A} as follows. Let $t \in \mathbb{T}^{\text{wilke}}A$ be the unravelling of a graph \mathfrak{G} with *n* strongly connected components. We define $\pi(t)$ by induction on *n*. Let *C* be the strongly connected component of \mathfrak{G} containing the root of *t*. For every vertex $v \notin C$, we can compute $\pi(t|_v)$ by inductive hypothesis. Let t' be the tree obtained from *t* by replacing every such subtree by its product. If $t' = \operatorname{sing}(a)$, we set

$$\pi_+(t) \coloneqq a$$

Otherwise, there exists a finite path *p* such that $t = p^{\omega}$ and we set

$$\pi_+(t) \coloneqq \pi(p)^\omega$$
 .

It remains to check that the function π_+ defined this way satisfies the axioms of a $\mathbb{T}^{\text{wilke}}$ -algebra and that the associated ω -operation coincides with the given one.

We start with the latter. Let $a \in A_{\{z\}}$. By definition of π_+ , there are numbers $m, n < \omega$ such that

$$\pi_+(aaa\dots)=\pi(a^m)\cdot\pi(a^n)^\omega=a^m\cdot(a^n)^\omega=a^\omega,$$

as desired.

For the unit law, it follows directly by definition that

$$\pi_+(\operatorname{sing}(a)) = a \, .$$

Next, let us show that π_+ is well-defined. Let \mathfrak{G} and \mathfrak{H} be two finite graphs with the same unravelling t. We have to show that we obtain the same result when defining $\pi_+(t)$ in terms of \mathfrak{G} and in terms of \mathfrak{H} . Since \mathfrak{H} and $\mathfrak{G} \times \mathfrak{H}$ have the same unravelling, we may assume that \mathfrak{G} is a quotient of \mathfrak{H} . We prove the claim by induction on the number of strongly connected components of \mathfrak{H} . Let C be the strongly connected component of \mathfrak{H} containing the root. For $v \notin C$, let \mathfrak{G}_v and \mathfrak{H}_v the subgraphs of \mathfrak{G} and \mathfrak{H} , respectively, reachable from the vertex v, and let $t|_v$ be their unravelling. By inductive hypothesis, the values of $\pi(t|_v)$ defined using these two graphs coincide. Hence, replacing each such subgraph by its product, we may assume that \mathfrak{H} has a single strongly connected component. Then so does \mathfrak{G} . If these components are singletons, the products are the same. Otherwise, there is some finite path p and numbers m, k such that \mathfrak{G} is a cycle consisting of m copies of p, and \mathfrak{H} consists of km copies. Setting $a := \pi(p)$, it follows that the product defined in terms of \mathfrak{G} is equal to $(a^m)^{\omega}$, while the one defined via \mathfrak{H} is $(a^{km})^{\omega}$. Since these values coincide, the claim follows.

Finally for associativity, fix $t \in \mathbb{T}^{\text{wilke}}\mathbb{T}^{\text{wilke}}A$ and let \mathfrak{G} be a finite graph with unravelling *t*. For each vertex *v* of \mathfrak{G} , we fix a finite graph \mathfrak{H}_v with unravelling t(v). Then flat(*t*) is the unravelling of the graph obtained from the disjoint union of all \mathfrak{H}_v by adding edges according to \mathfrak{G} . We prove that

$$\pi_+(\operatorname{flat}(t)) = \pi_+(\mathbb{T}^{\operatorname{wilke}}\pi_+(t))$$

by induction on the number of strongly connected components of the graph for flat(t). Let C be the strongly connected component of \mathfrak{G} containing the root. For every vertex $v \notin C$, it follows by inductive hypothesis that

$$\pi_+(\operatorname{flat}(t|_{\nu})) = \pi_+(\mathbb{T}^{\operatorname{wilke}}\pi_+(t|_{\nu})).$$

Replacing such subtrees by their respective products, we may therefore assume that \mathfrak{G} consists of a single strongly connected component *C*. If *C* is a singleton, we have $t = \operatorname{sing}(s)$ and

$$\pi_+(\text{flat}(t)) = \pi_+(s) = \pi_+(\mathbb{T}^{\text{wilke}}\pi_+(t)).$$

Hence, suppose that *C* is a cycle. Each graph \mathfrak{H}_{ν} consists of a finite path p_{ν} to which are possibly attached additional trees. By inductive hypothesis, associativity holds for these subtrees. Again, replacing each such subtree by its product, we may assume that \mathfrak{H}_{ν} is equal to p_{ν} . Consequently, flat(*t*) is a single infinite path consisting of the concatenation of all p_{ν} , while $\mathbb{T}^{\text{wilke}} \pi_{+}(t)$ is the infinite path labelled by the products $\pi(p_{\nu})$. The product of these two paths is the same. \Box

Corollary 5.11. Let $\mathfrak{A} = \langle A, \pi \rangle$ be a finitary \mathbb{T}^{fin} -algebra. There exists a bijection between all \mathbb{T}^{thin} -expansions of \mathfrak{A} and all functions $-^{\omega} : A_{\{z\}} \to A_{\emptyset}$ satisfying the axioms of a Wilke algebra.

Corollary 5.12. Every \mathbb{T}^{thin} -algebra \mathfrak{A} is uniquely determined by (1) its \mathbb{T}^{fin} -reduct and (11) the associated ω -semigroup.

5.2 EVALUATIONS WITH MERGING

When we try to go beyond \mathbb{T}^{thin} our machinery breaks down since we cannot use the results for semigroups anymore. The following counterexample shows that a naïve generalisation of our definitions does not work.

Lemma 5.13. There exists an MSO-definable \mathbb{T}^{reg} -algebra \mathfrak{A} and a tree $t \in \mathbb{T}A$ that has no simple \mathbb{T}^{reg} -evaluation.

Proof. Let \mathfrak{A} be the \mathbb{T}^{reg} -reduct of the Bojańczyk-Klin algebra from the example on page 25. Then the claim follows immediately from Corollary 5.6 (a). Nevertheless we give an explicit proof to see what exactly is going wrong. Set $\Delta := \mathbb{T}^{\text{fin}}\{a, b\}$ and recall that $\Delta \subseteq A$. We will prove by induction on α that

 $t \notin \operatorname{rng} \operatorname{term}_{\alpha}$, for all $t \in \mathbb{T}\Delta$ where every subtree has vertices of arbitrarily high arity.

For a contradiction, suppose otherwise. Let α be the minimal ordinal such that there is some simple evaluation $\gamma \in \mathbb{E}_{\alpha}(\mathfrak{A}, \mathbb{T})$ where every subtree of term(γ) has vertices of arbitrarily high arity. If $\alpha = 0$, then term(γ) = sing(a) in contradiction to our choice of γ . Hence, $\alpha = \beta + 1$, for some β . Fix $\nu \in \text{dom}(\gamma)$. Note that every subtree s of term_{β}($\gamma(\nu)$) has a simple evaluation in $\mathbb{E}_{\beta}(\mathfrak{A}, \mathbb{T})$ which, by inductive hypothesis, means that s has a subtree where the arity of the vertices is bounded. We claim that this implies that $t_{\nu} := \text{term}(\gamma(\nu))$ is finite. Suppose otherwise. Since t_{ν} has only finitely many variables, it has some infinite subtree swithout variables. But s is also a subtree of term(γ). By choice of γ this implies that the arities of the vertices of s are unbounded. A contradiction.

Hence, we have term $(\gamma(\nu)) \in \mathbb{T}^{\text{fin}} \Delta$, which implies that

 $\operatorname{val}(\gamma(\nu)) = \pi(\operatorname{term}(\gamma(\nu))) = \operatorname{term}(\gamma(\nu)).$

Furthermore, term($\gamma(\nu)$) being finite its arity is at least as high as the maximal arity of a vertex in dom($\gamma(\nu)$). It follows that, for every $n < \omega$, there is some

 $v \in \operatorname{dom}(\gamma)$ such that $\operatorname{val}(\gamma(v))$ has arity at least n. But $\gamma \in \mathbb{E}_{\alpha+1}(\mathfrak{A}, \mathbb{T})$ implies that $\operatorname{Tval}(\gamma) \in \operatorname{T}^{\operatorname{reg}} A$. In particular, $\operatorname{Tval}(\gamma)$ uses only finitely many different labels. This implies that their arity is bounded. A contradiction.

A closer look at the above proof reveals two possible reasons making simple evaluations impossible. Firstly, our counterexample uses a tree with infinitely many different labels. It still might be possible that trees with only finitely many different labels always have simple evaluations. Secondly, we made essential use of the fact that every factor of an infinite binary tree has a subtree that is itself an infinite binary tree. To be able to use factorisations of trees into pieces that are significantly simpler, we will probably have to allow more general factors, which then necessarily have infinitely many variables. Unfortunately, it is hard to combine these two modifications since factors with infinitely many variables usually give rise to infinitely many different elements of the algebra. What seems to be missing is some technique that, given a tree with infinitely many different labels, allows us to bound their arity by merging different variables (e.g., replacing a(x, y, z) by, say, a(x, x, z)).

This observation leads to the following attempt to allow for evaluations where variables are merged. To make our definitions precise we need a bit of terminology. First, as we want to identify variables, we need to work in \mathbb{T}^{\times} instead of \mathbb{T} . We also need a set of labels telling us which variables to identify.

Definition 5.14. (a) For a tree $t \in \mathbb{T}_{\zeta}^{\times}A$ and a function $\sigma : \zeta \to \xi$, we denote by ${}^{\sigma}t \in \mathbb{T}_{\xi_{o}}^{\times}A$ the tree obtained from *t* by replacing every variable *z* by $\sigma(z)$ (where $\xi_{o} \subseteq \xi$ is the range of σ).

If $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{\times}$ is closed under the operation $^{\sigma}$ - we can extend this operation to \mathbb{T}° -algebras \mathfrak{A} by setting

$${}^{\sigma}a \coloneqq \pi({}^{\sigma}\operatorname{sing}(a)), \text{ for } a \in A_{\zeta}.$$

(b) For a sort $\xi \in \Xi$, we set $\Gamma(\xi) := (\Gamma_{\zeta}(\xi))_{\zeta \in \Xi}$ where

$$\Gamma_{\zeta}(\xi) \coloneqq \{ \sigma \mid \sigma : \zeta \to \xi \}.$$

Given a tree $t \in \mathbb{T}^{\times}A$ we can choose some sort $\xi \in \Xi$ and functions $\sigma_{v} \in \Gamma(\xi)$, for every $v \in \text{dom}(t)$, and then replace every label t(v) by $\sigma_{v} t(v)$. The problem is that the resulting tree is not well-formed since the sorts do not match anymore. For instance, in the tree a(b, c) with $a = a(z_0, z_1)$ we can replace z_0 and z_1 by the same variable x. This produces the label a' := a(x, x) of arity $\{x\}$. Consequently, we need to produce a tree where the corresponding vertex has a single successor instead of two. Given the tree a(b, c) the only obvious choices for such a tree would be a'(b) or a'(c). This idea can be generalised as follows.

Definition 5.15. Let \mathfrak{A} be a \mathbb{T}° -algebra where $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{\times}$ is closed under the operations σ -, let $p : \Gamma(\xi) \times A \to \Gamma(\xi)$ and $q : \Gamma(\xi) \times A \to A$ be the two projections, and let $s \in \mathbb{T}^{\times}(\Gamma(\xi) \times A)$ be a tree. For $v \in \text{dom}(s)$, we set $\sigma_v := p(s(v))$. (If s(v) is a variable, we set $\sigma_v := \text{id}_{\emptyset}$.)

(a) A *choice function* for *s* is a family $\mu = (\mu_v)_{v \in \text{dom}(s)}$ of functions $\mu_v : \text{rng } \sigma_v \to \text{dom } \sigma_v$ such that $\sigma_v \circ \mu_v = \text{id.}$

(b) Given a choice function μ for *s*, we define the tree $s \parallel \mu \in \mathbb{T}^{\times}A$ as follows. For every $\nu \in \text{dom}(s)$,

- we delete from s all subtrees $s|_{suc_r(v)}$ with $x \notin rng \mu_v$, and
- for $x \in \operatorname{rng} \mu_{\nu}$, we change the *x*-successor of *v* to a $\sigma_{\nu}(x)$ -successor.
- we replace every label $s(v) = \langle \sigma, a \rangle$ by σa (which we consider to be an element of sort rng σ). If s(v) is a variable, we leave it unchanged.

Example. Let $t \in \mathbb{T}_{\emptyset}A$ be an infinite tree where all labels on the same level are equal, and let $s \in \mathbb{T}(\Gamma(\{x\}) \times A)$ be the tree with the same domain as t such that

 $s(v) = \langle \sigma_n, a_n \rangle$, for every vertex *v* with |v| = n,

where $a_n = t(v) \in A_{\zeta_n}$ is the label on level *n* of *t* and $\sigma_n : \zeta_n \to \{x\}$ is the function mapping all variables of a_n to the same variable *x*. For every choice function μ for *s*, we obtain a path

$$s \parallel \mu = b_0 b_1 \cdots$$
 with $b_n(x) \coloneqq a_n(x, \ldots, x)$.

Remark. Note that, in a tree of the form $s \parallel \mu$, every vertex has a sort which is a subset of ξ . In particular, the number of sorts used is finite.

We can produce well-formed trees $s \parallel \mu$ using a choice function μ . But which one do we take? The easiest case is if all choice functions produce the same result (cf. [23]), then it does not matter. (A more general construction will be presented further below.)

Definition 5.16. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{1} \subseteq \mathbb{T}^{\times}$ be submonads and \mathfrak{A} a \mathbb{T}° -algebra. (a) We say that a tree $s \in \mathbb{T}^{\times}(\Gamma(\xi) \times A)$ is *uniform* if

 $s \parallel \mu = s \parallel \mu'$, for all choice functions μ, μ' .

(b) A *uniform* \mathbb{T}° -*condensation* of $t \in \mathbb{T}^{1}A$ is a uniform tree $s \in \mathbb{T}^{1}(\Gamma(\xi) \times A)$ such that

 $\mathbb{T}^1 q(s) = t$ and $s \parallel \mu \in \mathbb{T}^o A$, for some/all choice functions μ .

(c) Let τ be a partial function mapping each tree *t* to some uniform \mathbb{T}° -condensation of *t* (if such a condensation exists). We set

 $\pi_{\tau}^{u}(t) \coloneqq \pi(\tau(t) \parallel \mu)$, where μ is an arbitrary choice function.

If $\tau(t)$ is undefined, we let $\pi_{\tau}^{u}(t)$ be undefined as well. We call π_{τ}^{u} -evaluations \mathbb{T}° -evaluations with uniform merging, and we denote the corresponding set by

 $\mathbb{E}^{\mathbf{u},\tau}_{\alpha}(\mathfrak{A},\mathbb{T}^{1}) \coloneqq \mathbb{E}_{\alpha}(\pi^{\mathbf{u}}_{\tau},\mathbb{T}^{1}).$

Example. The tree *s* from the preceding example is a uniform \mathbb{T}^{thin} -condensation of *t*.

Example. $\mathbb{T}^{\times \text{reg}}$ -evaluations with uniform merging were introduced in [23] where they were used to derive decidability results for trees. To do so Puppis considers trees $t \in \mathbb{T}^{\times}X$ such that (in our terminology), for every MSO-definable $\mathbb{T}^{\times \text{reg}}$ algebra \mathfrak{A} and every function $\beta : X \to A$, the image $\mathbb{T}^{\times}\beta(t)$ has an evaluation $\gamma \in \mathbb{E}_{n}^{\text{u, t}}(\mathfrak{A}, \mathbb{T}^{\times})$, for some number $n < \omega$ independent of β and \mathfrak{A} . The function τ chooses condensations based on the runs of an automaton recognising the product of \mathfrak{A} . (The details can be found in [23]. A similar construction is used at the beginning of the proof of Lemma 5.22 below.) Let us call such trees *reducible*.

By induction on n, we can transform every reducible tree t into a regular tree t_0 with the same value as t. Puppis considers reducible trees t where this transformation is computable using a particular algorithm. (We again omit the details.) Let us call such trees *effectively reducible*. [23] contains the following results.

- Every deterministic tree in the Caucal hierarchy is effectively reducible.
- The class of effectively reducible trees is closed under a number of natural operations.

L

Every effectively reducible tree has a decidable MSO-theory.

The key technical result of [23] is the following recipe of how to evaluate products in an MSO-definable \mathbb{T}^{\times} -algebra using uniform evaluations.

Proposition 5.17 (Puppis). Let \mathfrak{A} be an MSO-definable \mathbb{T}^{\times} -algebra and $\xi \in \Xi$ a sort. There exists an MSO-definable \mathbb{T}^{\times} -algebra \mathfrak{B} , a morphism $\rho : \mathbb{T}^{\times}A \to \mathfrak{B}$, and MSO-formulae φ_a , for $a \in A_{\xi}$, such that, given a tree $T \in \mathbb{T}_{\xi}^{\times}\mathbb{T}^{\times}A$ and a uniform \mathbb{T}^{\times} -condensation s of $\mathbb{T}^{\times}\rho(T)$, we have

$$\pi(\operatorname{flat}(T)) = a \quad \text{iff} \quad s \parallel \mu \vDash \varphi_a, \text{ for all } a \in A_{\xi} \text{ and all choice}$$

functions μ .

The proof uses similar techniques as that of Lemma 5.22 below.

Since evaluations with uniform merging generalise simple evaluations, they allow us to decompose more trees. Unfortunately, there are still trees left without an evaluation. We can generalise our evaluations even further by not requiring that all choice functions lead to the same tree, but only to one that is 'equivalent'.

Definition 5.18. Let $s \in \mathbb{T}^{\times}(\Gamma(\xi) \times A)$ and let $\pi : \mathbb{T}^{\circ}A \to A$ be a function. We call *s* π -*consistent* if

$s _{\nu} \parallel \mu \in \mathbb{T}^{\mathrm{o}}A$,	for every choice function μ of $s _{\nu}$ and each
	vertex $v \in \operatorname{dom}(s)$,
$\pi(s _{\nu} \parallel \mu) = \pi(s _{\nu} \parallel \mu'),$	for all choice functions μ , μ' of $s _{\nu}$ and each
	vertex $v \in \operatorname{dom}_{o}(s)$.

Definition 5.19. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{1} \subseteq \mathbb{T}$ be submonads, $\mathfrak{A} = \langle A, \pi \rangle$ a \mathbb{T}° -algebra, and $t \in \mathbb{T}^{1}A$.

(a) A consistent \mathbb{T}° -condensation of t is a π -consistent tree $s \in \mathbb{T}^{1}(\Gamma(\xi) \times A)$ such that $\mathbb{T}^{1}q(s) = t$.

(b) Let τ be a partial function mapping each tree *t* to some consistent \mathbb{T}° -condensation of *t* (if such a condensation exists). We set

 $\pi_{\tau}^{c}(t) \coloneqq \pi(\tau(t) \parallel \mu)$, where μ is an arbitrary choice function.

If $\tau(t)$ is undefined, we let $\pi_{\tau}^{c}(t)$ be undefined as well. We call π_{τ}^{c} -evaluations \mathbb{T}° -evaluations with consistent merging, and we denote the corresponding set by

$$\mathbb{E}^{\mathsf{c},\tau}_{\alpha}(\mathfrak{A},\mathbb{T}^1) \coloneqq \mathbb{E}_{\alpha}(\pi^{\mathsf{c}}_{\tau},\mathbb{T}^1).$$

Clearly, consistent merging generalises uniform merging. While $\mathbb{T}^{\times reg}$ -evaluations with uniform merging seem to exist only in special cases, our hope is that $\mathbb{T}^{\times reg}$ -evaluations with consistent merging always exist (at least for MSO-definable

algebras). At the moment we are only able to obtain partial results. To do so we need a bit of terminology. First, it is convenient to work with arbitrary directed graphs instead of just trees.

Definition 5.20. Let *A* be a sorted set.

(a) An *A*-labelled rooted graph g is a countable directed graph with a distinguished vertex r, the root of g, where every vertex is labelled by some element of A and where every edge is labelled by some variable in such a way that, if a vertex v is labelled by $a \in A_{\xi}$, then v has exactly one out-going edge labelled x, for every $x \in \xi$. As usual, we treat an A-labelled rooted graph as a function $g : \text{dom}(g) \to A$. We denote by $\mathbb{R}_{\xi}A$ the set of all $(A + \xi)$ -labelled rooted graphs (where, as usual, the variables are considered to be elements of sort \emptyset), and we set $\mathbb{R}A := (\mathbb{R}_{\xi}A)_{\xi}$.

(b) Given a graph $g \in \mathbb{RR}A$, we denote by flat(g) the graph obtained from the disjoint union $\sum_{v \in \text{dom}(g)} g(v)$ of all component graphs by replacing, in every component g(v), every x-labelled edge $u \to u'$ to a vertex labelled by some variable z to an x-labelled edge $u \to w$ where w is the root of the the component g(v') for the z-successor v' of v.

(c) We denote by $\mathbb{R}^{\text{thin}} A \subseteq \mathbb{R}A$ the set of all graphs whose unravelling is a thin tree.

We start with a technical lemma which is based on the following variant of a condensation. In a usual condensation we can only redirect edges to another successor of the same vertex. Below we will need a variant where we can also redirect edges to vertices that are farther away. We specify which destinations are allowed in such a redirection via a labelling σ of the tree by numbers.

Definition 5.21. Let $t \in \mathbb{T}_{\xi}^{\times} A$ be a tree and $\sigma : \operatorname{dom}(t) \to [N]$ a function. (a) The σ -parent $p_{\sigma}(v)$ of $v \in \operatorname{dom}(t)$ is the maximal vertex u satisfying

u < v and $\sigma(u) \ge \sigma(v)$.

If no such vertex exists, we set $p_{\sigma}(v) := \bot$ and we say that v does not have a σ -parent. We define a relation \approx_{σ} on dom(t) by

 $u \approx_{\sigma} v$: iff $p_{\sigma}(u) = p_{\sigma}(v)$.

(b) We say that a tree $t' \in \mathbb{T}^{\times} A$ is a σ -rewiring of t if it is the unravelling of some graph $g \in \mathbb{R}_{\xi} A$ satisfying the following conditions.

• $dom(g) \subseteq dom(t)$ is prefix-closed and non-empty.

- If v is the x-successor of u in t and u, v ∈ dom(g), then v is also the x-successor of u in g.
- If v is the x-successor of u in g and v' is its x-successor in t, then $v \approx_{\sigma} v'$.

Example. Let *t* be the tree on the left and σ the depicted labelling of *t*. Then the (unravelling of the) graph *g* on the right is a σ -rewiring of *t*.



After these preparations we can state and proof our key technical lemma.

Lemma 5.22. Let \mathfrak{A} be an MSO-definable \mathbb{T}^{\times} -algebra and $\xi \in \Xi$ a sort. There exists a constant $N < \omega$ with the following property. For every tree $t \in \mathbb{T}_{\xi}^{\times}A$, we can find a function $\sigma : \operatorname{dom}(t) \to [N]$ such that

$$\pi(t') = \pi(t)$$
, for every σ -rewiring t' of t.

Proof. Fix a tree $t \in \mathbb{T}_{\xi}^{\times}A$, let $\mathcal{A} = \langle Q, A + \xi, \Delta, q_o, \Omega \rangle$ be the automaton checking that the product of a given tree in $\mathbb{T}_{\xi}^{\times}A$ is equal to $\pi(t)$, and let \mathcal{G} be the corresponding Automaton-Pathfinder game for \mathcal{A} on the input tree t. Without loss of generality, we may assume that \mathcal{A} is a non-deterministic automaton. Since \mathcal{G} is a parity game, there exists a positional winning strategy τ for Automaton.

We use this strategy τ to define an additive labelling λ on t as follows. Given vertices u < v, we set

 $\lambda(u,v) \coloneqq \langle p,k,q \rangle,$

where p, k, q are chosen such that the unique play of \mathcal{G} conforming to the strategy τ that reaches the vertices u and v, contains the positions $\langle u, p \rangle$ and $\langle v, q \rangle$, and k is the minimal priority seen in the part of this play between these two positions (inclusive).

Let us properly define the semigroup \mathfrak{S} used by the labelling λ . The set of elements is

$$S \coloneqq Q \times D \times Q + \{\bot\},\$$

where $D := \operatorname{rng} \Omega$ is the set of priorities used by \mathcal{A} and \perp is used as an 'error element'. The multiplication is given by

$$\langle p, k, p' \rangle \cdot \langle q, l, q' \rangle \coloneqq \begin{cases} \langle p, m, q' \rangle & \text{if } p' = q \text{ and } m \coloneqq \min \{k, l\}, \\ \bot & \text{otherwise}, \end{cases}$$

and every product involving \perp evaluates to \perp .

By Theorem 3.4, there exists a weak Ramseyan split $\chi : \text{dom}(t) \rightarrow [N]$ for λ . We start by showing that

(†)
$$u \sqsubset_{\chi} v \sqsubset_{\chi} w$$
 implies $\lambda(u, v) = \lambda(v, w)$.

(This is a property of the particular semigroup \mathfrak{S} we are using.) Hence, suppose that $u \sqsubset_{\chi} v \sqsubset_{\chi} w$ and set $a \coloneqq \lambda(u, v)$ and $b \coloneqq \lambda(v, w)$. Since χ is weakly Ramseyan,

$$u \sqsubset_{\chi} v \sqsubset_{\chi} w$$
 implies $a \cdot b = a$ and $b \cdot a = b$.

These two equations imply that, either $a = \bot = b$, or we have $a = \langle p, k, p' \rangle$ and $b = \langle q, l, q' \rangle$ where

$$p' = q$$
, $p' = q'$, $k = \min\{k, l\}$,
 $q' = p$, $q' = p'$, $l = \min\{k, l\}$.

Hence, p = p' = q = q' and k = l, as desired.

We define a vertex labelling μ : dom $(t) \rightarrow [N] \times (Q \times D \times Q)$ by

 $\mu(v) \coloneqq \left\langle \chi(v), \lambda(p_{\chi}(v), v) \right\rangle.$

(In the case where $p_{\chi}(\nu) = \bot$, we set $\mu(\nu) := \langle \chi(\nu), \lambda(\langle \rangle, \nu) \rangle$.) Fix some bijection $h := \operatorname{rng} \mu \to [N]$ satisfying

$$i < j$$
 implies $h(\langle i, a \rangle) < h(\langle j, b \rangle)$.

We claim that $\sigma := h \circ \mu$ is the desired function.

Let t' be a σ -rewiring and let \mathcal{G}' be the Automaton-Pathfinder game on t'. We have to show that $\pi(t') = \pi(t)$. Note that the Automaton-Pathfinder game \mathcal{G}' for \mathcal{A} on t' can be obtained from the game \mathcal{G} for t by removing some positions and redirecting some of the edges. Consequently, the strategy τ for t induces a strategy τ' for the game on t'. For a path v (in \mathcal{G} or \mathcal{G}'), we denote by $\Omega(v)$ the least priority seen along v. We start by proving that, given two positions $\langle u, q \rangle$ and $\langle u', q' \rangle$ of Automaton in \mathcal{G}' that are both reachable by a play conforming to τ' and that satisfy

$$u < u'$$
 and $\sigma(u) \ge \sigma(x)$, for all $u \le x \le u'$,

we have

$$\Omega(v) = \Omega(v'), \text{ for all plays } v, v' \text{ of } \mathcal{G}' \text{ starting in } \langle u, q \rangle \text{ and}$$

ending in $\langle u', q' \rangle$ that both conform to τ' .

We prove the claim by a double induction, first on $\sigma(u')$ (starting with large values), and then on the lengths of v and v' (which are equal). We distinguish three cases.

(I) If *v* and *v'* share some intermediate position $\langle u'', q'' \rangle$ such that

$$\sigma(u'') \ge \sigma(x)$$
, for all $u'' \le x \le u'$,

we can split the plays at this position. Let v_0, v_1, v'_0, v'_1 be the corresponding parts. By inductive hypothesis, it follows that

$$\Omega(v_{o}) = \Omega(v'_{o})$$
 and $\Omega(v_{1}) = \Omega(v'_{1})$.

This implies that $\Omega(v) = \Omega(v')$.

(II) If v and v' have some common suffix, i.e., $v = v_0 v_*$ and $v' = v'_0 v_*$, we obtain $\Omega(v_0) = \Omega(v'_0)$ by inductive hypothesis, and it follows that $\Omega(v) = \Omega(v')$.

(III) Finally, suppose that cases (I) and (II) do not hold. Let $\langle v, s \rangle$ and $\langle v', s' \rangle$ be the second but last positions of Automaton in the plays v and v', respectively (see Figure 5 (a)). Since we are not in case (II), we have $v \neq v'$. Let w and w' be the successors of, respectively, v and v' in t that correspond to the edges $v \rightarrow u'$ and $v' \rightarrow u'$ in t'. Let ρ be the play that conforms to τ' , starts at $\langle u, q \rangle$, ends in $\langle w, q' \rangle$, and that uses only edges that are present in t. Similarly, let ρ' be the plays



Figure 5: (a) Left: the plays v and v'. (b) Right: the plays ρ_i and v'_i , projected to the game \mathcal{G} .

 v, v', ρ , and ρ' ending in the position $\langle v, s \rangle$ and $\langle v', s' \rangle$, respectively. By inductive hypothesis, we have

$$\Omega(\rho_{o}) = \Omega(v_{o})$$
 and $\Omega(\rho'_{o}) = \Omega(v'_{o})$.

Furthermore, by definition of a σ -rewiring, we have $w \approx_{\sigma} u' \approx_{\sigma} w'$, which implies that

$$p_{\sigma}(w) = p_{\sigma}(w') = p_{\sigma}(u').$$

As we are not in case (I), we also have $p_{\sigma}(u') = u$. This implies that $u \sqsubset_{\chi} w$ and $u \sqsubset_{\chi} w'$. By (†) and the fact that χ is weak Ramseyan, it follows that

$$\begin{split} \lambda(u,w) &= \lambda(u,p_{\chi}(w)) = \lambda(p_{\chi}(w),w) \\ &= \lambda(p_{\chi}(w'),w') = \lambda(u,p_{\chi}(w')) = \lambda(u,w') \,. \end{split}$$

Consequently, we have $\Omega(\rho) = \Omega(\rho')$, which implies that

$$\Omega(v) = \Omega(\rho) = \Omega(\rho') = \Omega(v').$$

Having proved the above claim we can now show that $\pi(t') = \pi(t)$, i.e., that \mathcal{A} accepts the tree t'. We claim that the strategy τ' defined above is winning. To

do so it is sufficient to show that, every play v' in \mathcal{G}' conforming to τ' induces a play v in \mathcal{G} conforming to τ such that the minimal priorities seen infinitely often along v and v' are the same.

Let *k* be the maximal number such that v' contains infinitely many positions $\langle v, q \rangle$ of Automaton with $\sigma(v) = k$, and let $\langle v_i, q_i \rangle_{i < \omega}$ be an increasing enumeration of these positions. We factorise the play as $v_0 = v'_* v'_0 v'_1 \dots$ where v'_* is the prefix ending in $\langle v_0, q_0 \rangle$ and v'_i is the part between the positions $\langle v_i, q_i \rangle$ and $\langle v_{i+1}, q_{i+1} \rangle$.

For every $i < \omega$, let ρ_i be the partial play of \mathcal{G}' from $\langle v_i, q_i \rangle$ to $\langle v_{i+1}, q_{i+1} \rangle$ that conforms to τ' and that uses only edges belonging to \mathcal{G} . (Such a play exists since $p_{\chi}(v_{i+1}) = v_i$ and t' is a σ -rewiring of t; see Figure 5 (b).) The composition

$$v \coloneqq v'_* \rho_0 \rho_1 \rho_2 \dots$$

forms a play in \mathcal{G} conforming to τ and, by the above claim, we have $\Omega(\rho_i) = \Omega(v'_i)$. Hence, the least priority seen infinitely often in v is the same as in v'.

Theorem 5.23. Let \mathfrak{A} be an MSO-definable \mathbb{T}^{\times} -algebra. Every tree $t \in \mathbb{T}^{\times}A$ has a consistent \mathbb{T}^{\times} -condensation s such that

 $\pi(t) = \pi(s \parallel \mu)$, for all choice functions μ .

Proof. Let σ : dom $(t) \rightarrow [N]$ be the function from Lemma 5.22. Fix a set of variables ξ with $|\xi| = N$ and let $\mu : [N] \rightarrow \xi$ be a bijection. We define the desired condensation *s* by

$$s(v) \coloneqq \langle \tau_v, t(v) \rangle$$
 with $\tau_v(x) \coloneqq \mu(\sigma(u_x))$,

where u_x is the *x*-successor of *v*. Then every tree of the form $s \parallel \mu$ is a σ -rewiring of *t*. By choice of σ this implies that $\pi(s \parallel \mu) = \pi(t)$.

In particular, it follows that every MSO-definable \mathbb{T}^{\times} -algebra has \mathbb{T}^{\times} -evaluations with consistent merging. Note that this statement is not as trivial as it sounds since trees can contain labels of arbitrarily high arity, while every \mathbb{T}^{\times} condensation produces a tree where these arities are bounded. In particular, the statement is false for \mathbb{T}^{\times} -evaluations with uniform merging.

Our hope is that a more elaborate version of the construction from the proof of Lemma 5.22 can be used to construct a $\mathbb{T}^{\times reg}$ -condensation instead of a \mathbb{T}^{\times} -one, or at least that we can iterate such a construction to obtain a $\mathbb{T}^{\times reg}$ -evaluation.

Conjecture. Let \mathfrak{A} be an MSO-definable $\mathbb{T}^{\times \operatorname{reg}}$ -algebra. Then every tree $t \in \mathbb{T}A$ has a $\mathbb{T}^{\times \operatorname{reg}}$ -evaluation with consistent merging.

If we relax our notion of an evaluation and of consistent merging a bit, we even obtain something similar to ' $\mathbb{T}^{\times \text{thin}}$ -evaluations'. (But note that, formally, this notion does not exist, since $\mathbb{T}^{\times \text{thin}}$ does not form a monad.)

Theorem 5.24. Let \mathfrak{A} be an MSO-definable \mathbb{T}^{\times} -algebra and $C \subseteq A$ a subset. For every $t \in \mathbb{T}_{\xi}^{\times}C$, there is some $s \in (\mathbb{T}^{\times \text{thin}})^n C$ with $\pi(\text{flat}^{n-1}(s)) = \pi(t)$, where the exponent $n < \omega$ only depends on \mathfrak{A} and ξ .

Proof. Let σ : dom $(t) \rightarrow [n]$ be the function from Lemma 5.22 and let

$$D := \{ {}^{\tau}c \mid c \in C_{\zeta}, \ \tau : \zeta \to \eta \text{ surjective, } \eta, \zeta \in \Xi \}$$

be the closure of *C* under variable substitutions. Without loss of generality we may assume that σ maps the root to n - 1. By induction on $i \le n$, we construct graphs $s_i \in (\mathbb{T}^{\times})^{n-i} (\mathbb{R}^{\text{thin}})^i D$ such that

- (the unravelling of) $\operatorname{flat}^{n-1}(s_i) \in \mathbb{R}D$ is a σ -rewiring of t and,
- if *v* is the vertex of *t* corresponding to the root of the graph s_i(u_o)…(u_{k-1}), then σ(v) ≥ n − k − 1.

Then we can produce the desired tree *s* as follows. As each $d \in D$ is of the form $d = \pi(r)$, for some $r \in \mathbb{T}^{\text{fin}}C$, there exists a function $\vartheta : D \to \mathbb{T}^{\text{fin}}C$ with $\pi \circ \vartheta = \text{id}$. We chose for $s \in (\mathbb{T}^{\times \text{thin}})^{n+1}C$ the unravelling (at all n + 1 levels) of the graph $\mathbb{R}^n \vartheta(s_n) \in (\mathbb{R}^{\text{thin}})^{n+1}C$. Since the unravelling s'_n of flatⁿ⁻¹(s_n) is a σ -rewiring of *t* it follows that

$$\pi(\operatorname{flat}^n(s)) = \pi(s'_n) = \pi(t),$$

as desired.

It remains to construct s_0, \ldots, s_n . For s_0 , we choose the \mathbb{T}^{\times} -evaluation of t induced by σ as in Lemma 5.2. For the inductive step, suppose that we have already defined s_i . For every component $s_i(u_0)\cdots(u_{n-i-2})$ there exists a canonical injection

$$\operatorname{dom}(s_i(u_0)\cdots(u_{n-i-2})) \to \operatorname{dom}(\operatorname{flat}^{n-1}(s_i))$$

mapping the root of every tree $s_i(u_0)\cdots(u_{n-i-2})(u_{n-i-1})$ to the corresponding vertex of flat^{*n*-1}(*s_i*). Since dom(flat^(*n*-1)(*s_i*)) is contained in dom(*t*), the restriction of σ to the range of this injection induces a function

$$\sigma_i: \operatorname{dom}(s_i(u_0)\cdots(u_{n-i-2})) \to [n].$$

By inductive hypothesis, we have $\sigma_i(v) \ge i$, for all v. Let s_{i+1} be obtained from s_i by replacing every component $s_i(u_0)\cdots(u_{n-i-2}) \in \mathbb{T}^{\times}(\mathbb{R}^{\text{thin}})^i A$ by a path $r \in (\mathbb{R}^{\text{thin}})^{i+1}A$ which we construct as follows. We define a branch v_0, v_1, \ldots of the tree $s_i(u_0)\cdots(u_{n-i-1})$ starting at the root v_0 . This branch will form the domain of r. For the inductive step, suppose that we have already defined the vertex v_j and the labels $r(v_k)$, for all k < j. Let

$$s_i(u_0)\cdots(u_{n-i-2})(v_j)=a(\bar{x},\bar{y}),$$

where \bar{x} are the variables such that the corresponding successors w satisfy $\sigma_i(w) = i$ and let \bar{y} be those with $\sigma_i(w) > i$. If \bar{x} is empty, we set

$$r(v_i) \coloneqq a(\bar{y})$$

and the construction terminates. Otherwise, we pick one variable $x' \in \bar{x}$, we choose for v_{j+1} the x'-successor of v_j , and we set

$$r(v_j) \coloneqq a(x' \dots x', \bar{y}).$$

Note that the flattening $\text{flat}^{n-1}(s_{i+1})$ of the resulting tree s_{i+1} is a σ -rewiring of *t*.

As already mentioned above, we cannot conclude that $\mathbb{T}^{\times \text{thin}}$ is dense in \mathbb{T}^{\times} since the former does not form a monad. But we obtain the following, weaker statements.

Corollary 5.25. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}^{\times}$ be the closure of $\mathbb{T}^{\times \text{thin}}$ under flat. Then \mathbb{T}° is dense in \mathbb{T}^{\times} over the class of all MSO-definable \mathbb{T}^{\times} -algebras.

Corollary 5.26. Let \mathfrak{A} be an MSO-definable \mathbb{T}^{\times} -algebra. A set $C \subseteq A$ induces a subalgebra of \mathfrak{A} if, and only if,

 $\pi(t) \in C$, for all $t \in \mathbb{T}^{\times \text{thin}} C$.

Corollary 5.27. The product of every MSO-definable \mathbb{T}^{\times} -algebra \mathfrak{A} is uniquely determined by its restriction to the set $\mathbb{T}^{\times \text{thin}}A$.

Proof. Suppose that there are two MSO-definable \mathbb{T}^{\times} -algebras $\mathfrak{A}_{\circ} = \langle A, \pi_{\circ} \rangle$ and $\mathfrak{A}_{1} = \langle A, \pi_{1} \rangle$ the same universe *A* and whose products have the same restriction to $\mathbb{T}^{\times \text{thin}} A$. Fix a tree $t \in \mathbb{T}^{\times} A$. We have to show that $\pi_{\circ}(t) = \pi_{1}(t)$.

Let $\delta : A \to A \times A$ be the diagonal map, let $\Delta := \operatorname{rng} \delta \subseteq A \times A$ be its range, and set $s := \mathbb{T}^{\times} \delta(t) \in \mathbb{T}^{\times} \Delta$. As the product $\mathfrak{A}_{0} \times \mathfrak{A}_{1}$ is also MSO-definable, we can use Theorem 5.24 to find a tree $r \in (\mathbb{T}^{\times \operatorname{thin}})^{n} \Delta$ with $\pi(\operatorname{flat}^{n}(r)) = \pi(s)$. Since π_{0} and π_{1} agree on all trees in $\mathbb{T}^{\times \operatorname{thin}} A$, we have

 $\pi(u) \in \Delta$, for all $u \in \mathbb{T}^{\times \text{thin}} \Delta$.

Consequently, we can evaluate

$$\pi(\operatorname{flat}^{n}(r)) = \pi(\mathbb{T}^{\times \operatorname{thin}} \pi(\mathbb{T}^{\times \operatorname{thin}} \mathbb{T}^{\times \operatorname{thin}} \pi(\cdots(\mathbb{T}^{\times \operatorname{thin}})^{n} \pi(r) \cdots)))$$

by recursion using only products of trees in $\mathbb{T}^{\times \text{thin}}\Delta$. In particular, we have $\pi(\text{flat}^n(r)) \in \Delta$ and it follows that

$$\langle \pi_{o}(t), \pi_{1}(t) \rangle = \pi(s) = \pi(\operatorname{flat}^{n}(r)) \in \Delta \quad \text{implies} \quad \pi_{o}(t) = \pi_{1}(t) \,. \quad \Box$$

6 CONSISTENT LABELLINGS

As we have seen in the previous section, we can construct expansions with the help of evaluations if the two monads in question are sufficiently well-behaved. What do we do if they are not? Let us turn to a second idea of how to prove that a \mathbb{T}° -algebra \mathfrak{A} has a \mathbb{T} -expansion: when we want to define the product of $t \in \mathbb{T}A$, we first annotate t with additional information that makes it easier to determine the value of the product. For instance, for each vertex v, we can guess the value $\pi(t|_v)$ of the corresponding subtree and then check that these guesses are correct.

Definition 6.1. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}$ be a submonad, $\mathfrak{A}^{\circ} = \langle A, \pi^{\circ} \rangle$ a \mathbb{T}° -algebra, and $t \in \mathbb{T}_{\xi}A$.

(a) A *labelling* of t is a function $\lambda : \text{dom}(t) \to A$ (not necessarily arity-preserving) such that, for every vertex v,

 $\lambda(v) \in A_{\zeta}$ iff ζ is the set of variables appearing in $t|_{v}$.

(b) A labelling $\lambda : \operatorname{dom}(t) \to A$ is *weakly* \mathbb{T}° -consistent if, for every factor $[u, \bar{v})$ with $t[u, \bar{v}) \in \mathbb{T}^{\circ}A$,

L

$$\lambda(u) = \pi(t[u, \bar{v}))(\lambda(v_{o}), \lambda(v_{1}), \dots).$$

Example. For every \mathbb{T} -algebra \mathfrak{A} and every tree $t \in \mathbb{T}A$, we can define a labelling by

$$\lambda(v) \coloneqq \pi(t|_v).$$

This labelling is obviously weakly \mathbb{T} -consistent and, hence, weakly \mathbb{T}° -consistent for every $\mathbb{T}^{\circ} \subseteq \mathbb{T}$. In particular, if a \mathbb{T}° -algebra has a \mathbb{T} -expansion, then every tree has at least one weakly \mathbb{T}° -consistent labelling.

Weak consistency is based on factors with finitely many variables. In many situations this is not sufficient and we have to use the following stronger version of consistency where we also allow factors with infinitely many variables.

Definition 6.2. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}$ be a submonad, $\mathfrak{A}^{\circ} = \langle A, \pi^{\circ} \rangle$ a \mathbb{T}° -algebra, and $t \in \mathbb{T}_{\xi}A$.

(a) Given a factor $[u, \bar{v})$ of t, possibly with infinitely many holes \bar{v} , we denote by $t[u, \bar{v})(a_0, a_1, ...)$ the tree obtained from $t[u, \bar{v})$ by replacing each leaf labelled by a variable x_i by the tree sing (a_i) .

(b) A labelling $\lambda : \text{dom}(t) \to A$ is strongly \mathbb{T}° -consistent if, for every factor $[u, \bar{v})$, possibly with infinitely many holes \bar{v} , with $t[u, \bar{v})(\lambda(v_\circ), \lambda(v_1), \dots) \in \mathbb{T}^\circ A$, we have

L

$$\lambda(u) = \pi(t[u, \bar{v})(\lambda(v_{o}), \lambda(v_{1}), \dots)).$$

We start with the following easy observation.

Lemma 6.3. Let \mathfrak{A} be a finitary \mathbb{T}^{fin} -algebra. Every tree $t \in \mathbb{T}A$ has a strongly \mathbb{T}^{fin} -consistent labelling.

Proof. We call a labelling λ of some tree *t* locally consistent if

 $\lambda(v) = t(v) \big(\lambda(u_{o}), \ldots, \lambda(u_{n-1}) \big),$

for every vertex ν with successors u_0, \ldots, u_{n-1} . Fix an increasing sequence $P_0 \subset P_1 \subset \cdots \subset \text{dom}(t)$ of finite prefixes of t with $\bigcup_i P_i = \text{dom}(t)$, and let Λ_i be the set of all locally consistent labellings of P_i , for $i < \omega$. Then $\Lambda := \bigcup_i \Lambda_i$ ordered by \subset forms a finitely-branching tree. By Kőnig's Lemma, there exists an infinite branch $\lambda_0 \subset \lambda_1 \subset \cdots$. Let λ be its limit. Then λ is locally consistent.

It therefore, remains to prove that every locally consistent labelling of t is \mathbb{T}^{fin} consistent. Consider a finite factor H of t with root v and leaves u_0, \ldots, u_{m-1} . By
induction on |H| it follows that

$$\lambda(v) = \pi((t \upharpoonright H)(\lambda(u_0), \dots, \lambda(u_{m-1}))).$$

Next, let us show how to use consistent labellings to characterise possible \mathbb{T} -expansions of a given \mathbb{T}° -algebra. We need the following additional property.

Definition 6.4. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}$.

(a) A weak labelling scheme for a \mathbb{T}° -algebra \mathfrak{A} is a function σ assigning to each tree $t \in \mathbb{T}A$ a weakly \mathbb{T}° -consistent labelling $\sigma(t)$ of t. Similarly, a *strong labelling scheme* σ assigns to each tree $t \in \mathbb{T}A$ a strongly \mathbb{T}° -consistent labelling $\sigma(t)$.

(b) A labelling scheme σ for \mathfrak{A} is *associative* if, for every tree $T \in \mathbb{TT}A$, we have

$$\sigma(t) = \sigma(\operatorname{flat}(T)) \circ \mu,$$

where μ : dom(T) \rightarrow dom(flat(T)) maps each vertex $\nu \in$ dom(T) to the vertex of flat(T) corresponding to the root of $T(\nu)$, and t is the tree such that

$$t(v) \coloneqq \sigma(T(v))(\langle \rangle), \quad \text{for } v \in \text{dom}(T).$$

Example. There are algebras with several associative strong labelling schemes. Let \mathfrak{A} be the \mathbb{T}^{thin} -algebra with domains $A_{\xi} := [n]$, for some fixed number $n < \omega$, where the product is just the maximum

$$\pi(t) \coloneqq \max\left\{ t(v) \mid v \in \operatorname{dom}(t) \right\}.$$

For every k < n, we obtain an associative labelling scheme σ_k defined by

$$\sigma_k(t)(v) \coloneqq \begin{cases} \pi(t|_v) & \text{if } t|_v \in \mathbb{T}^{\text{thin}}A, \\ \max\{k\} \cup \{t(u) \mid u \ge v\}, & \text{otherwise}. \end{cases}$$

There is a tight connection between \mathbb{T} -expansions and associative labelling schemes (weak or strong, it does not matter).

Proposition 6.5. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}$ and let \mathfrak{A} be a \mathbb{T}° -algebra.

- (a) Every associative weak labelling scheme for \mathfrak{A} is strong.
- (b) There exists a bijective correspondence between associative labelling schemes σ and T-expansions of A.

Proof. We define two mutually-inverse functions mapping (I) every associative weak labelling scheme to a \mathbb{T} -expansion of \mathfrak{A} and (II) every such expansion to an associative strong labelling scheme.

(I) Given a weak scheme σ we define the corresponding function π_+ by

$$\pi_+(t) \coloneqq \sigma(t)(\langle \rangle), \quad \text{for } t \in \mathbb{T}A.$$

Then π_+ extends π since weak \mathbb{T}° -consistency of σ implies that

$$\pi(t) = \sigma(t)(\langle \rangle) = \pi_+(t), \text{ for } t \in \mathbb{T}^{\circ}A.$$

Hence, it remains to show that π_+ is associative. Fix $T \in \mathbb{TT}A$. By the definition of associativity, we have

$$\sigma(\mathbb{T}\pi_+(T)) = \sigma(\operatorname{flat}(T)) \circ \mu,$$

which in particular implies that

$$\pi_{+}(\mathbb{T}\pi_{+}(T)) = \sigma(\mathbb{T}\pi_{+}(T))(\langle \rangle)$$
$$= \sigma(\operatorname{flat}(T))(\langle \rangle) = \pi_{+}(\operatorname{flat}(T)).$$

(II) Conversely, given a product π_+ : $\mathbb{T}A \to A$ we define a strong scheme σ by

$$\sigma(t)(v) \coloneqq \pi_+(t|_v)$$
, for $t \in \mathbb{T}A$ and $v \in \text{dom}(t)$.

To show that this function σ is a strong labelling scheme, fix a factor [u, v] of some tree $t \in \mathbb{T}A$. Then it follows by associativity of π_+ that

$$\begin{aligned} \sigma(u) &= \pi_+(t|_u) \\ &= \pi_+(t[u,\bar{v})(\pi_+(t|_{v_o}),\pi_+(t|_{v_1}),\dots)) \\ &= \pi_+(t[u,\bar{v})(\sigma(v_o),\sigma(v_1),\dots)), \end{aligned}$$

as desired. To show that σ is associative, let $T \in \mathbb{TT}A$, $v \in \text{dom}(T)$, and let *t* be the tree from the definition of associativity. Then

$$\begin{aligned} \sigma(t)(v) &= \pi_{+}(t|_{v}) \\ &= \pi_{+}(\mathbb{T}\pi_{+}(t)|_{v}) \\ &= \pi_{+}(\operatorname{flat}(T|_{v})) \\ &= \pi_{+}(\operatorname{flat}(T)|_{\mu(v)}) = \sigma(\operatorname{flat}(T))(\mu(v)) \,. \end{aligned}$$

Finally, note that the mappings $\sigma \mapsto \pi_+$ and $\pi_+ \mapsto \sigma$ are clearly inverse to each other.

In particular, if labellings are unique, so is the expansion. In fact, we do not need to assume associativity here.

Proposition 6.6. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}$ and let \mathfrak{A} be a \mathbb{T}° -algebra satisfying at least one of *the following two conditions.*

- (I) Every tree $t \in \mathbb{T}A$ has a unique weak \mathbb{T}° -consistent labelling.
- (II) Every tree $t \in \mathbb{T}A$ has a unique strong \mathbb{T}° -consistent labelling.

Then \mathfrak{A} *has a unique* \mathbb{T} *-expansion.*

Proof. Let σ be the unique labelling scheme (weak or strong). By Proposition 6.5, it is sufficient to prove that σ is associative. Hence, fix a tree $T \in \mathbb{TT}A$ and let *t* and μ be as in the definition of associativity. We claim that the labelling

$$\tau \coloneqq \sigma(\operatorname{flat}(T)) \circ \mu$$

is a \mathbb{T}° -consistent labelling of *t*. Then uniqueness of labellings implies that

 $\sigma(t) = \tau = \sigma(\operatorname{flat}(T)) \circ \mu,$

as desired.

For the proof, fix a factor $[u, \bar{v})$ of *t*. We have to show that

$$\tau(u) = \pi(t[u,\bar{v})(\tau(v_{o}),\tau(v_{1}),\ldots))$$

Since every tree T(w) only contains finitely many variables, we can replace in T(w) some subtrees $T(w)|_{w'}$ (without variables) by the corresponding constant $\sigma(T(w))(w')$. Let P(w) be a finite tree obtained in this way from T(w). Using the consistency of $\sigma(P(w))$ and $\sigma(T(w))$, we can show by induction on w'(starting at the leaves) that

$$\sigma(P(w))(w') = \sigma(T(w))(w'), \text{ for all } w' \in \operatorname{dom}(P(w)).$$

Consequently,

$$\pi(P(w)) = \sigma(P(w))(\langle \rangle) = \sigma(T(w))(\langle \rangle) = t(w).$$

We regard the family $(P(w))_w$ as a tree $P \in \mathbb{TT}A$ with domain

 $\operatorname{dom}(P) = [u, \bar{v}) \cup \{v_0, v_1, \dots\},\$

where, for the leaves v_i , we choose

$$P(v_i) \coloneqq \tau(v_i)$$

Then the domain of *P* is that of a tree in \mathbb{T}° . Hence, we have $P \in \mathbb{T}^{\circ}\mathbb{T}^{\operatorname{fin}}A \subseteq \mathbb{T}^{\circ}\mathbb{T}^{\circ}A$. This implies that $\operatorname{flat}(P) \in \mathbb{T}^{\circ}A$ and

$$\sigma(\operatorname{flat}(P))(\langle \rangle) = \pi(\operatorname{flat}(P)) = \pi(\mathbb{T}^{\circ}\pi(P))$$
$$= \pi(t[u,\bar{v})(\tau(v_{\circ}),\tau(v_{1}),\ldots))),$$

where the last step follows by the fact that

$$\pi(P(w)) = \begin{cases} t(w) & \text{if } w \in [u, \bar{v}), \\ \tau(v_i) & \text{if } w = v_i. \end{cases}$$

Hence,

$$\pi(t[u,\bar{v})(\tau(v_{o}),\tau(v_{1}),\ldots))$$

= $\sigma(\operatorname{flat}(P))(\langle\rangle)$
= $\sigma(\operatorname{flat}(T))(\mu(u))$
= $\tau(u)$,

where the second step follows by the uniqueness of labellings.

As an application let us show how to use consistent labellings to prove that an algebra is definable.

Proposition 6.7. *Every finitary* \mathbb{T}^{thin} *-algebra is* MSO-*definable.*

Proof. Let \mathfrak{A} be a finitary \mathbb{T}^{thin} -algebra and $t \in \mathbb{T}^{\text{thin}}A$. We construct a formula guessing the labelling $\lambda : \text{dom}(t) \to A$ induced by the product $\lambda(v) := \pi(t|_v)$ and then verifying the correctness of its guess by checking for each vertex v that

- $\lambda(v) = t(v)(\lambda(u_0), \dots, \lambda(u_{n-1}))$, where u_0, \dots, u_{n-1} are the successors of v,
- $\lambda(v) = \pi(s_{\beta})$, for every branch β starting at v, where s_{β} is the path obtained from $t|_{v}$ by replacing every vertex u not belonging to β by the constant $\lambda(u)$.

The latter condition can be expressed in MSO since this logic can evaluate products in ω -semigroups. By induction on the Cantor-Bendixson rank of t it follows that the above checks ensure that the guessed labelling coincides with the intended one.

Corollary 6.8. Let \mathfrak{A} be a finitary \mathbb{T} -algebra where every tree $t \in \mathbb{T}A$ has exactly one \mathbb{T}^{thin} -consistent labelling. Then \mathfrak{A} is MSO-definable.

Proof. It follows by Proposition 6.7 that being \mathbb{T}^{thin} -consistent can be expressed in MSO. To evaluate a given product $\pi(t)$ in MSO, we can therefore guess a \mathbb{T}^{thin} -consistent labelling of t and take the label at the root.

7 UNAMBIGUOUS ALGEBRAS

In Corollary 5.11, we have obtained a complete classification of all \mathbb{T}^{thin} -expansions of a \mathbb{T}^{fin} -algebra. In the present section, we use consistent labellings to study the inclusion $\mathbb{T}^{\text{thin}} \subseteq \mathbb{T}$. First let us remark that it is not dense.

Lemma 7.1. There exists a \mathbb{T}^{thin} -algebra \mathfrak{A} with two different MSO-definable \mathbb{T} -expansions.

Proof. Let *A* be the set with two elements o_{ξ} , i_{ξ} for every sort ξ . We consider two different products π_0 , $\pi_1 : \mathbb{T}A \to A$ on this set. The first one is just the maximum operation:

$$\pi_{o}(t) \coloneqq \max \{ t(v) \mid v \in \operatorname{dom}(t) \}, \quad \text{for } t \in \mathbb{T}A.$$

The second one is given by

$$\pi_1(t) := \begin{cases} o & \text{if } t \in \mathbb{T}^{\text{thin}} C, \\ 1 & \text{if } t \in \mathbb{T} A \smallsetminus \mathbb{T}^{\text{thin}} C, \end{cases}$$

where $C \subseteq A$ is the subset consisting of the elements o_{ξ} , $\xi \in \Xi$. Since there exists an MSO-formula expressing that a given tree is thin, both products are MSOdefinable. Furthermore, π_0 is clearly associative. To show that so is π_1 , fix a tree $t \in \mathbb{TT}A$. We distinguish three cases.

- If there is some $v \in \text{dom}(t)$ with $t(v) \in \mathbb{T}A \setminus \mathbb{T}^{\text{thin}}C$, we have $\text{flat}(t) \notin \mathbb{T}^{\text{thin}}C$ and $\pi_1(\mathbb{T}\pi_1(t)) = 1 = \pi_1(\text{flat}(t))$.
- If $t \in \mathbb{T}^{\text{thin}}\mathbb{T}^{\text{thin}}C$, then $\text{flat}(t) \in \mathbb{T}^{\text{thin}}C$ and $\pi_1(\mathbb{T}\pi_1(t)) = 0 = \pi_1(\text{flat}(t))$.
- Finally, suppose that $t \in \mathbb{TT}^{\text{thin}}C \setminus \mathbb{T}^{\text{thin}}C$. Then $\text{flat}(t) \in \mathbb{T}C \setminus \mathbb{T}^{\text{thin}}C$ and $\pi_1(\mathbb{T}\pi_1(t)) = 1 = \pi_1(\text{flat}(t))$.

Consistent labellings have been used in [2] to study unambiguous tree languages. Let us give a brief overview over these results. The central notion is the following one.

Definition 7.2. Let $\mathbb{T}^{\circ} \subseteq \mathbb{T}$. A \mathbb{T}° -algebra \mathfrak{A} is *unambiguous* if every tree $t \in \mathbb{T}A$ has at most one strongly \mathbb{T}° -consistent labelling.

Remarks. (a) For $\mathbb{T}^{\circ} = \mathbb{T}^{\text{thin}}$ these algebras were introduced in [2] under the name *prophetic thin algebras.*

(b) The fact that a given tree has a unique strongly \mathbb{T}^{thin} -consistent labelling is expressible in MSO. And so are the facts that every tree has a unique \mathbb{T}^{thin} -consistent labelling and that the corresponding labelling scheme is associative.

First, note that there exist $\mathbb{T}^{\mathrm{thin}}\text{-algebras}$ which are not unambiguous.

Example. Let \mathfrak{A} be the \mathbb{T}^{thin} -algebra generated by the elements 0, 1 (of arity 0), b_0, b_1, c_0, c_1 (of arity 1), and *a* (of arity 2) subject to the following equations.

$$\begin{aligned} b_i(j) &= j, & a(x,i) = b_i(x), \\ b_i b_j(x) &= b_{\max\{i,j\}}(x), & a(i,x) = c_i(x), \\ c_i(j) &= i, & b_i^{\omega} = 1 - i, \\ c_i(b_j(x)) &= c_i(x), & c_i^{\omega} = i, \\ c_i(c_j(x)) &= c_i(x), \end{aligned}$$

for $i, j \in \{0, 1\}$. This algebra is not unambiguous since the (unique) tree $t \in \mathbb{T}_{\emptyset}\{a\}$ has several consistent labellings, including

 $\lambda(w) \coloneqq |w|_1 \mod 2$ and $\mu(w) \coloneqq (|w|_1 + 1) \mod 2$,

where $|w|_1$ denotes the number of letters 1 in $w \in \{0, 1\}^*$.

The connection between unambiguous \mathbb{T}^{thin} -algebras and unambiguous tree languages is given by the following theorem.

Definition 7.3. (a) A tree automaton is *unambiguous* if it has at most one accepting run on each given input tree.

(b) A language $K \subseteq \mathbb{T}_{\xi}\Sigma$ is called *bi-unambiguous* if both K and $\mathbb{T}_{\xi}\Sigma \setminus K$ are recognised by unambiguous automata.

Theorem 7.4 (Bilkowski, Skrzypczak [2]). A language $K \subseteq \mathbb{T}_{\xi}\Sigma$ is bi-unambiguous *if, and only if, it is recognised by a* \mathbb{T}^{thin} *-morphism* $\varphi : \mathbb{T}\Sigma \to \mathfrak{A}$ *to a finitary unambiguous* \mathbb{T}^{thin} *-algebra.*

Unfortunately, the question of whether all trees have strong \mathbb{T}^{thin} -consistent labellings is still an open problem, one which turns out to be equivalent to the existence of the following kind of choice functions.

Definition 7.5. The *Thin Choice Conjecture* states that there does *not* exist an MSO-formula $\varphi(x; Z)$ such that, for every thin (unlabelled) tree *t* and every non-empty set $P \subseteq \text{dom}(t)$ of parameters, the formula $\varphi(x; P)$ defines a unique element of *P*.

Theorem 7.6 (Bilkowski, Skrzypczak [2]). The following statements are equivalent.

- (1) The Thin Choice Conjecture holds.
- (2) All trees have strongly T^{thin}-consistent labellings, for every finitary T^{thin}algebra A.
- (3) The unique tree in \mathbb{T}_{\emptyset} {*a*} has a strongly \mathbb{T}^{thin} -consistent labelling, for every \mathbb{T}^{thin} -algebra \mathfrak{A} and every $a \in A$.
- (4) For every morphism φ : A → B of T^{thin}-algebras and every strongly T^{thin}-consistent labelling β of some tree t ∈ TB, there exists a strongly T^{thin}-consistent labelling α with φ ∘ α = β.

Theorem 7.7 (Bilkowski, Skrzypczak [2]). *Suppose that the Thin Choice Conjecture holds.*

- (a) Every finitary unambiguous \mathbb{T}^{thin} -algebra \mathfrak{A} has a unique \mathbb{T} -extension. Furthermore, this extension is MSO-definable.
- (b) The class of unambiguous \mathbb{T}^{thin} -algebras forms a pseudo-variety.
- (c) A language $K \subseteq \mathbb{T}\Sigma$ is bi-unambiguous if, and only if, (the \mathbb{T}^{thin} -reduct of) its syntactic algebra Syn(K) is unambiguous.

8 BRANCH-CONTINUOUS ALGEBRAS

In this final section, we take a look at a few other natural classes of \mathbb{T}^{thin} -algebras where unique \mathbb{T} -expansions exist. The simplest example consists of algebras that are constructed from an ω -semigroup as follows.

Definition 8.1. (a) Let $\mathfrak{S} = \langle S, S_{\omega} \rangle$ be an ω -semigroup. We denote by TA(\mathfrak{S}) the \mathbb{T} -algebra $\langle A, \pi \rangle$ with domains

$$A_{\xi} \coloneqq S_{\omega} + S \times \xi, \quad \text{for } \xi \in \Xi.$$

For elements $\langle a, x \rangle \in S \times \xi$, we will use the more suggestive notation a(x). The product is defined as follows. Given $t \in \mathbb{T}_{\xi}A$, let $\beta = (v_i)_i$ be the path defined as follows. We start with the root v_0 of t. Having chosen v_i , we take a look at its label $t(v_i)$. If $t(v_i) = a_i(z_i) \in S \times \zeta_i$, we choose $v_{i+1} := \operatorname{suc}_{z_i}(v_i)$. Otherwise, the path ends at v_i . Let $(a_i)_i$ be the corresponding sequence of labels. (If the path is finite, the last label a_n is either an element of S_{ω} or a variable.) We set

$$\pi(t) \coloneqq \prod_i a_i \, .$$

Note that this product can be of one the following forms:

- an infinite product $a_0 \cdot a_1 \cdots \in S_\omega$ with $a_i \in S$,
- a finite product $a_0 \cdots a_n \in S_\omega$ with $a_0, \ldots, a_{n-1} \in S$ and $a_n \in S_\omega$,
- a finite product $\langle a_0 \cdots a_{n-1}, a_n \rangle \in S \times \xi$ with $a_0, \ldots, a_{n-1} \in S$ and $a_n \in \xi$ is a variable.

(b) A \mathbb{T} -algebra \mathfrak{A} is *semigroup-like* if it is isomorphic to TA(\mathfrak{S}), for some ω -semigroup \mathfrak{S} . Similarly, for a submonad $\mathbb{T}^{\circ} \subseteq \mathbb{T}$, we call a \mathbb{T}° -algebra \mathfrak{A} *semigroup-like* if it is the \mathbb{T}° -reduct of a semigroup-like \mathbb{T} -algebra.

Lemma 8.2. Every semigroup-like \mathbb{T}^{thin} -algebra is unambiguous and has a unique \mathbb{T} -expansion. This expansion is again semigroup-like.

Proof. Let \mathfrak{A} be a semigroup-like \mathbb{T}^{thin} -algebra and let \mathfrak{S} be the ω -semigroup such that $\mathfrak{A} \cong \text{TA}(\mathfrak{S})|_{\mathbb{T}^{\text{thin}}}$. First, note that \mathfrak{A} has a semigroup-like \mathbb{T} -expansion: the algebra $\text{TA}(\mathfrak{S})$. Hence, we can use Proposition 6.5 (b) to find a strong labelling scheme for \mathfrak{A} . It therefore remains to prove that this labelling scheme is unique. Fix a tree $t \in \mathbb{T}A$ and let λ and μ be two strongly \mathbb{T}^{thin} -consistent labellings of t. Given a vertex $v \in \text{dom}(t)$, let β be the path starting at v that we constructed in the definition of $\pi_+(t|_v)$. We choose a thin factor p of $t|_v$ containing this path. Then \mathbb{T}^{thin} -consistency implies that $\lambda(v) = \pi(p) = \mu(v)$. Hence, $\lambda = \mu$.

This lemma is hardly surprising, since the product of a semigroup-like algebra only depends on a single branch of the given tree. We can extend this result to more complicated classes of algebras as follows. So far, we have mostly ignored the fact that our algebras are ordered. (The ordering is needed to characterise logics that are not closed under negation, something we are not concerned with in the present article.) The next two classes of examples on the other hand make essential use of the ordering. We start by introducing some notation concerning meets and joins.

Definition 8.3. Let *A* be a sorted set.

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

 $\Uparrow X := \{ a \in A \mid a \ge x \text{ for some } x \in X \},$ $\downarrow X := \{ a \in A \mid a \le x \text{ for some } x \in X \}.$

The set *X* is *upwards closed* if $\uparrow X = X$, and it is *downwards closed* if $\downarrow X = X$.

(b) We define two functors \mathbb{U} and \mathbb{D} as follows. For sorted sets *A*, we set

 $\mathbb{U}A \coloneqq \{ I \subseteq A \mid I \text{ is upwards closed } \},$ $\mathbb{D}A \coloneqq \{ I \subseteq A \mid I \text{ is downwards closed } \}.$

For functions $f : A \rightarrow B$, we define

$$\mathbb{U}f(I) \coloneqq \{ f(a) \mid a \in I \},\\ \mathbb{D}f(I) \coloneqq \{ f(a) \mid a \in I \}.$$

(c) For $t \in \mathbb{T}A$ and $T \in \mathbb{T}UA$ or $T \in \mathbb{T}DA$, we write

 $t \in^{\mathbb{T}} T$: iff t and T have the same domain and if $t(v) \in T(v)$, for all vertices v.

(d) Let $C \subseteq A$. We denote by $\langle \! \langle C \rangle \! \rangle_{inf}$ the closure of *C* under arbitrary meets and by $\langle \! \langle C \rangle \! \rangle_{sup}$ its closure under arbitrary joins. *C* is a set of *meet-generators* if $\langle \! \langle C \rangle \! \rangle_{inf} = A$ and a set of *join-generators* if $\langle \! \langle C \rangle \! \rangle_{sup} = A$.

The next, more interesting class of algebras we take a look at is the class of *deterministic* algebras, which was introduced in [7] to give an algebraic characterisation of the class of MSO-definable \mathbb{T} -algebras. Here, we are interested in the fact that their product is determined by its \mathbb{T}^{thin} -reduct. The definition is as follows.

Definition 8.4. (a) We define a function dist : $\mathbb{TU}A \rightarrow \mathbb{UT}A$ by

dist $(t) := \{ s \in \mathbb{T}A \mid s \in^{\mathbb{T}} t \}.$

(b) A function $g : \mathfrak{A} \to B$ from a \mathbb{T} -algebra \mathfrak{A} to a completely ordered sorted set *B* is *meet-distributive* if *g* preserves meets and there exists a function $\sigma : \mathbb{T}\langle\langle \operatorname{rng} g \rangle\rangle_{\inf} \to B$ such that

$$\sigma \circ \mathbb{T}(\inf \circ \mathbb{U}g) = g \circ \inf \circ \mathbb{U}\pi \circ \text{dist}.$$

A completely ordered \mathbb{T} -algebra \mathfrak{A} is *meet-distributive* if the identity $id : \mathfrak{A} \to \mathfrak{A}$ is meet-distributive. *Join-distributivity* is defined analogously with \mathbb{D} and sup instead of \mathbb{U} and inf.

(c) A \mathbb{T}^{thin} -algebra \mathfrak{A} is *deterministic* if it is meet-distributive and it has a semigroup-like subalgebra \mathfrak{C} such that *C* forms a set of meet-generators of \mathfrak{A} and the inclusion $C \to A$ is meet-distributive. Meet-distributive algebras are those where the product commutes with meets. More generally, an embedding $e : \mathfrak{A} \to \mathfrak{B}$ of \mathbb{T} -algebras is meet-distributive if the restriction of \mathfrak{B} to rng e is meet-distributive. The name 'deterministic algebra' stems from the fact that such algebras correspond to deterministic automata. A typical example of a deterministic algebra is one where every element is of the form

$$a_{\circ}(x_{\circ}) \sqcap \cdots \sqcap a_{m-1}(x_{m-1}) \sqcap b_{\circ} \sqcap \cdots \sqcap b_{n-1}$$

where $a_0, \ldots, a_{m-1} \in S$ and $b_0, \ldots, b_{n-1} \in S_{\omega}$ are elements of some ω -semigroup $\mathfrak{S} = \langle S, S_{\omega} \rangle$ and the x_0, \ldots, x_{m-1} are variables.

We start with two technical lemmas. The first one is trivial.

Lemma 8.5. A completely ordered \mathbb{T} -algebra \mathfrak{A} is meet-distributive if, and only if,

 $\pi \circ \mathbb{T}$ inf = inf $\circ \mathbb{U}\pi \circ \text{dist}$.

Proof. (\Leftarrow) In the definition of meet-distributivity, we can take $\sigma \coloneqq \pi$.

(⇒) Let σ be the function from the definition of meet-distributivity. Given a tree $t \in \mathbb{T}A$, let $T \in \mathbb{T}\mathbb{U}A$ be the tree with labels $T(v) = \{ a \in A \mid a \ge t(v) \}$. Then

$$\sigma(t) = \sigma(\mathbb{T}\inf(T)) = \inf \{ \pi(s) \mid s \in \mathbb{T} T \} = \inf \{ \pi(t) \} = \pi(t) .$$

Hence, we have

$$\pi \circ \mathbb{T}\inf = \sigma \circ \mathbb{T}(\inf \circ \mathbb{U}id)$$
$$= id \circ \inf \circ \mathbb{U}\pi \circ dist = \inf \circ \mathbb{U}\pi \circ dist.$$

Meet-distributive functions can be used to transfer a \mathbb{T} -algebra product from their domain to their codomain.

Lemma 8.6. Let $\varphi : \mathfrak{C} \to A$ be a meet-distributive function such that $\operatorname{rng} \varphi$ is a set of meet-generators of A. There exists a unique function $\sigma : \mathbb{T}A \to A$ such that (A, σ) is a meet-distributive \mathbb{T} -algebra and φ a morphism of \mathbb{T} -algebras.

Proof. To make our proof more concise, we use some properties of the function dist : $\mathbb{TU} \Rightarrow \mathbb{UT}$. We have shown in [9] that dist is what is called a *distributive law*, which means it is a natural transformation satisfying the equations

dist \circ flat = Uflat \circ dist \circ Tdist,	dist \circ sing = Using,
dist $\circ \mathbb{T}$ union = union $\circ \mathbb{U}$ dist \circ dist ,	dist $\circ \mathbb{T}$ pt = pt,

where union : $\mathbb{UU}A \to \mathbb{U}A$ maps a set of sets to its union and pt : $A \to \mathbb{U}A$ is defined by pt $(a) := \Uparrow\{a\}$.

Let $\sigma : \mathbb{T}A \to A$ be the function from the definition of meet-distributivity. To see that $\langle A, \sigma \rangle$ is a \mathbb{T} -algebra, note that

 $\sigma \circ \operatorname{sing} \circ (\operatorname{inf} \circ \mathbb{U}\varphi)$ $= \sigma \circ \mathbb{T}(\inf \circ \operatorname{pt}) \circ \operatorname{sing} \circ (\inf \circ \mathbb{U}\varphi)$ $[inf \circ pt = id]$ $= \sigma \circ \mathbb{T}(\inf \circ \operatorname{pt}) \circ \mathbb{T}(\inf \circ \mathbb{U}\varphi) \circ \operatorname{sing}$ [sing nat. trans.] $= \sigma \circ \mathbb{T}inf \circ \mathbb{T}\mathbb{U}(inf \circ \mathbb{U}\varphi) \circ \mathbb{T}pt \circ sing$ [pt nat. trans.] $= \sigma \circ \mathbb{T}inf \circ \mathbb{T}(union \circ \mathbb{U}\mathbb{U}\varphi) \circ \mathbb{T}pt \circ sing$ $[\inf \circ \mathbb{U}\inf = \inf \circ \operatorname{union}]$ $= \sigma \circ \mathbb{T}inf \circ \mathbb{T}(\mathbb{U}\varphi \circ union) \circ \mathbb{T}pt \circ sing$ [union nat. trans.] $= \varphi \circ \inf \circ \mathbb{U}\pi \circ \operatorname{dist} \circ \mathbb{T} \operatorname{union} \circ \mathbb{T} \operatorname{pt} \circ \operatorname{sing}$ $[\varphi \text{ meet-dist.}]$ $= \varphi \circ \inf \circ \mathbb{U}\pi \circ \operatorname{dist} \circ \operatorname{sing}$ $[union \circ pt = id]$ $= \varphi \circ \inf \circ \mathbb{U}\pi \circ \mathbb{U}sing$ [dist dist. law] $= \varphi \circ \inf$ [unit law for π] $= \inf \circ \mathbb{U}\varphi$, $[\varphi \text{ meet-dist.}]$ $\sigma \circ \mathbb{T} \sigma \circ \mathbb{T} \mathbb{T} (\inf \circ \mathbb{U} \varphi)$ $= \sigma \circ \mathbb{T}(\varphi \circ \inf \circ \mathbb{U}\pi \circ \operatorname{dist})$ $[\varphi \text{ meet-dist.}]$ $= \sigma \circ \mathbb{T}(\inf \circ \mathbb{U}\varphi) \circ \mathbb{T}\mathbb{U}\pi \circ \mathbb{T}$ dist [inf nat. trans.] $= \varphi \circ \inf \circ \mathbb{U}\pi \circ \operatorname{dist} \circ \mathbb{T}\mathbb{U}\pi \circ \mathbb{T}\operatorname{dist}$ $[\varphi \text{ meet-dist.}]$ $= \varphi \circ \inf \circ \mathbb{U}\pi \circ \mathbb{U}\mathbb{T}\pi \circ \operatorname{dist} \circ \mathbb{T}\operatorname{dist}$ [dist nat. trans.] $= \varphi \circ \inf \circ \mathbb{U}\pi \circ \mathbb{U}$ flat $\circ dist \circ \mathbb{T}$ dist $[\mathfrak{C} \mathbb{T}$ -algebra] $= \varphi \circ \inf \circ \mathbb{U}\pi \circ \operatorname{dist} \circ \operatorname{flat}$ [dist dist. law] $= \sigma \circ \mathbb{T}inf \circ \mathbb{T}\mathbb{U}\varphi \circ flat$ $[\varphi \text{ meet-dist.}]$ $= \sigma \circ \text{flat} \circ \mathbb{TT}(\inf \circ \mathbb{U}\varphi)$. [flat nat. trans.]

Since $\inf \circ \mathbb{U}\varphi$ is surjective and \mathbb{T} preserves surjectivity, it follows that

 $\sigma \circ \operatorname{sing} = \operatorname{id} \quad \operatorname{and} \quad \sigma \circ \mathbb{T} \sigma = \sigma \circ \operatorname{flat}.$

To see that $\langle A, \sigma \rangle$ is meet-distributive, note that

$\sigma \circ \mathbb{T}\mathrm{inf} \circ \mathbb{T}\mathbb{U}(\mathrm{inf} \circ \mathbb{U}\varphi)$	
$= \sigma \circ \mathbb{T}(\inf \circ \operatorname{union} \circ \mathbb{UU}\varphi)$	$[\inf \circ \mathbb{U}inf = \inf \circ union]$
$= \sigma \circ \mathbb{T}(\inf \circ \mathbb{U}\varphi \circ \text{union})$	[union nat. trans.]
$= \varphi \circ \inf \circ \mathbb{U}\pi \circ \operatorname{dist} \circ \mathbb{T}$ union	$[\varphi \text{ meet-dist.}]$
$= \varphi \circ \inf \circ \mathbb{U}\pi \circ \operatorname{union} \circ \mathbb{U}$ dist \circ dist	[dist dist. law]
$= \varphi \circ \inf \circ union \circ \mathbb{U}\mathbb{U}\pi \circ \mathbb{U}dist \circ dist$	[union nat. trans.]
$= \varphi \circ \inf \circ \mathbb{U} \inf \circ \mathbb{U} \mathbb{U} \pi \circ \mathbb{U} \operatorname{dist} \circ \operatorname{dist}$	$[\inf \circ \mathbb{U}inf = \inf \circ union]$
$= \inf \circ \mathbb{U}(\varphi \circ \inf \circ \mathbb{U}\pi \circ \operatorname{dist}) \circ \operatorname{dist}$	$[\varphi \text{ meet-dist.}]$
$= \inf \circ \mathbb{U}(\sigma \circ \mathbb{T}(\inf \circ \mathbb{U}\varphi)) \circ \text{dist}$	$[\varphi \text{ meet-dist.}]$
$= \inf \circ \mathbb{U}\sigma \circ \operatorname{dist} \circ \mathbb{TU}(\inf \circ \mathbb{U}\varphi) .$	[dist nat. trans.]

By surjectivity of $\mathbb{TU}(\inf \circ \mathbb{U}\varphi)$, this implies that

 $\sigma \circ \mathbb{T}$ inf = inf $\circ \mathbb{U}\sigma \circ \text{dist}$.

Hence, the claim follows by Lemma 8.5.

To see that φ is a morphism of \mathbb{T} -algebras, note that

$\sigma \circ \mathbb{T} \varphi = \sigma \circ \mathbb{T} (\inf \circ \mathrm{pt} \circ \varphi)$	$[\inf \circ pt = id]$
$= \sigma \circ \mathbb{T}(\inf \circ \mathbb{U}\varphi \circ \mathrm{pt})$	[pt nat. trans.]
$= \varphi \circ \inf \circ \mathbb{U}\pi \circ \operatorname{dist} \circ \mathbb{T}\mathrm{pt}$	$[\varphi \text{ meet-dist.}]$
$= \varphi \circ \inf \circ \mathbb{U}\pi \circ \mathrm{pt}$	[dist dist. law]
$= \varphi \circ \inf \circ \operatorname{pt} \circ \pi$	[pt nat. trans.]
$= \varphi \circ \pi$.	$[\inf \circ pt = id]$

Finally, for uniqueness suppose that $\sigma': \mathbb{T}A \to A$ is another function like this. Then it follows by Lemma 8.5 that

 $\sigma \circ \mathbb{T}(\inf \circ \mathbb{U}\varphi) = (\inf \circ \mathbb{U}\pi \circ \operatorname{dist}) \circ \mathbb{T}\mathbb{U}\varphi = \sigma' \circ \mathbb{T}(\inf \circ \mathbb{U}\varphi).$

Hence, the fact that $\mathbb{T}(\inf \circ \mathbb{U}\varphi)$ is surjective implies that $\sigma = \sigma'$.

Theorem 8.7. Every deterministic \mathbb{T}^{thin} -algebra has a unique meet-distributive \mathbb{T} -expansion.

Proof. Let $\mathfrak{A} = \langle A, \pi \rangle$ be a deterministic \mathbb{T}^{thin} -algebra and let $\mathfrak{C} \subseteq \mathfrak{A}$ be the corresponding semigroup-like subalgebra. We can use Lemma 8.2 to find a unique \mathbb{T} -expansion \mathfrak{C}_+ of \mathfrak{C} , and Lemma 8.6 to find a unique meet-distributive algebra $\mathfrak{A}_+ = \langle A, \pi_+ \rangle$ with universe *A* that contains \mathfrak{C}_+ as a subalgebra.

It therefore remains to prove that \mathfrak{A} is the \mathbb{T}^{thin} -reduct of this algebra \mathfrak{A}_+ . Hence, let $t \in \mathbb{T}^{\text{thin}}A$ and fix a tree $T \in \mathbb{T}^{\text{thin}}\mathbb{U}C$ such that $t = \mathbb{T}^{\text{thin}}\inf(T)$. By meet-distributivity and the fact that the products π and π_+ agree on trees in $\mathbb{T}^{\text{thin}}C$, it follows that

$$\pi_{+}(t) = \inf \{ \pi_{+}(s) \mid s \in^{\mathbb{T}} T \} = \inf \{ \pi(s) \mid s \in^{\mathbb{T}} T \} = \pi(t),$$

as desired.

Corollary 8.8. Every deterministic \mathbb{T} -algebra is uniquely determined by its \mathbb{T}^{thin} -reduct.

 \square

We can generalise deterministic algebras by also allowing joins. The resulting algebras are called *branch-continuous*. They were introduced in [4] as an algebraic analogue to tree automata.

Definition 8.9. A \mathbb{T}^{thin} -algebra \mathfrak{A} is *branch-continuous* if it is join-distributive and it has a deterministic subalgebra \mathfrak{C} such that *C* forms a set of join-generators of \mathfrak{A} and the inclusion $C \to A$ is join-distributive.

Example. A typical example of a branch-continuous algebra consists of an algebra of (profiles of) certain games. A *regular game* is played on a directed graph the form $\mathfrak{G} = \langle V_{\rm I}, V_{\rm II}, E, \lambda, \mu, \nu_{\rm o} \rangle$ where the set $V = V_{\rm I} + V_{\rm II}$ of vertices is divided into two parts, one for each player, and the vertices and edges are labelled by elements of some finite ω -semigroup $\mathfrak{S} = \langle S, S_{\omega} \rangle$. This labelling is given by the functions

$$\lambda: E \to S$$
 and $\mu: V \to \mathcal{P}(S_{\omega})$.

We assume that $\mu(\nu) \neq \emptyset$, for every leaf ν . The vertex $\nu_0 \in V$ denotes the initial position of the game.

Since we want to compose games we allow some of the leaves of \mathfrak{G} to be labelled by (distinct) variables. For such leaves v, we assume that $\mu(v) = \emptyset$.

The game is played between two players, Player I and Player II and proceeds as follows. It starts in the position v_0 . If the game has reached some position $v \in V$, the player to whom v belongs chooses either some semigroup element $c \in \mu(v)$

or some outgoing edge $v \rightarrow u$. In the first case the game terminates, otherwise it continues in position u. It follows that each play of the game produces a finite or infinite path starting at the root. The labelling of this path is of one of the following forms.

$$a_0a_1a_2\cdots, a_0\cdots a_{n-1}c, a_0\cdots a_{n-1}x,$$

where $a_i \in S$, $c \in S_\omega$, and x is one of the variables. Each such sequence can be multiplied to either an element of S_ω or an element of the form a(x) with $a \in S$ and x a variable. We call this product the *outcome* of the play.

The set of all games (over some fixed ω -semigroup \mathfrak{S}) forms a \mathbb{T} -algebra. But here we are interested in a finitary quotient of this algebra. When we are only interested in the possible outcomes of a game and not in its game graph, we can represent each game \mathfrak{G} as a term of the form

$$\bigsqcup_{\sigma_{\rm I}} \prod_{\sigma_{\rm II}} a_{\sigma_{\rm I},\sigma_{\rm II}}$$

where σ_{I} and σ_{II} range over all strategies for the respective player and $a_{\sigma_{I},\sigma_{II}}$ is the outcome of the game when both players play according to the indicated strategies. Let us call such a term the *profile* of the game \mathfrak{G} .

The set of all possible profiles (again, for some fixed ω -semigroup \mathfrak{S}), forms a \mathbb{T} -algebra \mathfrak{A} which is branch-continuous. The elements of the form $a_{\sigma_{\mathrm{I}},\sigma_{\mathrm{II}}}$ form a semigroup-like subalgebra, those of the form $\prod_{\sigma_{\mathrm{II}}} a_{\sigma_{\mathrm{I}},\sigma_{\mathrm{II}}}$ form a deterministic subalgebra, and every element of *A* is a join of such elements.

Using join-distributivity and meet-distributivity, one can show that a product $\pi(t)$ in a branch-continuous algebra can be computed by taking a join over meets over products along single branches of t (see [4] for details). In particular, a product of this form is MSO-definable. Together with the translation of automata into branch-continuous T-algebras, this leads to the following two results from [4].

Proposition 8.10. *Every finitary branch-continuous* **T***-algebra is* MSO-*definable.*

Theorem 8.11. A language $K \subseteq \mathbb{T}\Sigma$ is regular if, and only if, it is recognised by a morphism into a finitary branch-continuous \mathbb{T} -algebra.

Thus branch-continuous algebras play a similar role as the MSO-definable ones. The reason we usually work with the latter is that the former do not form a pseudo-variety: the class of branch-continuous algebras is not closed under finitely-generated subalgebras. Here, we are more interested in the fact that branchcontinuous algebras have unique expansions.

Theorem 8.12. Every branch-continuous \mathbb{T}^{thin} -algebra has a unique join-distributive \mathbb{T} -expansion.

Proof. Let \mathfrak{A} be a branch-continuous \mathbb{T}^{thin} -algebra and let \mathfrak{C} be the corresponding deterministic subalgebra. By Theorem 8.7, \mathfrak{C} has a unique meet-distributive \mathbb{T} -expansion \mathfrak{C}_+ . By the dual version of Lemma 8.6, there exist a unique joindistributive \mathbb{T} -algebra \mathfrak{A}_+ extending \mathfrak{C}_+ . By the same argument as in the proof of Theorem 8.7, it follows that \mathfrak{A}^+ is a \mathbb{T} -expansion of \mathfrak{A} .

Corollary 8.13. *Every branch-continuous* \mathbb{T} *-algebra is uniquely determined by its* \mathbb{T}^{thin} *-reduct.*

9 CONCLUSION

We have presented several approaches to the expansion problem for tree algebras. In each cases, we could use the existing combinatorial theory for ω -semigroups to solve the problem for \mathbb{T}^{thin} -expansions, but we always hit a wall when considering the problem for expansions to non-thin trees.

In particular, the methods we developed seem to work well if there exists a unique expansion (or at least a unique expansion with a certain property, like a unique MSO-definable expansion, or a unique branch-continuous one), but there is currently no approach to prove the existence of several expansions.

Promising next steps towards further progress seem to include

- trying to generalise some of our existing tools to general trees; and/or
- finding counterexamples delineating the parameter space where such generalisations do not exist any more.

As a problem to work on, let us mention a proof Simon's Factorisation-Tree Theorem to trees. But this seems to be a very hard problem. There are two technical frameworks that might be of help here: (I) one can try to flesh out the theory of Green's relations for tree algebras, and (II) one can try to make use of Tame Congruence Theory [19]. While developing these two theories for tree algebras is not that difficult, it is not obvious how to apply them to concrete problems, like the one mentioned above.

References

- P. A. ABDULLA, P. KRCÁL, AND W. YI, *R-Automata*, in CONCUR 2008 Concurrency Theory, 19th International Conference, CONCUR 2008, Toronto, Canada, August 19–22, 2008. Proceedings, F. van Breugel and M. Chechik, eds., vol. 5201 of Lecture Notes in Computer Science, Springer, 2008, pp. 67–81.
- [2] M. BILKOWSKI AND M. SKRZYPCZAK, Unambiguity and Uniformization Problems on Infinite Trees, in Computer Science Logic 2013 (CSL 2013), CSL 2013, September 2–5, 2013, Torino, Italy, vol. 23 of LIPIcs, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2013, pp. 81–100.
- [3] A. BLUMENSATH, Abstract Algebraic Language Theory. unpublished notes, https://www.fi.muni.cz/~blumens/ALT.pdf.
- [4] _____, Branch-Continuous Tree Algebras. arXiv:1807.04568, unpublished.
- [5] ——, Recognisability for Algebras of Infinite Trees, Theoretical Computer Science, 412 (2011), pp. 3463–3486.
- [6] —, An Algebraic Proof of Rabin's Tree Theorem, Theoretical Computer Science, 478 (2013), pp. 1–21.
- [7] ——, *Regular Tree Algebras*, Logical Methods in Computer Science, 16 (2020), pp. 16:1–16:25.
- [8] ——, Algebraic Language Theory for Eilenberg–Moore Algebras, Logical Methods in Computer Science, 17 (2021), pp. 6:1–6:60.
- [9] —, The Power-Set Operation for Tree Algebras, Logical Methods in Computer Science, 19 (2023), pp. 9:1–9:47.
- [10] 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.
- [11] M. BOJAŃCZYK, F. CAVALLARI, T. PLACE, AND M. SKRZYPCZAK, Regular Tree Languages in Low Levels of the Wadge Hierarchy, Log. Methods Comput. Sci., 15 (2019), pp. 27:1–27:61.
- [12] M. BOJAŃCZYK AND T. IDZIASZEK, Algebra for Infinite Forests with an Application to the Temporal Logic EF, in Proc. 20th International Conference on Concurrency Theory, CONCUR, LNCS 5710, 2009, pp. 131–145.
- [13] M. BOJAŃCZYK, T. IDZIASZEK, AND M. SKRZYPCZAK, Regular Languages of Thin Trees, in Proc. 30th International Symposium on Theoretical Aspects of Computer Science, STACS 2013, 2013, pp. 562–573.
- [14] M. BOJAŃCZYK AND B. KLIN, A Non-Regular Language of Infinite Trees That is Recognizable by a Finite Algebra, Logical Methods in Computer Science, 15 (2019).

- [15] J. R. BÜCHI AND L. H. LANDWEBER, Solving Sequential Conditions by Finite-State Strategies, Trans. Amer. Math. Soc., 139 (1969), pp. 367–378.
- [16] O. CARTON, T. COLCOMBET, AND G. PUPPIS, An Algebraic Approach to MSO-Definability on Countable linear Orderings, J. Symb. Log., 83 (2018), pp. 1147–1189.
- [17] T. COLCOMBET, A Combinatorial Theorem for Trees, in Proc. 34th Int. Colloquium on Automata, Languages and Programming, ICALP, 2007, pp. 901–912.
- [18] _____, The Factorisation Forest Theorem, in Handbook of Automata Theory, J.-É. Pin, ed., European Mathematical Society, 2021, pp. 653–693.
- [19] D. HOBBY AND R. MCKENZIE, *The Structure of Finite Algebras*, American Mathematical Society, 1988.
- [20] C. LÖDING, Automata on Infinite Trees, in Handbook of Automata Theory, J.-É. Pin, ed., European Mathematical Society, 2021, pp. 265–302.
- [21] K. R. MILLIKEN, A Ramsey Theorem for Trees, Journal of Combinatorial Theory, Series A, 26 (1979), pp. 215–237.
- [22] D. PERRIN AND J.-É. PIN, Infinite Words Automata, Semigroups, Logic and Games, Elsevier, 2004.
- [23] G. PUPPIS, Automata for Branching and Layered Temporal Structures: An Investigation into Regularities of Infinite Transition Systems, vol. 5955 of Lecture Notes in Computer Science, Springer, 2010.
- [24] I. SIMON, Factorization Forests of Finite Height, Theoretical Computer Science, 72 (1990), pp. 65–94.
- [25] ——, On Semigroups of Matrices over the Tropical Semiring, RAIRO Theor. Informatics Appl., 28 (1994), pp. 277–294.
- [26] W. THOMAS, *A Combinatorial Approach to the Theory of ω-Automata*, Information and Control, 48 (1981), pp. 261–283.
- [27] ——, *Automata on Infinite Objects*, in Handbook of Theoretical Computer Science, J. van Leeuwen, ed., vol. B, Elsevier, Amsterdam, 1990, pp. 135–191.
- [28] —, Languages, Automata, and Logic, in Handbook of Formal Languages,
 G. Rozenberg and A. Salomaa, eds., vol. 3, Springer, New York, 1997, pp. 389–455.