Recognisability for algebras of infinite trees

Achim Blumensath*
TU Darmstadt, Mathematik, AG Logik
Schloßgartenstraße 7
64289 Darmstadt
Germany
Phone: +49-6151-162215
Fax: +49-6151-163317
blumensath@mathematik.tu-darmstadt.de

28th February 2011

We develop an algebraic language theory for languages of infinite trees. We define a class of algebras called ω-hyperclones and we show that a language of infinite trees is regular if, and only if, it is recognised by a finitary path-continuous ω-hyperclone.

Keywords. infinite trees; recognisability; tree automata; monadic second-order logic

Instead of using finite automata to develop the theory of regular languages, one can also employ semigroup theory. By now this approach has a long tradition and there exists an extended structure theory connecting varieties of languages with finite semigroups. This theory is particularly effective if one is interested in characterising subclasses of the class of all regular languages. For instance, the only known decidable characterisation of the class of first-order definable languages is based on semigroup theory.

*Work supported by ESF project 'Automata: from Mathematics to Applications'.

Naturally, there have been attempts to generalise this theory to other notions of regularity. For languages of ω-words, such a generalisation has largely been achieved. A detailed account can be found in the book of Pin and Perrin [8]. There also have been several contributions to an algebraic theory for languages of finite trees [11, 13, 12, 9, 5, 10, 3, 1, 7]. But the resulting theory is still fragmentary with several competing approaches and formalisations. Our own work has been influenced in particular by the following two articles: Ésik and Weil [6] have developed an approach using *preclones,* while Bojańczyk and Walukiewicz [3] use *forest algebras.* As far as the algebraic setting is concerned, the formalisation in the present article most closely resembles the work on *clones* by Ésik [4].

So far, an algebraic theory for languages of *infinite trees* is still missing. The main obstacle is the lack of appropriate combinatorial tools, like Ramseyan factorisation theorems for infinite trees. In particular, a purely combinatorial proof that every nonempty regular language of infinite trees contains a regular tree is still missing. There is recent work of Bojańczyk and Idziaszek [2] on characterisation results for classes of infinite trees that manages to circumvent these problems by a technical trick: since every regular language of infinite trees is determined by the regular trees it contains, it is sufficient to consider only regular trees.

This paper provides a first step in the development of an algebraic theory for recognisability of classes of infinite trees. Inspired by the work of Ésik and Weil on preclones, we define suitable algebras of infinite trees called *ω-hyperclones.* We can show that every regular language is recognised by some homomorphism into such a (finitary, path-continuous) ω-hyperclone.

The proof is performed in two steps. First, we define a special class of ω-hyperclones called *path-hyperclones* that directly correspond to tree-automata. The problem with path-hyperclones is that their definition is not axiomatic, but syntactic. That is, given an arbitrary ω-hyperclone we cannot tell from the definition whether or not this ω-hyperclone is isomorphic to some path-hyperclone.

In the second step, we therefore give an algebraic characterisation of the main properties of such path-hyperclones (they are *path-continuous*). Using this result we can transfer our characterisation from path-hyperclones to path-continuous ω-hyperclones.

Finally, we prove that the class of path-continuous ω-hyperclones is closed under products and a certain power-set operation. From these results we can deduce a second (equivalent) version of our main theorem: recognisability by finitary path-continuous ω-hyperclones is the same as definability in monadic second-order logic.

Acknowledgements. I would like to thank Thomas Colcombet and Arnaud Carayol for the intensive discussions that started this work. In particular, the notion of a path-hyperclone is partially based on their ideas.

1 Preliminaries

Let us fix our notation and recall basic definitions. We set $[n] := \{0, \dots, n-1\}$. Frequently, we do not distinguish between a tuple $\bar{a} = \langle a_0, \dots, a_{n-1} \rangle$ and the set $\{a_0, \dots, a_{n-1}\}$ of its components. In particular, we sometimes write $\bar{a} \subseteq A$ instead of $\bar{a} \in A^n$. We denote the power set of a set A by $\wp(A)$.

Definition 1.1. (a) An *ω-semigroup* is a two-sorted structure $\mathfrak{S} = \langle S, S_\omega, \cdot, \alpha, \pi \rangle$ where

- $\langle S, \cdot \rangle$ forms a semigroup,
- $\alpha : S \times S_\omega \to S_\omega$ is a left action of S on S_ω, and
- $\pi : S^\omega \to S_\omega$ is a function satisfying the following associative law:

$$\pi\big((s_i)_{i<\omega}\big) = \alpha\big(s_0 \cdots s_{k_0-1}, \pi\big((s_{k_i} \cdots s_{k_{i+1}-1})_{i<\omega}\big)\big),$$

for every sequence $(s_i)_i \in S^\omega$ and all increasing sequences of indices $k_0 < k_1 < \cdots < \omega$.

Usually, we will omit α and $\cdot$ and write sw and st instead of, respectively, $\alpha(s, w)$ and $s \cdot t$.

(b) A *morphism* $\mu : \mathfrak{S} \to \mathfrak{T}$ of ω-semigroups is a pair $\mu = \langle \mu_0, \mu_\omega \rangle$ of maps $\mu_0 : S \to T$ and $\mu_\omega : S_\omega \to T_\omega$ such that μ_0 is a morphism of semigroups and we have

$$\mu_\omega(sw) = \mu_0(s)\mu_\omega(w) \quad \text{and} \quad \mu_\omega(\pi(s_i)_i) = \pi(\mu_0(s_i))_i\,.$$

(c) A semigroup S operates in a canonical way on S^n and on $\wp(S)$ by

$$a \cdot \langle b_0, \dots, b_{n-1} \rangle := \langle ab_0, \dots, ab_{n-1} \rangle \quad \text{and} \quad a \cdot P := \{\, ab \mid b \in P \,\},$$

respectively.

For more information about ω-semigroups we recommend the book [8].

Definition 1.2. Let Σ be a set.

(a) A *tree domain* is a prefix closed subset $T \subseteq \omega^{<\omega}$ such that $wd \in T$, for $w \in \omega^{<\omega}$ and $d < \omega$, implies $wc \in T$, for all $c < d$. The element $wd \in T$ is called the *d-th successor* of $w \in \omega^{<\omega}$.

(b) A *Σ-labelled tree* is a function $t : \mathrm{dom}(t) \to \Sigma$ where $\mathrm{dom}(t)$ is a tree domain.

(c) Let Σ be a *signature,* i.e., a set of function symbols, and let X be a set of variable symbols. A *Σ-term* with variables X is a $(\Sigma \cup X)$-labelled tree t where

- every internal vertex $x \in \mathrm{dom}(t)$ is labelled by some function symbol $t(x) \in \Sigma$ and the number of immediate successors of x coincides with the arity of the symbol $t(x)$,
- every leaf x of t is labelled either by a variable from X or a constant symbol (i.e., a 0-ary function symbol) from Σ.

The set of all Σ-terms with variables from X is denoted by $T_\omega[\Sigma, X]$, the set of all finite Σ-terms by $T[\Sigma, X]$.

Definition 1.3. A *parity automaton* is a tuple $\mathcal{A} = \langle Q, \Sigma, \delta, q_0, \Omega \rangle$ where Q is the finite set of states, Σ is a signature, $q_0 \in Q$ is the initial state, $\Omega : Q \to \omega$ is a priority function, and $\Delta \subseteq Q \times \Sigma \times Q^*$ is the transition relation.

Let $t \in T_\omega[\Sigma, \varnothing]$. A *run* of $\mathcal{A}$ on t is a Q-labelled tree $\rho : \mathrm{dom}(t) \to Q$ with the same domain as t such that $\rho(\langle\rangle) = q_0$ and

$$\big\langle \rho(x), t(x), \rho(y_0) \dots \rho(y_{n-1}) \big\rangle \in \Delta\,,$$

for every vertex $x \in \mathrm{dom}(t)$ with immediate successors $y_0, \dots, y_{n-1}$.

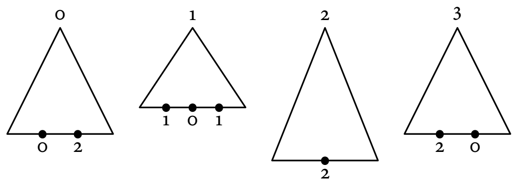
Such a run is *accepting* if every infinite branch $x_0, x_1, \dots$ in $\mathrm{dom}(t)$ satisfies the following *parity condition:*

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

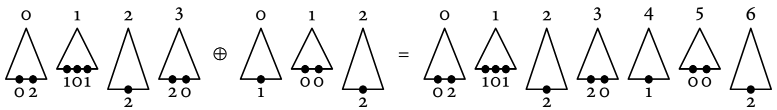
The *language recognised* by $\mathcal{A}$ consist of all trees $t \in T_\omega[\Sigma, \varnothing]$ for which there is an accepting run. Languages recognised by some parity automaton are called *regular.*

2 Hyperclones

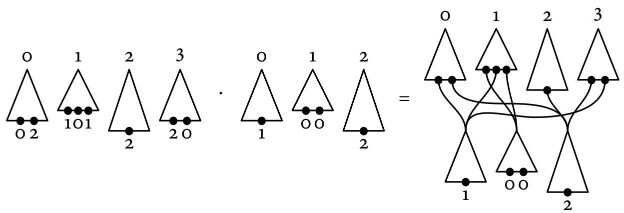
Before introducing the notion of a hyperclone let us first give some intuition. A hyperclone is an algebra where each element can be thought of as a tuple of objects each of which has a number of *ports*. Each port is labelled by a natural number. For instance, the objects could be terms where each occurrence of a variable x_i corresponds to a port with label i. In particular, the ports are arranged in a left-to-right fashion and there may be several ports with the same label. For simplicity, we assume that each object has only finitely many ports. Hence, to each object we can associate a finite tuple of natural numbers. We can depict an element of a hyperclone consisting of four objects with ports $\langle 0, 2\rangle$, $\langle 1, 0, 1\rangle$, $\langle 2\rangle$, and $\langle 2, 0\rangle$, respectively, as in the following diagram.



There are two main operations in a hyperclone: horizontal and vertical composition. Horizontal composition $\oplus$ is just the concatenation of tuples.



For vertical composition $\cdot$ we plug in the i-th object of the second tuple into every port of the first tuple with label i. For instance, if the objects are terms then vertical composition might correspond in substituting the i-th term of the second tuple for the variable x_i in each term of the first tuple.



For technical reasons, we will not use the basic vertical composition $\cdot$, but a more complicated operation $:_{I,\tau}$ where we can plug in objects into only some of the ports, while changing the numbers of the remaining ports.

The motivating example of a hyperclone is the algebra consisting of all terms. We define this algebra first, before giving the general definition of a hyperclone.

Definition 2.1. (a) Let Σ be a signature and set $X := \{x_0, x_1, \dots\}$. We denote by $T_\omega[\Sigma]$ the set of all terms $t \in T_\omega[\Sigma, X]$ with only finitely many occurrences of variables from X. As usual, $T[\Sigma]$ is the subset of $T_\omega[\Sigma]$ consisting of all finite terms.

For a term $t \in T_\omega[\Sigma]$, we denote by $\operatorname{var}(t)$ the sequence of all (indices of) variables appearing in t in left-to-right order. Formally, we define var as the unique function $T_\omega[\Sigma] \to \omega^{<\omega}$ satisfying the following equations:

$$\begin{aligned} \operatorname{var}(t) &= \langle\rangle, \quad \text{if } t \text{ does not contain a variable}, \\ \operatorname{var}(x_i) &= \langle i\rangle, \\ \text{and}\quad \operatorname{var}(f(t_0, \dots, t_{n-1})) &= \operatorname{var}(t_0) \dots \operatorname{var}(t_{n-1}). \end{aligned}$$

(b) The *free ω-hyperclone* over Σ is the many-sorted structure

$$\mathfrak{F}_\omega[\Sigma] := \big\langle (F_{\bar u}[\Sigma])_{\bar u \in (\omega^{<\omega})^{<\omega}}, \oplus, 0, (\lambda_\sigma)_\sigma, (:_{I,\tau})_{I,\tau}, \pi, \leq \big\rangle$$

where I ranges over finite subsets of ω, σ over all functions $[m] \to [n]$, for $m, n < \omega$, and τ over all functions $[m] \to \omega$, for $m < \omega$. The set of sorts is $(\omega^{<\omega})^{<\omega}$ where the domain of sort $\bar u = \langle u_0, \dots, u_{n-1}\rangle \in (\omega^{<\omega})^{<\omega}$ is the set

$$F_{\bar u}[\Sigma] := \big\{ \langle t_0, \dots, t_{n-1}\rangle \bigm| t_i \in T_\omega[\Sigma] \smallsetminus \{x_0, x_1, \dots\} \text{ with } \operatorname{var}(t_i) = u_i \big\}$$

of all finite tuples of non-trivial terms such that the i-th term has variables u_i. $\mathfrak{F}_\omega[\Sigma]$ has the following operations:

- $\oplus$ is the concatenation of tuples.
- $:_{I,\tau}$, for $I \subseteq [n]$ and $\tau : [k] \to \omega$, is defined as
  $$\langle s_0, \dots, s_{m-1}\rangle :_{I,\tau} \langle t_0, \dots, t_{n-1}\rangle := \langle u_0, \dots, u_{m-1}\rangle,$$
  where u_l is obtained from s_l by replacing every occurrence of a variable x_i with $i \in I$ by the term t_i. Variables x_i with $i \notin I$ are replaced by the variable $x_{\tau(i)}$.
- λ_σ reorders its argument according to $\sigma : [m] \to [n]$:
  $$\lambda_\sigma\langle t_0, \dots, t_{n-1}\rangle = \langle t_{\sigma(0)}, \dots, t_{\sigma(m-1)}\rangle.$$
- $\pi(a_0, a_1, \dots)$ is the limit of the terms a_0, $(a_0 \cdot a_1)$, $(a_0 \cdot a_1 \cdot a_2)$,…, where, for an m-tuple $\bar s$ and an n-tuple $\bar t$, the simple version of the vertical composition is defined by $\bar s \cdot \bar t := \bar s :_{[n],\mathrm{id}} \bar t$.

- o denotes the empty tuple of terms.
- The order $\leq$ is trivial: $a \leq b$ iff $a = b$.

(c) The *free hyperclone* over Σ is the restriction $\mathfrak{F}[\Sigma]$ of $\mathfrak{F}_\omega[\Sigma]$ where we omit the infinite product π and restrict the domains to finite terms only.

Remark. We have omitted trivial terms $t = x_i$ consisting of a single variable from our algebra in order to avoid technical difficulties with the infinite product. (The product of $x_0, x_0, x_0, \dots$ is the term x_0, but it should be a term without variables.) The price we have to pay for this choice is that we need a more complication version $:_{I,\tau}$ of the vertical composition to build terms like

$$f(x_2, g(x_2)) = f(x_0, x_1) :_{\{1\},\tau} \langle h(x_7), g(x_2)\rangle \quad \text{where } \tau(0) = 2\,.$$

Before stating the general definition of a hyperclone, we need a bit of notation. To manipulate finite sequences we define the following operations.

Definition 2.2. Let A, B be sets.

(a) For $\bar{a} = \langle a_0, \dots, a_{n-1}\rangle \in A^n$ and functions $\sigma : [m] \to [n]$ and $\tau : A \to B$, we define

$$\bar{a}^\sigma := \langle a_{\sigma(0)}, \dots, a_{\sigma(m-1)}\rangle \in A^m$$
$$\text{and} \quad \tau(\bar{a}) := \langle \tau(a_0), \dots, \tau(a_{n-1})\rangle \in B^n.$$

Note that $(\bar{a}^\sigma)^{\sigma'} = \bar{a}^{\sigma \circ \sigma'}$ and $\tau(\tau'(\bar{a})) = (\tau \circ \tau')(\bar{a})$, for $\sigma' : [n] \to [k]$ and $\tau' : B \to C$.

(b) For readability, we introduce a *concatenation operation*

$$\operatorname{cat} : (A^{<\omega})^{<\omega} \to A^{<\omega} \quad \text{by} \quad \operatorname{cat}(\bar{a}_0, \dots, \bar{a}_{n-1}) := \bar{a}_0 \dots \bar{a}_{n-1}\,.$$

Since the definition of a hyperclone is rather long, we isolate some parts and treat them separately. We start by axiomatising the horizontal composition $\oplus$.

Definition 2.3. Let S be a set and $\mathfrak{C}$ an $S^{<\omega}$-sorted structure with domains $C_{\bar{u}}$, for $\bar{u} \in S^{<\omega}$.

(a) The *width* of an element $a \in C_{\bar{u}}$ with $\bar{u} \in S^n$ is the number n.

(b) A *free monoid structure* on $\mathfrak{C}$ consists of

- a constant $\text{o} \in C_{\langle\rangle}$;
- binary operations $\oplus : C_{\bar{u}} \times C_{\bar{v}} \to C_{\bar{u}\bar{v}}$ for all $\bar{u}, \bar{v} \in S^{<\omega}$;
- unary operations $\lambda_\sigma : C_{\bar{u}} \to C_{\bar{u}^\sigma}$ for all $\bar{u} \in S^{<\omega}$ and every function $\sigma : [m] \to [n]$ with $n = |\bar{u}|$ and $m < \omega$

that satisfy the following conditions:

(M1) The monoid laws for $\oplus$:

$$(a \oplus b) \oplus c = a \oplus (b \oplus c) \quad \text{and} \quad \text{o} \oplus a = a = a \oplus \text{o}\,,$$

for all $a, b, c \in C$.

(M2) o is the only element of sort $\langle\rangle$:

$$C_{\langle\rangle} = \{\text{o}\}\,.$$

(M3) The laws of the action of λ_σ on C:

$$\lambda_\sigma(\lambda_\tau(a)) = \lambda_{\tau\circ\sigma}(a) \quad \text{and} \quad \lambda_{\mathrm{id}}(a) = a$$

for all $a \in C$ and all suitable σ and τ.

(M4) We can decompose every element $a \in C_{\bar{u}}$ of width m into a sum of elements in $C_{\langle u_0\rangle}, \dots, C_{\langle u_{m-1}\rangle}$ of width 1:

$$a = \lambda_{\tau_0}(a) \oplus \dots \oplus \lambda_{\tau_{m-1}}(a)\,, \quad \text{where } \tau_i : [1] \to [m] : 0 \mapsto i\,.$$

(M5) For $a \in C_{\bar{u}}$ of width m and $b \in C_{\bar{v}}$ of width n

$$\lambda_\sigma(a \oplus b) = a \quad \text{and} \quad \lambda_\tau(a \oplus b) = b\,,$$

where $\sigma : [m] \to [m+n] : i \mapsto i$
and $\tau : [n] \to [m+n] : i \mapsto m + i$.

Lemma 2.4. *Let* o, $\oplus$, $(\lambda_\sigma)_\sigma$ *be a free monoid structure on an* $S^{<\omega}$*-sorted structure* $\mathfrak{C}$. *For every sort* $\bar{u} = \langle u_0, \dots, u_{n-1}\rangle \in S^{<\omega}$, *the function*

$$C_{\langle u_0\rangle} \times \dots \times C_{\langle u_{n-1}\rangle} \to C_{\bar{u}} : \langle a_0, \dots, a_{n-1}\rangle \mapsto a_0 \oplus \dots \oplus a_{n-1}$$

is bijective. Furthermore,

$$\lambda_\sigma(a_0 \oplus \dots \oplus a_{n-1}) = a_{\sigma(0)} \oplus \dots \oplus a_{\sigma(m-1)}\,,$$

for all $a_i \in C_{\langle u_i\rangle}$ *with* $u_i \in S$ *and all functions* $\sigma : [m] \to [n]$.

Proof. The above function is surjective by (M4). For injectivity, it is sufficient to prove that

$$a \oplus b = a' \oplus b' \quad \text{implies} \quad a = a' \text{ and } b = b' ,$$

for all $a, a' \in C_{\bar u}$ and $b, b' \in C_{\bar v}$. This follows from (M5) since

$$a \oplus b = a' \oplus b' \quad \text{implies} \quad a = \lambda_\sigma(a \oplus b) = \lambda_\sigma(a' \oplus b') = a'$$

and similarly for b and b'.

For the second claim, we consider an element $a = a_0 \oplus \cdots \oplus a_{n-1}$ of width n. For $\tau_k : [1] \to [n] : 0 \mapsto k$ and $\sigma_k : [n-k] \to [n] : i \mapsto i + k$ it follows by (M5) that

$$\begin{aligned} \lambda_{\tau_k}(a) &= \lambda_{\sigma_k \circ \tau_0}(a_0 \oplus \cdots \oplus a_{n-1}) \\ &= \lambda_{\tau_0}\big(\lambda_{\sigma_k}\big((a_0 \oplus \cdots \oplus a_{k-1}) \oplus (a_k \oplus \cdots \oplus a_{n-1})\big)\big) \\ &= \lambda_{\tau_0}(a_k \oplus \cdots \oplus a_{n-1}) \\ &= a_k . \end{aligned}$$

For $\sigma : [m] \to [n]$, we therefore obtain

$$\begin{aligned} \lambda_\sigma(a) &= \lambda_{\tau_0}(\lambda_\sigma(a)) \oplus \cdots \oplus \lambda_{\tau_{m-1}}(\lambda_\sigma(a)) \\ &= \lambda_{\sigma \circ \tau_0}(a) \oplus \cdots \oplus \lambda_{\sigma \circ \tau_{m-1}}(a) \\ &= \lambda_{\tau_{\sigma(0)}}(a) \oplus \cdots \oplus \lambda_{\tau_{\sigma(m-1)}}(a) \\ &= a_{\sigma(0)} \oplus \cdots \oplus a_{\sigma(m-1)} . \end{aligned}$$

□

This lemma tells us that, in every algebra with a free monoid structure, we can regard elements of width n as n-tuples of elements of width 1. This identification is formalised in the following definition.

Definition 2.5. Let $\mathfrak{C}$ be a structure with a free monoid structure $0, \oplus, (\lambda_\sigma)_\sigma$. The *decomposition* of an element $a \in C$ of width m is the tuple $\langle a_0, \dots, a_{m-1}\rangle$ of elements a_i of width 1 such that $a = a_0 \oplus \cdots \oplus a_{m-1}$.

The free hyperclones defined above are $S^{<\omega}$-sorted algebras for $S = \omega^{<\omega}$. We define the following notation regarding this set of sorts.

Definition 2.6. Let $S := (\omega^{<\omega})^{<\omega}$.

(a) We write elements $\bar u \in S$ as tuples $\bar u = \langle u_0, \dots, u_{n-1}\rangle$ of functions $u_i : [m_i] \to \omega$.

(b) Let $\bar u = \langle u_0, \dots, u_{n-1}\rangle \in S$. The *width* of $\bar u$ is its length n. The *support* of $\bar u$ is the set

$$\operatorname{supp}(\bar u) := \operatorname{rng} u_0 \cup \cdots \cup \operatorname{rng} u_{n-1} .$$

(c) For two tuples $\bar u = \langle u_0, \dots, u_{m-1}\rangle$ and $\bar v = \langle v_0, \dots, v_{n-1}\rangle$ in S, a set $I \subseteq [m]$, and a function $\sigma : [k] \to \omega$ such that

$$\operatorname{supp}(\bar u) \cap I \subseteq [n] \quad \text{and} \quad \operatorname{supp}(\bar u) \setminus I \subseteq [k] ,$$

we define the *substitution operation*

$$\bar u :_{I,\sigma} \bar v := \langle \operatorname{cat}(\bar w^0), \dots, \operatorname{cat}(\bar w^{m-1})\rangle$$

where

$$w^k_i := \begin{cases} v_{u_k(i)} & \text{if } u_k(i) \in I , \\ \langle \sigma(u_k(i))\rangle & \text{if } u_k(i) \notin I . \end{cases}$$

Example. Let $\bar u = \big\langle \langle 0, 2, 1\rangle, \langle 1, 2\rangle \big\rangle$, and $\bar v = \big\langle \langle 0, 1\rangle, \langle 1, 2\rangle, \langle 0, 5\rangle \big\rangle$. Then

$$\bar u :_{\{0,2\},\sigma} \bar v = \big\langle \langle 0, 1, 0, 5, \sigma(1)\rangle, \langle \sigma(1), 0, 5\rangle \big\rangle .$$

In the next step, we introduce the axioms for the substitution operation $:_{I,\sigma}$.

Definition 2.7. Let $S := (\omega^{<\omega})^{<\omega}$ and let $\mathfrak{C}$ be an S-sorted algebra with free monoid structure $0, \oplus, (\lambda_\sigma)_\sigma$.

(a) A *substitution operation* is a family of binary operations

$$:_{I,\sigma} : C_{\bar u} \times C_{\bar v} \to C_{\bar u :_{I,\sigma} \bar v} , \quad \begin{array}{l} \text{for all finite } I \subseteq \omega, \text{ all } \sigma : [m] \to \omega, \text{ and} \\ \text{all } \bar u, \bar v \in S \text{ such that } \bar u :_{I,\sigma} \bar v \text{ is defined,} \end{array}$$

that satisfy the following conditions.

(S1) For $b = b_0 \oplus \cdots \oplus b_{m-1}$ of width m and $c = c_0 \oplus \cdots \oplus c_{n-1}$ of width n, we have the associative law

$$(a :_{I,\sigma} b) :_{J,\tau} c = a :_{I \cup \sigma^{-1}[J], \tau \circ \sigma} (d_0 \oplus \cdots \oplus d_{m-1}) ,$$

where

$$d_i := \begin{cases} b_i :_{J,\tau} c & \text{if } i \in I\,, \\ c_{\sigma(i)} & \text{if } i \notin I\,. \end{cases}$$

(s2) The distributive law

$$(a \oplus b) :_{I,\sigma} c = (a :_{I,\sigma} c) \oplus (b :_{I,\sigma} c)\,.$$

(s3) $a :_{\varnothing,\mathrm{id}} b = a$

(s4) Let $\bar{u}, \bar{v} \in S$ be sorts and $\sigma : [m] \to \omega$ a function such that $\bar{u} :_{\varnothing,\sigma} \bar{v}$ is defined. For every $a \in C_{\sigma(\bar{u})}$, there are elements $b \in C_{\bar{u}}$ and $c \in C_{\bar{v}}$ with $a = b :_{\varnothing,\sigma} c$.

(s5) For $a \in C_{\bar{u}}$ and $b, b' \in C_{\bar{v}}$ with decompositions $b = b_0 \oplus \dots \oplus b_{n-1}$ and $b' = b'_0 \oplus \dots \oplus b'_{n-1}$,

$$a :_{I,\sigma} b = a :_{J,\tau} b'\,,$$

for all sets I, J and functions σ, τ such that

- $b_i = b'_i$, for all $i \in I \cap \mathrm{supp}(\bar{u})$,
- $I \cap \mathrm{supp}(\bar{u}) = J \cap \mathrm{supp}(\bar{u})$,
- $\sigma \restriction (\mathrm{supp}(\bar{u}) \setminus I) = \tau \restriction (\mathrm{supp}(\bar{u}) \setminus I)$.

(b) For a substitution operation $(:_{I,\sigma})_{I,\sigma}$, we define the following abbreviations. For $a \in C_{\bar{u}}$ and $b \in C_{\bar{v}}$, we set

$$\begin{aligned} a :_I b &:= a :_{I,\mathrm{id}} b\,, \\ a \cdot b &:= a :_{\mathrm{supp}(\bar{u})} b\,, \\ \rho_\sigma(a) &:= a :_{\varnothing,\sigma} c\,, \qquad \text{for an arbitrary } c\,. \end{aligned}$$

Note that, by (s5), the value of $\rho_\sigma(a)$ does not depend on the chosen element c. Let us collect some properties of the operation ρ_σ.

Lemma 2.8. *Let $(:_{I,\sigma})_{I,\sigma}$ be a substitution operation for $\mathfrak{C}$. For all suitable values of a, b, σ, τ, and I,*

(a) $\rho_{\mathrm{id}}(a) = a$

(b) $\rho_\sigma(\rho_\tau(a)) = \rho_{\sigma\circ\tau}(a)$

(c) $\lambda_\sigma(\rho_\tau(a)) = \rho_\tau(\lambda_\sigma(a))$

(d) $\lambda_\tau(a :_{I,\sigma} b) = \lambda_\tau(a) :_{I,\sigma} b$

(e) $\rho_\tau(a :_{I,\sigma} b) = a :_{I,\tau\circ\sigma} \rho_\tau(b)$

(f) $\rho_\tau(a) :_{I,\sigma} b = a :_{\tau^{-1}[I],\sigma\circ\tau} \lambda_\tau(b)$

(g) *$\rho_\sigma : C_{\bar{u}} \to C_{\sigma(\bar{u})}$ is bijective, for all suitable $\bar{u}$ and σ.*

(h) *if $a \in C_{\bar{u}}$ and $\sigma \restriction \mathrm{supp}(\bar{u}) = \tau \restriction \mathrm{supp}(\bar{u})$, then $\rho_\sigma(a) = \rho_\tau(a)$.*

Proof. (a) is just a restatement of (s3).

(b) We have

$$\begin{aligned} \rho_\sigma(\rho_\tau(a)) &= (a :_{\varnothing,\tau} b) :_{\varnothing,\sigma} c \\ &= a :_{\varnothing,\sigma\circ\tau} (d_0 \oplus \dots \oplus d_{m-1}) = \rho_{\sigma\circ\tau}(a)\,, \end{aligned}$$

where $d_0, \dots, d_{m-1}$ are as in (s1).

(c) Let $a = a_0 \oplus \dots \oplus a_{m-1}$ and $\sigma : [n] \to [m]$. Then

$$\begin{aligned} \lambda_\sigma(\rho_\tau(a)) &= \lambda_\sigma(\rho_\tau(a_0 \oplus \dots \oplus a_{m-1})) \\ &= \lambda_\sigma(\rho_\tau(a_0) \oplus \dots \oplus \rho_\tau(a_{m-1})) \\ &= \rho_\tau(a_{\sigma(0)}) \oplus \dots \oplus \rho_\tau(a_{\sigma(n-1)}) \\ &= \rho_\tau(a_{\sigma(0)} \oplus \dots \oplus a_{\sigma(n-1)}) \\ &= \rho_\tau(\lambda_\sigma(a))\,. \end{aligned}$$

(d) Let $\tau : [m] \to [n]$ and let $a = a_0 \oplus \dots \oplus a_{n-1}$ be the decomposition of a. Then

$$\begin{aligned} \lambda_\tau(a :_{I,\sigma} b) &= \lambda_\tau((a_0 \oplus \dots \oplus a_{n-1}) :_{I,\sigma} b) \\ &= \lambda_\tau(a_0 :_{I,\sigma} b \oplus \dots \oplus a_{n-1} :_{I,\sigma} b) \\ &= (a_{\tau(0)} :_{I,\sigma} b) \oplus \dots \oplus (a_{\tau(m-1)} :_{I,\sigma} b) \\ &= (a_{\tau(0)} \oplus \dots \oplus a_{\tau(m-1)}) :_{I,\sigma} b \\ &= \lambda_\tau(a) :_{I,\sigma} b\,. \end{aligned}$$

(e) Let $b = b_0 \oplus \cdots \oplus b_{m-1}$ and $c = c_0 \oplus \cdots \oplus c_{n-1}$. Setting

$$d_i := \begin{cases} b_i :_{\varnothing,\tau} c & \text{if } i \in I, \\ c_{\sigma(i)} & \text{if } i \notin I \end{cases}$$

as in (s1), it follows that

$$\begin{aligned} \rho_\tau(a :_{I,\sigma} b) &= (a :_{I,\sigma} b) :_{\varnothing,\tau} c \\ &= a :_{I,\tau\circ\sigma} (d_0 \oplus \cdots \oplus d_{m-1}) \\ &= a :_{I,\tau\circ\sigma} ((b_0 :_{\varnothing,\tau} c) \oplus \cdots \oplus (b_{m-1} :_{\varnothing,\tau} c)) \\ &= a :_{I,\tau\circ\sigma} ((b_0 \oplus \cdots \oplus b_{m-1}) :_{\varnothing,\tau} c) \\ &= a :_{I,\tau\circ\sigma} \rho_\tau(b) . \end{aligned}$$

(f) Let $b = b_0 \oplus \cdots \oplus b_{m-1}$ and $c = c_0 \oplus \cdots \oplus c_{n-1}$. Setting

$$d_i := \begin{cases} c_i :_{I,\sigma} b & \text{if } i \in \varnothing, \\ b_{\tau(i)} & \text{if } i \notin \varnothing \end{cases}$$

as in (s1), it follows that

$$\begin{aligned} \rho_\tau(a) :_{I,\sigma} b &= (a :_{\varnothing,\tau} c) :_{I,\sigma} b \\ &= a :_{\tau^{-1}[I],\sigma\circ\tau} (d_0 \oplus \cdots \oplus d_{n-1}) \\ &= a :_{\tau^{-1}[I],\sigma\circ\tau} (b_{\tau(0)} \oplus \cdots \oplus b_{\tau(n-1)}) \\ &= a :_{\tau^{-1}[I],\sigma\circ\tau} \lambda_\tau(b) . \end{aligned}$$

(g) Consider $\sigma : [m] \to [n]$. It follows immediately from (s4) that ρ_σ is surjective. For injectivity, suppose that $\rho_\sigma(a) = \rho_\sigma(b)$. Let $\tau : [n] \to [1]$ be the constant function with value 0. Since $\tau \circ \sigma : [m] \to [1]$ is surjective, there exists a function $\tau' : [1] \to [m]$ such that $\tau \circ \sigma \circ \tau' = \mathrm{id}$. By (s4), we can find elements a_0 and b_0 such that $a = \rho_{\tau'}(a_0)$ and $b = \rho_{\tau'}(b_0)$. It follows that

$$\begin{aligned} a_0 = \rho_{\mathrm{id}}(a_0) = \rho_{\tau\circ\sigma\circ\tau'}(a_0) = \rho_\tau(\rho_\sigma(\rho_{\tau'}(a_0))) = \rho_\tau(\rho_\sigma(a)) \\ = \rho_\tau(\rho_\sigma(b)) = \rho_\tau(\rho_\sigma(\rho_{\tau'}(b_0))) = \rho_{\tau\circ\sigma\circ\tau'}(b_0) = \rho_{\mathrm{id}}(b_0) = b_0 . \end{aligned}$$

Consequently, $a = \rho_{\tau'}(a_0) = \rho_{\tau'}(b_0) = b$.

(h) follows from (s5):

$$\rho_\sigma(a) = a :_{\varnothing,\sigma} b = a :_{\varnothing,\tau} b = \rho_\tau(a) . \qquad \square$$

Mainly we are interested in the simple version $\cdot$ of the substitution operation. The corresponding laws take the following simpler form.

Corollary 2.9. *Let $(:_{I,\sigma})_{I,\sigma}$ be a substitution operation for $\mathfrak{C}$. Then*

(a) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

(b) $(a \cdot b) :_{I,\sigma} c = a \cdot (b :_{I,\sigma} c)$

(c) $\lambda_\tau(a \cdot b) = \lambda_\tau(a) \cdot b$

(d) $\rho_\tau(a \cdot b) = a \cdot \rho_\tau(b)$

(e) $\rho_\tau(a) \cdot b = a \cdot \lambda_\tau(b)$

Definition 2.10. Let $(:_{I,\sigma})_{I,\sigma}$ be a substitution operation for $\mathfrak{C}$ and let $a \in C_{\bar{u}}$ be an element of sort $\bar{u} = u_0 \dots u_{m-1}$.

(a) We say that a is *in separation normal form* if there are numbers $0 = k_0 \leq \cdots \leq k_m < \omega$ such that

$$u_i = \langle k_i, k_i + 1, \dots, k_{i+1} - 1 \rangle , \quad \text{for all } i < m .$$

(b) The *separation normal form* of a is an element b in separation normal form such that $a = \rho_\sigma(b)$, for some σ. We denote this normal form by $\mathrm{sep}(a)$.

Remark. Since the operations ρ_σ are bijective, every element $a \in C_{\bar{u}}$ has a unique separation normal form. Hence, $\mathrm{sep}(a)$ is well-defined.

Example. In the free hyperclone we have

$$\mathrm{sep}\langle f(x_0, f(x_2, x_0)),\ f(x_0, x_2)\rangle = \langle f(x_0, f(x_1, x_2)),\ f(x_3, x_4)\rangle .$$

After these preparations we can finally state the definition of a hyperclone.

Definition 2.11. (a) Let $S := (\omega^{<\omega})^{<\omega}$. A *hyperclone* is an S-sorted structure

$$\mathfrak{C} = \langle (C_{\bar{u}})_{\bar{u}\in S},\ \oplus,\ (\ :_{I,\sigma}\ )_{I,\sigma},\ 0,\ (\lambda_\sigma)_{\sigma\in\omega^{<\omega}}, \leq \rangle$$

such that

- $0, \oplus, (\lambda_\sigma)_\sigma$ is a free monoid structure on $\mathfrak{C}$,
- $(\ :_{I,\sigma}\ )_{I,\sigma}$ is a substitution structure on $\mathfrak{C}$, and

- $\leq$ is a (family of) partial orders of each domain $C_{\bar{u}}$ such that, for all suitable a, b, c,

$$a \leq b \quad \text{implies} \quad \begin{cases} a \oplus c \leq b \oplus c\,, & a :_{I,\sigma} c \leq b :_{I,\sigma} c\,, \\ c \oplus a \leq c \oplus b\,, & c :_{I,\sigma} a \leq c :_{I,\sigma} b\,. \end{cases}$$

(b) To simplify notation, we write

$$\begin{matrix} \sigma a := \lambda_\sigma(a)\,, & ab := a \cdot b\,, \\ a\sigma := \rho_\sigma(a)\,, & \sigma\tau := \tau \circ \sigma\,, \end{matrix} \quad \text{for } a, b \in C \text{ and } \sigma, \tau \in \omega^{<\omega}.$$

(c) The *support* of an element $a \in C_{\bar{u}}$ is the support of its sort $\bar{u}$:

$$\operatorname{supp}(a) := \operatorname{supp}(\bar{u})\,.$$

(d) A hyperclone $\mathfrak{C}$ is *finitary* if every domain $C_{\bar{u}}$ is finite.

(e) Morphisms between hyperclones are defined in the usual way, i.e., they are sort-preserving maps that are compatible with every operation and preserve the ordering.

Remark. In the simplified notation, we have the following associative laws for a hyperclone:

$$\begin{matrix} \sigma(\tau a) = (\sigma\tau)a\,, & \sigma(a\tau) = (\sigma a)\tau\,, \\ (a\sigma)\tau = a(\sigma\tau)\,, & (a\sigma)b = a(\sigma b)\,. \end{matrix}$$

To check that the above definition of a hyperclone captures our intuition of generalising $\mathfrak{F}[\Sigma]$ we outline a proof of the fact that the 'free hyperclone' $\mathfrak{F}[\Sigma]$ is really free.

Remark. Note that the free hyperclone $\mathfrak{F}[\Sigma]$ is generated by all terms of the form $\sigma(x_0, \dots, x_{n-1})$, for $\sigma \in \Sigma$. In the following, we identify the symbols $\sigma \in \Sigma$ with these terms $\sigma(x_0, \dots, x_{n-1})$. Hence, we regard Σ as a subset of $\mathfrak{F}[\Sigma]$.

Lemma 2.12. *$\mathfrak{F}[\Sigma]$ is the free hyperclone generated by Σ.*

Proof. To show that $\mathfrak{F}[\Sigma]$ is free consider a hyperclone $\mathfrak{C}$ and a sort-preserving function $f : \Sigma \to C$. We have to extend f to a morphism $\varphi : \mathfrak{F}[\Sigma] \to \mathfrak{C}$.

First, we define $\varphi(t)$, for single terms t. We do so by induction on t. If $t = \sigma(x_{i_0}, \dots, x_{i_{m-1}})$ for some $\sigma \in \Sigma$, we define

$$\varphi(t) := \rho_\tau(f(\sigma))\,, \quad \text{where} \quad \tau : [m] \to \omega : k \mapsto i_k\,.$$

For the inductive step, consider a term $t = \sigma(s_0, \dots, s_{m-1})$ with $\sigma \in \Sigma$. We define $I \subseteq [m]$ and $\tau : [m] \to \omega$ by

$$I := \{\, i < m \mid s_i \text{ is a non-trivial term} \,\}\,,$$

$$\tau(i) := \begin{cases} k & \text{if } s_i = x_k\,, \\ 0 & \text{if } i \in I\,. \end{cases}$$

We define

$$\varphi(t) := f(\sigma) :_{I,\tau} b\,,$$

where $b := b_0 \oplus \dots \oplus b_{m-1}$ for

$$b_i := \begin{cases} \varphi(s_i) & \text{for } i \in I\,, \\ \text{arbitrary} & \text{otherwise}\,. \end{cases}$$

For tuples $\bar{t} = \langle t_0, \dots, t_{m-1} \rangle$ with more than one component, we can now set

$$\varphi(\bar{t}) := \varphi(t_0) \oplus \dots \oplus \varphi(t_{m-1})\,.$$

For the empty tuple, we define $\varphi(\langle\rangle) := 0$.

It remains to prove that the mapping φ defined in this way is indeed a morphism of hyperclones. Immediately from the definition of φ we see that

$$\varphi(0) = 0 \quad \text{and} \quad \varphi(\bar{s} \oplus \bar{t}) = \varphi(\bar{s}) \oplus \varphi(\bar{t})\,.$$

For λ_σ, we have

$$\varphi(\lambda_\sigma(\bar{t})) = \varphi(\bar{t}^\sigma) = \varphi(t_{\sigma(0)}) \oplus \dots \oplus \varphi(t_{\sigma(m-1)}) = \lambda_\sigma(\varphi(\bar{t}))\,.$$

If t is a single term, we can show by induction on t that

$$\varphi(t :_{I,\sigma} \bar{s}) = \varphi(t) :_{I,\sigma} \varphi(\bar{s})\,.$$

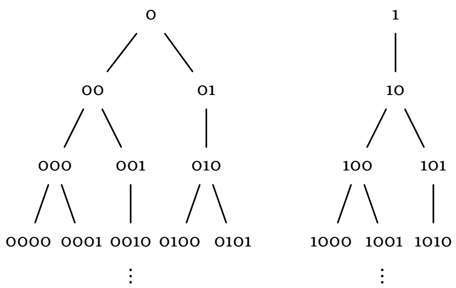
This implies the same statements for tuples $\bar{t}$. □

3 ω-HYPERCLONES

Our main object of study are so-called *ω-hyperclones* which are obtained from hyperclones by adding an infinite vertical product $\pi(a^0, a^1, a^2, \dots)$. To formally state the associative law for such infinite products we have to deal with different factorisations of a sequence $(a^n)_{n<\omega}$. Note that we can split such a sequence not only by cutting it 'horizontally' (i.e., $\langle a^0, \dots, a^{k-1}\rangle, \langle a^k, \dots, a^{l-1}\rangle, \dots$), but also 'vertically' (i.e, $a^n = b^n \oplus c^n$, for all n) and in many other intermediate ways. To make this idea precise we associate with $(a^n)_n$ a forest as follows.

Let $a^n = a^n_0 \oplus \dots \oplus a^n_{m(n)-1}$ be the decomposition of a^n into elements of width 1. We construct the graph whose vertices are pairs of indices $\langle n, i\rangle$ (representing a^n_i), for all $n < \omega$ and $i < m(i)$, where we add an edge from $\langle n, i\rangle$ to $\langle n{+}1, k\rangle$ if $k \in \operatorname{supp}(a^n_i)$. The *branch tree* of $(a^n)_n$ is the tree-unfolding of this graph.

Example. In the free ω-hyperclone, consider the sequence $(a^n)_{n<\omega}$ where $a^n = \langle f(x_0, x_1),\ g(x_0)\rangle$, for every n. The first levels of the branch tree of this sequence are



Each partition of this branch tree into connected parts corresponds to a factorisation of the product $\pi(a^0, a^1, \dots)$. In the following definition we will encode such a partition by a set H containing the least element (in the tree order) of each class. The associative law below states that every such factorisation results in the same product.

Definition 3.1. Let $\mathfrak{C}$ be a hyperclone.

(a) We use the notation $a^\square$ for a sequence $(a^n)_{n<\alpha}$ of length $\alpha \leq \omega$ such that the product $a^n \cdot a^{n+1}$ is defined for every n.

(b) We say that a sequence $a^\square = (a^n)_{n<\alpha}$ is *in separation normal form* if every a^n is in separation normal form.

(c) Let $a^\square = (a^n)_{n<\alpha}$ be a sequence, let $m(n)$ be the width of a^n, and suppose that $a^n = a^n_0 \oplus \dots \oplus a^n_{m(n)-1}$ is the decomposition of a^n into elements of width 1.

The *branch tree* of $a^\square$ is the forest

$$\Lambda(a^\square) := \big\{\, \eta \in \omega^{<\alpha} \;\big|\; |\eta| > 0,\ \eta(0) < m(0),\ \text{and} \quad \eta(n+1) \in \operatorname{supp}(a^n_{\eta(n)}),\ \text{for all } n \,\big\}\,.$$

A *branch* of $a^\square$ is a sequence $\beta \in \omega^{\leq\omega}$ of maximal length such that every finite prefix $\eta \preceq \beta$ belongs to $\Lambda(a^\square)$.

(d) We define functions $\mu : \Lambda(a^\square) \to C$ and $\mu_0 : \Lambda(a^\square) \to \omega \times \omega$ by setting

$$\mu_0(\eta) := \langle n-1, \eta(n-1)\rangle \quad \text{and} \quad \mu(\eta) := a^{n-1}_{\eta(n-1)}\,, \quad \text{for } n := |\eta| > 0\,.$$

Hence, $\mu(\eta) = a^n_k$, for $\langle n, k\rangle := \mu_0(\eta)$.

A subset $U \subseteq \Lambda(a^\square)$ is *connected* if there exists an element $\eta \in U$ such that

$$\zeta \in U \quad \text{implies} \quad \eta \preceq \zeta \text{ and } \xi \in U \text{ for all } \eta \preceq \xi \preceq \zeta\,.$$

We extend μ to finite connected subsets $U \subseteq \Lambda(a^\square)$ as follows. We define an element $\mu[U] \in C$ by induction on $|U|$. For $U = \{\eta\}$, we set

$$\mu[\{\eta\}] := \operatorname{sep}(\mu(\eta))\,.$$

For $|U| > 1$, let η be the minimal element of U and let $\zeta_0, \dots, \zeta_{m-1}$ be its successors in $\Lambda(a^\square)$. We define

$$\mu[U] := \operatorname{sep}(b :_I c)\,.$$

where

$$\begin{aligned} I &:= \{\, i < m \mid \zeta_i \in U \,\}\,, \\ b &:= \mu[\{\eta\}]\,, \\ c &:= \mu[U_0] \oplus \dots \oplus \mu[U_{m-1}] \quad \text{for} \quad U_i := \{\zeta_i\} \cup \{\, \xi \in U \mid \xi \succeq \zeta_i \,\}\,. \end{aligned}$$

(e) Let $H \subseteq \Lambda(a^\square)$ be a set such that

- H contains all roots of $\Lambda(a^\square)$ (i.e., all sequences $\eta \in \Lambda(a^\square)$ with $|\eta| = 1$),
- every infinite branch of $\Lambda(a^\square)$ contains infinitely many elements of H,
- $\mu_0(\eta) = \mu_0(\zeta)$ implies $\eta \in H \Leftrightarrow \zeta \in H$, for $\eta, \zeta \in \Lambda(a^\square)$.

Then H *induces a factorisation* $b^\square$ of $a^\square$ as follows.

For each $\eta \in H$, we set

$$\|\eta\| := \big|\{\, \zeta \in H \mid \zeta \prec \eta \,\}\big|,$$
$$U_\eta := \{\, \zeta \in \Lambda(a^\square) \mid \eta \preceq \zeta \text{ and there is no } \eta' \in H \text{ with } \eta \prec \eta' \preceq \zeta \,\},$$
$$b_\eta := \mu[U_\eta].$$

We define

$$b^n := \operatorname{sep}(b_{\eta^n_0} \oplus \dots \oplus b_{\eta^n_l}),$$

where $\eta^n_0, \dots, \eta^n_l$ is an enumeration, from left to right, of all $\eta \in H$ with $\|\eta\| = n$.

(f) The *unravelling* of $a^\square$ is the factorisation $b^\square$ of $a^\square$ induced by the set $H := \Lambda(a^\square)$.

Remark. Note that the unravelling $b^\square$ of a sequence $a^\square$ is always in separation normal form. Suppose that a^n has width $m(n)$ and let $\eta^n_0, \dots, \eta^n_{k(n)-1}$ be an enumeration, from left to right, of all elements of $\Lambda(a^\square)$ of length $n+1$. Defining the functions $\sigma_n : [k(n)] \to [m(n)]$ with

$$\sigma_n(i) := \eta^n_i(n),$$

we obtain

$$b^n \sigma_{n+1} = \mu(\eta^n_0) \oplus \dots \oplus \mu(\eta^n_{k(n)-1}) = \sigma_n a^n, \quad \text{for all } n < \omega.$$

Definition 3.2. (a) An ω*-hyperclone* is an expansion of a hyperclone $\mathfrak{C}$ by (a family of) ω-ary operations

$$\pi : \prod_{n<\omega} C_{\bar u^n} \to C_{\bar v},$$

for all $\bar v, \bar u^0, \bar u^1, \dots$ such that

- $\operatorname{supp}(\bar u^n) \subseteq [m_{n+1}]$, where m_{n+1} is the width of $\bar u^{n+1}$, and
- $\bar v = \langle \varnothing, \dots, \varnothing \rangle$ is the tuple of the same width as $\bar u^0$ consisting entirely of empty tuples $\varnothing : \varnothing \to \omega$.

We require $\mathfrak{C}$ to satisfy the following conditions, for every sequence $a^\square$:

(11) For all $k_0 < k_1 < k_2 < \dots < \omega$, we have the associative laws

$$\pi(a^0, a^1, a^2, \dots) = a^0 \cdot \pi(a^1, a^2, \dots),$$
$$\pi(a^0, a^1, a^2, \dots) = \pi((a^0 \cdots a^{k_0-1}), (a^{k_0} \cdots a^{k_1-1}), (a^{k_1} \cdots a^{k_2-1}), \dots).$$

(12) Let $b^\square$ be the factorisation of $a^\square$ induced by some $H \subseteq \Lambda(a^\square)$. Then

$$\pi(b^\square) = \pi(a^\square).$$

(13) Let $m(n)$ be the width of a^n and set $b^n := \pi(a^n, a^{n+1}, \dots)$. For all $I_n \subseteq [m(n)]$, we have

$$\pi(a^0, a^1, \dots) = \pi(a^0 :_{I_1} b^1,\ a^1 :_{I_2} b^2,\ a^2 :_{I_3} b^3, \dots).$$

(14) Let $m(n)$ be the width of a^n. If $\sigma_n : [m(n)] \to [m(n)]$, $n < \omega$, are functions such that

$$\sigma_n \pi(a^n, a^{n+1}, \dots) = \pi(a^n, a^{n+1}, \dots),$$

then

$$\pi(\sigma_0 a^0, \sigma_1 a^1, \sigma_2 a^2, \dots) = \pi(a^0, a^1, a^2, \dots).$$

(15) For all sequences $a^\square$ and $b^\square$,

$$a^n \leq b^n, \text{ for all } n < \omega, \quad \text{implies} \quad \pi(a^0, a^1, \dots) \leq \pi(b^0, b^1, \dots).$$

Remark. Note that, if $b^\square$ is the unravelling of $a^\square$, then (12) implies that $\pi(b^\square) = \pi(a^\square)$. More generally, if $a^\square$ and $b^\square$ are sequences such that $b^n \sigma^{n+1} = \sigma^n a^n$, for suitable functions σ^n, $n < \omega$, then $\pi(b^\square) = \sigma^0 \pi(a^\square)$, since the sequences $b^\square$ and $\sigma^0 a^0, a^1, a^2, \dots$ have the same unravelling.

Remark. Every ω-semigroup $\langle S, S_\omega \rangle$ can be turned into an ω-hyperclone with domains

$$C_{\langle u \rangle} := \begin{cases} S_\omega & \text{if } u = \langle\rangle, \\ S & \text{if } u = \langle k \rangle, \\ \varnothing & \text{otherwise}, \end{cases}$$
$$C_{\langle u_0, \dots, u_{m-1} \rangle} := C_{\langle u_0 \rangle} \times \dots \times C_{\langle u_{m-1} \rangle}, \quad \text{for } m \neq 1.$$

The definition of $\oplus$ and λ_σ is canonical; $:_{I,\tau}$ and π correspond to the products of the ω-semigroup.

Morphisms between ω-hyperclones are defined in the obvious way. We conclude this section by showing that the free ω-hyperclone $\mathfrak{F}_\omega[\Sigma]$ defined above really is free.

Lemma 3.3. *$\mathfrak{F}_\omega[\Sigma]$ is the free ω-hyperclone generated by Σ.*

Proof. To show that $\mathfrak{F}_\omega[\Sigma]$ is free, consider an ω-hyperclone $\mathfrak{C}$ and a sort-preserving function $f : \Sigma \to C$. We have to extend f to a morphism $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$.

We have already seen in Lemma 2.12 that we can extend f to a function $\varphi_0 : F[\Sigma] \to C$ that is a morphism of hyperclones. Hence, it remains to define $\varphi(t)$ for infinite terms t and to show that it is also a morphism of ω-hyperclones.

We define $\varphi(t)$ by induction on the number of variables in t. If t does not contain a variable, we set

$$\varphi(t) := \pi(a^0, a^1, \dots)\,,$$

where the a^n are defined as follows. Let $v^n_0, \dots, v^n_{m(n)-1}$ be an enumeration (in lexicographic order) of all vertices $v \in \operatorname{dom}(t)$ of length $|v| = n$, and set $\sigma^n_k := t(v^n_k)$. We define

$$a^n := \operatorname{sep}(f(\sigma^n_0) \oplus \dots \oplus f(\sigma^n_{m(n)-1}))\,.$$

For an infinite term t with variables, we factorise t as

$$t = s :_{I,\sigma} \bar{t}'$$

where s is a finite term and $\bar{t}'$ is a tuple of infinite terms with fewer variables than t. We define

$$\varphi(t) := \varphi(s) :_{I,\sigma} \varphi(\bar{t}')\,.$$

It remains to check that φ is a morphism of ω-hyperclones. We only give a sketch for the proof that φ commutes with infinite products. Let $\bar{t}^n \in F_{\bar{u}_n}[\Sigma]$, $n < \omega$, be a sequence such that the product $\bar{t}^\omega := \pi(\bar{t}^\square)$ is defined. We have to show that

$$\varphi(\bar{t}^\omega) = \pi(\varphi(\bar{t}^0), \varphi(\bar{t}^1), \dots)\,.$$

Note that, by definition of φ,

$$\begin{aligned}\varphi(\bar{t}^\omega) &= \varphi(t^\omega_0) \oplus \dots \oplus \varphi(t^\omega_{m-1})\,,\\ \varphi(\bar{t}^0) &= \varphi(t^0_0) \oplus \dots \oplus \varphi(t^0_{m-1})\,,\end{aligned}$$

where $\bar{t}^\omega = t^\omega_0 \oplus \dots \oplus t^\omega_{m-1}$ and $\bar{t}^0 = t^0_0 \oplus \dots \oplus t^0_{m-1}$ are the decompositions of $\bar{t}^\omega$ and $\bar{t}^0$, respectively. Since for all suitable $a, b, c^1, c^2, \dots$, it follows by (I1) and (S2) that

$$\begin{aligned}\pi(a \oplus b, c^1, c^2, \dots) &= (a \oplus b) \cdot \pi(c^1, c^2, \dots)\\ &= a \cdot \pi(c^1, c^2, \dots) \oplus b \cdot \pi(c^1, c^2, \dots)\\ &= \pi(a, c^1, c^2, \dots) \oplus \pi(b, c^1, c^2, \dots)\,,\end{aligned}$$

it is therefore sufficient to prove the claim for elements $\bar{t}^\omega = t^\omega$ of width 1.

Let $\bar{s}^\square$ be the unravelling of $\bar{t}^\square$. Then $(\varphi(\bar{s}^n))_n$ is the unravelling of $(\varphi(\bar{t}^n))_n$ and, by (I2), we have

$$\begin{aligned}\pi(\bar{s}^\square) &= \pi(\bar{t}^\square) = t^\omega\,,\\ \text{and}\quad \pi(\varphi(\bar{s}^0), \varphi(\bar{s}^1), \dots) &= \pi(\varphi(\bar{t}^0), \varphi(\bar{t}^1), \dots)\,.\end{aligned}$$

Replacing $\bar{t}^\square$ by $\bar{s}^\square$ we may therefore assume w.l.o.g. that the sequence $\bar{t}^\square$ is in separation normal form.

Note that

$$t^\omega = \pi(a^\square) \quad \text{and} \quad \varphi(t^\omega) = \pi(\varphi(a^0), \varphi(a^1), \dots)\,,$$

where, as above,

$$a^n := \operatorname{sep}(t^\omega(v^n_0) \oplus \dots \oplus t^\omega(v^n_{m(n)-1}))\,,$$

for an enumeration $v^n_0, \dots, v^n_{m(n)-1}$ of all vertices of $\operatorname{dom}(t^\omega)$ of length n. Since the term t^ω is the product of the $\bar{t}^n$, there exists, for every term t^n_i, an embedding $\operatorname{dom}(t^n_i) \to \operatorname{dom}(t^\omega)$ such that the images of these embeddings form a partition of $\operatorname{dom}(t^\omega)$. Furthermore, there is a canonical isomorphism $\operatorname{dom}(t^\omega) \cong \Lambda(a^\square)$. Let $H \subseteq \Lambda(a^\square)$ be the set of all vertices that correspond to the root of some t^n_i under the corresponding embedding. Then $\bar{t}^\square$ is the factorisation of $a^\square$ induced by H, and $(\varphi(a^n))_n$ is the factorisation of $(\varphi(t^n))_n$ induced by H. By (I2), it follows that

$$\varphi(t^\omega) = \pi(\varphi(a^0), \varphi(a^1), \dots) = \pi(\varphi(\bar{t}^0), \varphi(\bar{t}^1), \dots)\,. \qquad \square$$

We are interested in using ω-hyperclones to recognise languages of infinite terms. In the following sections we will isolate a class of finitary ω-hyperclones that recognise exactly the regular languages.

Definition 3.4. Let $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ be a morphism of ω-hyperclones. A subset $L \subseteq F_{\langle\varnothing\rangle}[\Sigma]$ is *recognised* by φ if there exists some set $P \subseteq C_{\langle\varnothing\rangle}$ such that $L = \varphi^{-1}[P]$. Similarly, we say that L is recognised by $\mathfrak{C}$ if there exists a morphism $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ recognising L.

Let us give an example of an ω-hyperclone recognising a language which will be used in the proof of Theorem 6.9 below.

Lemma 3.5. *Let $\Gamma \subseteq \Sigma$ be signatures. Then $L := F_{\langle\varnothing\rangle}[\Gamma]$, regarded as a subset of $F_{\langle\varnothing\rangle}[\Sigma]$, is recognised by a finitary ω-hyperclone.*

Proof. Let us give some intuition first. For each tree, we only need to compute one bit of information: whether or not it contains a symbol from $\Sigma \setminus \Gamma$. Hence, we map an n-tuple $\bar{t}$ of trees to an n-tuple $\bar{b}$ of bits where $b_i = 1$ if, and only if, t_i only contains symbols from Γ.

Consequently, we use the following ω-hyperclone $\mathfrak{C}$. The domain of sort $\bar{u}$ is

$$C_{\bar{u}} := \{0,1\}^n \quad \text{where} \quad n := |\bar{u}|\,.$$

For $\bar{a} \in C_{\bar{u}}$ of width m and $\bar{b} \in C_{\bar{v}}$ of width n, we define the operations as follows:

$$\begin{aligned} &\bar{a} \oplus \bar{b} := \bar{a}\bar{b}\,, \quad 0 := \langle\rangle\,, \quad \lambda_\sigma(\bar{a}) := \bar{a}^\sigma\,, \\ &\bar{a} :_{I,\sigma} \bar{b} := \bar{c}\,, \qquad \text{where} \quad c_i := \min\big(\{a_i\} \cup \{\, b_k \mid k \in I \cap \operatorname{supp}(u_i)\,\}\big)\,. \end{aligned}$$

The infinite product is defined as

$$\pi(\bar{a}^0, \bar{a}^1, \bar{a}^2, \dots) := \bar{b}\,, \quad \text{where} \quad b_i := \lim_{n\to\infty} (a_i^0 \cdot \bar{a}^1 \cdot \bar{a}^2 \cdot \dots \cdot \bar{a}^n)\,.$$

The morphism $\varphi : \mathfrak{F}_\omega[\Gamma] \to \mathfrak{C}$ recognising L is defined by

$$\varphi(t) := \begin{cases} 1 & \text{if } t \in F_{\langle\varnothing\rangle}[\Gamma]\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

and $\varphi(\bar{t}) := \langle \varphi(t_0), \dots, \varphi(t_{n-1})\rangle$. It follows that $L = \varphi^{-1}(1)$. □

4 Path-hyperclones

In this section we prove our first characterisation result. We define path-hyperclones and we show that they are equivalent to parity automata. Intuitively, a path-hyperclone is an ω-hyperclone where the elements are (tuples of) trees whose edges are labelled by elements of a given ω-semigroup $\langle S, S_\omega\rangle$. We label a term t with n variables by an n-tuple $\bar{s} \in S^n$ of semigroup elements, one for each path from the root to a variable. Furthermore, we label the infinite paths of t by elements of S_ω. These are collected in a set $P \subseteq S_\omega$. Hence, the labelling of t is a pair $(\bar{s}, P) \in S^n \times \wp(S_\omega)$. In fact, with each term t we associate a *set* of possible labellings. Hence, the actual labels are sets in $\wp(S^n \times \wp(S_\omega))$.

Definition 4.1. Let $\mathfrak{S} = \langle S, S_\omega\rangle$ be an ω-semigroup. We associate with $\mathfrak{S}$ an ω-hyperclone $\mathfrak{C}$ called the *path-hyperclone* of $\mathfrak{S}$. The domains are

$$\begin{aligned} &C_{\langle u\rangle} := \wp(S^n \times \wp(S_\omega))\,, && \text{for } u : [n] \to \omega \text{ of width } 1, \\ \text{and} \quad &C_{\bar{u}} := C_{u_0} \times \dots \times C_{u_{m-1}}\,, && \text{for } \bar{u} = u_0 \dots u_{m-1} \text{ of width } m \neq 1. \end{aligned}$$

Horizontal composition $\bar{a} \oplus \bar{b} := \bar{a}\bar{b}$, the constant $0 := \langle\rangle$, and the action $\lambda_\sigma(\bar{a}) := \bar{a}^\sigma$ are defined canonically. For $a \in C_{\langle u\rangle}$ and $\bar{b} \in C_{\bar{v}}$ with $u : [n] \to \omega$ of width 1, we define the substitution operation

$$a :_{I,\sigma} \bar{b} := \Big\{ \big(\operatorname{cat}(\bar{z}_0, \dots, \bar{z}_{n-1}),\ P \cup \textstyle\bigcup_{i : u(i) \in I} s_i \cdot Q_i\big) \ \Big|\ (\bar{s}, P) \in a,\ (\bar{t}^i, Q_i) \in b_{u(i)},\ \bar{z}_i := \begin{cases} s_i \cdot \bar{t}^i & \text{if } u(i) \in I\,, \\ \langle s_i\rangle & \text{if } u(i) \notin I\,. \end{cases} \Big\}\,.$$

($s_i \cdot \bar{t}^i$ and $s_i \cdot Q_i$ refer to the canonical action of S on, respectively, $S^{<\omega}$ and $\wp(S)$.) We extend this definition to tuples $\bar{a} \in C_{\bar{u}}$ and $\bar{b} \in C_{\bar{v}}$ by setting

$$\bar{a} :_{I,\sigma} \bar{b} := \langle a_0 :_{I,\sigma} \bar{b}, \dots, a_{|\bar{u}|-1} :_{I,\sigma} \bar{b}\rangle\,.$$

We define an order on $C_{\bar{u}}$ by setting

$$\bar{a} \leq \bar{b} \quad :\text{iff} \quad a_i \subseteq b_i\,, \quad \text{for all } i\,.$$

To define the infinite product consider a sequence $\bar{a}^\square$ in C. Let $\mu : \Lambda(a^\square) \to C$ be the function from Definition 3.1. A *run* on $a^\square$ is a function

$$\chi : \Lambda(a^\square) \to S^{<\omega} \times \wp(S_\omega) \quad \text{such that} \quad \chi(\eta) \in \mu(\eta)\,, \quad \text{for all } \eta\,.$$

Let χ be a run on $\bar{a}^\square$ and let $\eta \in \Lambda(a^\square)$ be a vertex with $|\eta| = k$. We denote the prefix of η of length n by η_n. Suppose that η_{n+1} is the d_n-th successor of η_n in $\Lambda(a^\square)$, and let $\chi(\eta_n) = (\bar{s}^n, P_n)$. We set

$$\Pi_\chi(\eta) := \{\, s^0_{d_0} \cdots s^{k-1}_{d_{k-1}} w \mid w \in P_k \,\}\,.$$

For infinite branches $\beta = (\eta_n)_{n<\omega}$ of $\Lambda(a^\square)$ where η_{n+1} is the d_n-th successor of η_n and $\chi_k(\eta_n) = (\bar{s}^n, P_n)$, we set

$$\pi_\chi(\beta) := \pi(s^0_{d_0}, s^1_{d_1}, \dots)\,.$$

Finally, we set

$$\Pi(\chi) := \bigcup_{\eta \in \Lambda(a^\square)} \Pi_\chi(\eta) \cup \{\, \pi_\chi(\beta) \mid \beta \text{ an infinite branch of } \Lambda(a^\square) \,\}\,.$$

We define the infinite product as follows. Suppose that $\bar{a}^0 = \langle a^0_0, \dots, a^0_{m-1} \rangle$ has width m. We set

$$\pi(\bar{a}^0, \bar{a}^1, \dots) := \langle b_0, \dots, b_{m-1} \rangle\,,$$

where

$$b_i := \big\{\, (\langle\rangle, \Pi(\chi)) \;\big|\; \chi \text{ a run on } a^0_i, \bar{a}^1, \bar{a}^2, \dots \big\}\,.$$

With each automaton we can associate a corresponding path-hyperclone.

Definition 4.2. Suppose that $\mathcal{A} = \langle Q, \Sigma, \Delta, q_0, \Omega \rangle$ is a parity automaton and set $D := \operatorname{rng} \Omega$. The *hyperclone* $\mathfrak{C}_\mathcal{A}$ *for* $\mathcal{A}$ is the path-hyperclone associated with the ω-semigroup $\langle S, S_\omega \rangle$ where

$$S := Q \times D \times Q \cup \{\bot\} \quad \text{and} \quad S_\omega := Q \cup \{\bot\}\,.$$

The multiplication $S \times S \to S$ is defined by

$$(p,d,q) \cdot (p',d',q') := \begin{cases} (p, \min\{d,d'\}, q') & \text{if } q = p'\,, \\ \bot & \text{otherwise}\,. \end{cases}$$

$$s \cdot \bot = \bot \cdot s := \bot\,, \quad \text{for all } s \in S\,.$$

The multiplication $S \times S_\omega \to S_\omega$ is defined by

$$(p,d,q) \cdot r := \begin{cases} p & \text{if } q = r \neq \bot\,, \\ \bot & \text{otherwise}\,. \end{cases}$$

$$\bot \cdot r := \bot\,.$$

For $s_n = (p_n, d_n, q_n) \in S$, $n < \omega$, we define the infinite product by

$$\pi(s_0, s_1, \dots) := \begin{cases} p_0 & \text{if } q_n = p_{n+1}\,, \text{ for all } n, \text{ and } \liminf_{n\to\infty} d_n \text{ is even}\,, \\ \bot & \text{otherwise}\,. \end{cases}$$

If some $s_n = \bot$ then we set

$$\pi(s_0, s_1, \dots) := \bot\,.$$

The following two theorems show that path-hyperclones recognise the same languages as parity automata.

Theorem 4.3. *For every automaton $\mathcal{A}$, there exists a morphism $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ into a path-hyperclone $\mathfrak{C}$ associated with a finite ω-semigroup such that φ recognises $L(\mathcal{A})$.*

Proof. Suppose that $\mathcal{A} = \langle Q, \Sigma, \Delta, q_0, \Omega \rangle$ is a nondeterministic parity automaton and let $\mathfrak{C} := \mathfrak{C}_\mathcal{A}$ be the hyperclone for $\mathcal{A}$. We define the morphism φ as follows. Recall that $\mathfrak{F}_\omega[\Sigma]$ is freely generated by Σ. Hence, it is sufficient to define φ for elements of Σ. If $\sigma \in \Sigma$ is of arity n, we set

$$\varphi(\sigma) := \big\{\, (\bar{s}, \varnothing) \in S^n \times \wp(S_\omega) \;\big|\; \text{there is some } (p, \sigma, \bar{q}) \in \Delta \text{ with } s_i := (p, \Omega(q_i), q_i) \text{ for all } i \,\big\}\,.$$

We claim that $L(\mathcal{A}) = \varphi^{-1}[P]$ where

$$P := \big\{\, X \in C_{\langle\rangle} \;\big|\; (\langle\rangle, \{q_0\}) \in X \,\big\}\,.$$

Let $t \in T^0_\omega[\Sigma]$ be a tree and let $v^n_0, \dots, v^n_{m(n)-1}$ be an enumeration from left to right of all vertices $v \in \operatorname{dom}(t)$ of length $|v| = n$. Setting $a^n_i := \varphi(t(v^n_i))$ it follows that

$$\varphi(t) = \pi(a^0, a^1, \dots) \quad \text{where} \quad a^n := \operatorname{sep}(a^n_0 \oplus \dots \oplus a^n_{m(n)-1})\,.$$

Furthermore, note that the trees $\operatorname{dom}(t)$ and $\Lambda(a^\square)$ are isomorphic. Let

$$\mu : \operatorname{dom}(t) \to \Lambda(a^\square)$$

be the corresponding isomorphism. We have to show that $\varphi(t) \in P$ iff $t \in L(\mathcal{A})$.

($\Leftarrow$) Let $\rho : \mathrm{dom}(t) \to Q$ be an accepting run of $\mathcal{A}$ on t. We obtain a run χ on $(\varphi(a^n))_n$ as follows. Let $\eta \in \Lambda(a^\square)$ be a vertex with successors $\zeta_0, \dots, \zeta_{k-1}$ and suppose that $\rho(\mu^{-1}(\eta)) = p$ and $\rho(\mu^{-1}(\zeta_i)) = q_i$. We set

$$\chi(\eta) := (\bar{s}, \varnothing) \quad \text{where} \quad s_i := (p, \Omega(q_i), q_i)\,, \text{ for all } i < k\,.$$

χ is a run on $a^\square$. Since ρ is accepting it follows that $\Pi(\chi) = \{q_0\}$. Hence, $\varphi(t) \in P$.

($\Rightarrow$) Since $\varphi(t) \in P$ there is some run χ on $a^\square$ with $\Pi(\chi) = \{q_0\}$. Let $w \in \mathrm{dom}(t)$ and suppose that $\chi(\mu(w)) = (\bar{s}, \varnothing)$. There are states $p, q_0, q_1, \dots \in Q$ such that $s_i = (p, \Omega(q_i), q_i)$. We set

$$\rho(w_n) := p\,.$$

Then ρ is a run of $\mathcal{A}$ on t. It is accepting by choice of χ. ☐

The proof of the converse result is split into several lemmas.

Lemma 4.4. *Let $\mathfrak{S} = \langle S, S_\omega \rangle$ be a finite ω-semigroup. For every $u \in S_\omega$, there exists an ω-automaton $\mathcal{A}_u$ recognising the language*

$$L(\mathcal{A}_u) := \{\, (a_n)_n \in S^\omega \mid \pi(a_0, a_1, \dots) = u \,\}\,.$$

Proof. Let $(a_n)_n \in S^\omega$ be the input word. By the Theorem of Ramsey, we can find an increasing sequence of indices $k_0 < k_1 < \dots < \omega$ such that

$$a_{k_i} \dots a_{k_{i+1}-1} = a_{k_j} \dots a_{k_{j+1}-1}\,, \quad \text{for all } i, j < \omega\,.$$

Hence, the automaton $\mathcal{A}_u$ can check that $\pi(a_0, a_1, \dots) = u$ by guessing two elements $s, e \in S$ with $se^\omega = u$ and then checking that there are indices $k_0 < k_1 < \dots < \omega$ such that

$$a_0 \cdots a_{k_0-1} = s \quad \text{and} \quad a_{k_i} \cdots a_{k_{i+1}-1} = e\,, \quad \text{for all } i < \omega\,. \qquad ☐$$

Lemma 4.5. *Let $\mathfrak{C}$ be the path-hyperclone associated with a finite ω-semigroup $\langle S, S_\omega \rangle$ and let $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ be a homomorphism. For every $Q \subseteq S_\omega$, there exists an automaton $\mathcal{A}_Q$ such that*

$$(\langle\rangle, Q) \in \varphi(t) \quad \text{iff} \quad t \in L(\mathcal{A}_Q)\,, \quad \textit{for all trees } t \in F_{\langle\varnothing\rangle}[\Sigma]\,.$$

Proof. Let $t \in F_{\langle\varnothing\rangle}[\Sigma]$ be a tree and let $v_0^n, \dots, v_{m(n)-1}^n$ be an enumeration from left to right of all vertices $v \in \mathrm{dom}(t)$ of length $|v| = n$. Setting $a_i^n := \varphi(t(v_i^n))$ it follows that

$$\varphi(t) = \pi(a^0, a^1, \dots) \quad \text{where} \quad a^n := \mathrm{sep}(a_0^n \oplus \dots \oplus a_{m(n)-1}^n)\,.$$

As in the proof of Theorem 4.3 there is an isomorphism

$$\mu : \Lambda(a^\square) \to \mathrm{dom}(t)\,.$$

If χ is a run on $a^\square$, then we have

$$\chi(\eta) \in \varphi(t(\mu(\eta)))\,, \quad \text{for all } \eta\,.$$

Consequently, our automaton $\mathcal{A}_Q$ can guess a run χ by guessing a labelling $\chi : \mathrm{dom}(t) \to C$ and checking that $\chi(v) \in \varphi(t(v))$, for all $v \in \mathrm{dom}(t)$. Having done so, it uses Lemma 4.4 to verify that

- $\Pi_\chi(v) \subseteq Q$, for all vertices v of χ;
- $\pi_\chi(\beta) \in Q$, for every infinite branch β of χ;
- for every $s \in Q$, there is some vertex v of χ with $s \in \Pi_\chi(v)$, or there is an infinite branch β of χ with $\pi_\chi(\beta) = s$. ☐

Theorem 4.6. *Let $\mathfrak{C}$ be a path-hyperclone associated with a finite ω-semigroup. For every homomorphism $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ into $\mathfrak{C}$ and for every subset $P \subseteq C_{\langle\varnothing\rangle}$, there exists an automaton $\mathcal{A}$ recognising the set*

$$L(\mathcal{A}) = \varphi^{-1}[P]\,.$$

Proof. We use the automata $\mathcal{A}_Q$ provided by the preceding lemma. Given an input tree t, the automaton $\mathcal{A}$ guesses some $a \in P$ and checks that

- $\mathcal{A}_Q$ accepts t, for every Q with $(\langle\rangle, Q) \in a$, and
- $\mathcal{A}_Q$ does not accept t, for every Q with $(\langle\rangle, Q) \notin a$. ☐

5 Path-continuity

In Theorem 4.3 and 4.6, we have characterised the regular tree languages as those that are recognised by morphisms into path-hyperclones. This result is somewhat unsatisfactory since the definition of a path-hyperclone is not 'algebraic' in the sense that path-hyperclones are not closed under isomorphisms. In particular, there is no known axiomatic characterisation of path-hyperclones, say, via a set of equations. Therefore, we define a second class of ω-hyperclones (this time by axioms that are invariant under isomorphisms) and we show that these ω-hyperclones can also be used to characterise regular tree languages.

Our approach is as follows. We isolate several special properties of path-hyperclones and we show that these properties are sufficient for constructing an automaton that evaluates infinite products. In the proof of Lemma 4.5 the main ingredient was the fact that an infinite product in a path-hyperclone basically reduces to a family of infinite products along branches of the term. Therefore, we will consider ω-hyperclones where we can evaluate an infinite product by considering each branch separately.

The idea is as follows. Given a sequence $a^\square$ we would like to compute the product $\pi(a^\square)$ by guessing the values $b^n := \pi(a^n, a^{n+1}, a^{n+2}, \dots)$, for all $n < \omega$, and then checking that our guess was correct. A necessary condition for $(b^n)_{n<\omega}$ are the equations $b^n = a^n b^{n+1}$. In general, this condition is not sufficient, even for path-hyperclones. But adding an additional consistency condition on every branch of $a^\square$, we obtain a condition that is sufficient for path-hyperclones.

Definition 5.1. Let $\mathfrak{C}$ be an ω-hyperclone and $J \subseteq C$ a subset.

(a) A sequence $(b^n)_{n<\omega}$ where $b^n \in C_{\langle\varnothing,\dots,\varnothing\rangle}$ is *locally consistent* with $a^\square$ if $b_n = a_n b_{n+1}$, for every $n < \omega$. We denote by $\mathrm{LC}_J(a^\square)$ the set of all locally consistent sequences $(b^n)_{n<\omega}$ with $b^n \in J$, for all n.

(b) Let $a^\square$ be a sequence where a^n has the decomposition $a^n = a^n_0 \oplus \dots \oplus a^n_{m(n)-1}$, let $(b^n)_n \in \mathrm{LC}_J(a^\square)$ be locally consistent with $a^\square$, and let β be a branch of $a^\square$. The *trace* of $(b^n)_n$ along β is the sequence $c^\square$ with

$$c^n := \mathrm{sep}(a^n_{\beta(n)} :_{I_n} b^{n+1}),$$

where

$$I_n := \begin{cases} [m(n+1)] \smallsetminus \{\beta(n+1)\} & \text{if } n+1 \in \mathrm{dom}(\beta), \\ [m(n+1)] & \text{otherwise}. \end{cases}$$

We set

$$\begin{aligned} \mathrm{Tr}_J(\beta) &:= \{ \pi(c^\square) \mid c^\square \text{ the trace of some } (b^n)_n \in \mathrm{LC}_J(a^\square) \text{ along } \beta \}, \\ \mathrm{BT}_J(a^\square) &:= \{ \mathrm{Tr}_J(\beta) \mid \beta \text{ a branch of } \Lambda(a^\square) \}. \end{aligned}$$

Remark. Let $c^\square$ be the trace of $(b^n)_n$ over $a^\square$ along the branch β and let $m(n)$ be the width of a^n. Each element c^n has width 1. Setting

$$\begin{aligned} \sigma_n &: [m(n+1)] \to [1] : x \mapsto 0, \\ \tau_n &: [1] \to [m(n)] : 0 \mapsto \beta(n), \end{aligned}$$

we have

$$\begin{aligned} \pi(c^\square) &= \pi\big(\tau_0(a^0 :_{I_0,\sigma_0} b^1),\ \tau_1(a^1 :_{I_1,\sigma_1} b^2), \dots\big) \\ &= \tau_0\pi\big((a^0 :_{I_0,\sigma_0} b^1)\tau_1,\ (a^1 :_{I_1,\sigma_1} b^2)\tau_2, \dots\big) \\ &= \tau_0\pi\big(a^0 :_{I_0} b^1,\ a^1 :_{I_1} b^2, \dots\big). \end{aligned}$$

In particular, for the special case that $b^n = \pi(a^n, a^{n+1}, a^{n+2}, \dots)$, we have

$$\pi(c^\square) = \tau_0\pi(a^\square).$$

Note that, for $J = C$, there is at least one sequence $(b^n)_n \in \mathrm{LC}_J(a^\square)$: the sequence with $b^n = \pi(a^n, a^{n+1}, a^{n+2}, \dots)$. Unfortunately, in most cases this 'real' sequence is not the only one. Let us mention two simple cases, where we always have a unique locally consistent sequence.

- Every sequence $a^\square$ in the free ω-hyperclone $\mathfrak{F}_\omega[\Sigma]$ has this property.
- A sequence $a^\square$ in an arbitrary ω-hyperclone has this property if the branch tree $\Lambda(a^\square)$ is finite. (Hence, the product $\pi(a^\square)$ is not truly infinite.)

In the general case, we thus face the problem of singling out the 'real' sequence from among the sequences in $\mathrm{LC}_J(a^\square)$. We will introduce a class of ω-hyperclones, where this can be done by checking the branches separately. To obtain the precise definitions we will first take a look at path-hyperclones. We start with collecting some technical results on infinite products in path-hyperclones.

Definition 5.2. Let $\mathfrak{C}$ be a path-hyperclone. An element $a \in C_u$ of width 1 is *subminimal* if $|a| \leq 1$. An element $a = a_0 \oplus \dots \oplus a_{n-1} \in C_{\bar u}$ of width $n > 1$ is *subminimal* if every component a_i is subminimal. Finally, we call a sequence $a^\square$ of elements *subminimal* if every a^n is subminimal.

Lemma 5.3. *Let $\mathfrak{C}$ be a path-hyperclone and $\bar{a}^\square$ a subminimal sequence where $\bar{a}^0$ has width 1, and let $(\bar{b}^n)_n \in \mathrm{LC}_J(\bar{a}^\square)$, where J is the set of all subminimal elements.*

(a) *There is at most one run on $\bar{a}^\square$.*

(b) $|\pi(\bar{a}^\square)| \leq 1$.

(c) *There exists a run on $\bar{a}^\square$ if, and only if, $\pi(\bar{a}^\square) \neq \emptyset$.*

(d) *If $\bar{c}^\square$ is the trace of $(\bar{b}^n)_n$ over $\bar{a}^\square$ along some branch β then $\bar{c}^\square$ is also subminimal.*

(e) *If $b^0_0 \neq \emptyset$ and if $\bar{c}^\square$ is the trace of $(\bar{b}^n)_n$ over $\bar{a}^\square$ along some branch β, then $\pi(\bar{c}^\square) \neq \emptyset$.*

Proof. (a) holds since $|a^n_k| \leq 1$, for all n and k.

(b),(c) By (a), there is at most one run χ on $(\bar{a}^n)_n$. Consequently, $\pi(\bar{a}^0, \bar{a}^1, \dots)$ is either empty or $\pi(\bar{a}^0, \bar{a}^1, \dots) = \{(\langle\rangle, \Pi(\chi))\}$ is a singleton.

(d) follows by the definition of a trace.

(e) Let $I := \mu_0[\Lambda(a^\square)]$ where $\mu_0 : \Lambda(a^\square) \to \omega \times \omega$ is the function from Definition 3.1. We claim that

$$a^n_i, b^n_i, c^n_i \neq \emptyset, \quad \text{for all } \langle n, i\rangle \in I .$$

We start by proving, by induction on n, that $b^n_i \neq \emptyset$. For $n = 0$, we have $b^0_0 \neq \emptyset$, by assumption. Hence, suppose that $n > 0$. Let $\eta \in \Lambda(a^\square)$ be an element with $\mu_0(\eta) = \langle n, i\rangle$ and let η_0 be the predecessor of η. Then $\mu_0(\eta_0) = \langle n-1, k\rangle$, for some k. By induction hypothesis, we have $b^{n-1}_k \neq \emptyset$. Hence, $b^{n-1}_k = a^{n-1}_k \cdot \bar{b}^n$ implies $b^n_i \neq \emptyset$, as desired.

The two remaining claims follow. For a^n_i, note that

$$b^n_i \neq \emptyset \quad \text{and} \quad b^n_i = a^n_i \cdot \bar{b}^{n+1} \quad \text{implies} \quad a^n_i \neq \emptyset .$$

It follows that $c^n_i = \mathrm{sep}(a^n_i :_{I_n} \bar{b}^{n+1}) \neq \emptyset$, for $\langle n, i\rangle \in I$. Consequently, $\pi(\bar{c}^\square) \neq \emptyset$. ☐

One important property of path-hyperclones is the fact that the traces of subminimal sequences uniquely determine the value of an infinite product.

Lemma 5.4. *Let $\mathfrak{C}$ be a path-hyperclone, $J \subseteq C$ the set of all subminimal elements, and $\bar{a}^\square$ a subminimal sequence where $\bar{a}^0$ has width 1.*

(a) *$\pi(\bar{a}^\square) \in \mathrm{Tr}_J(\beta)$, for every branch β.*

(b) *Suppose that $\pi(\bar{a}^\square) = \{(\langle\rangle, P)\}$. For every $s \in P$, there exists some branch β of $\Lambda(\bar{a}^\square)$ such that $s \in Q$, for every set Q with $\{(\langle\rangle, Q)\} \in \mathrm{Tr}_J(\beta)$.*

(c) *If $\pi(\bar{a}^\square) = \emptyset$, there exists a branch β of $\Lambda(\bar{a}^\square)$ such that $\mathrm{Tr}_J(\beta) = \{\emptyset\}$.*

(d) *The value of the product $\pi(\bar{a}^\square)$ is uniquely determined by the set $\mathrm{BT}_J(\bar{a}^\square)$.*

Proof. (a) Let $\bar{b}^n := \pi(\bar{a}^n, \bar{a}^{n+1}, \dots)$. By Lemma 5.3 (b), we have $\bar{b}^n \in J$. Hence, $(\bar{b}^n)_n \in \mathrm{LC}_J(\bar{a}^\square)$ and $\pi(\bar{a}^\square) = \bar{b}^0 \in \mathrm{Tr}_J(\beta)$.

(b) By Lemma 5.3 (c), there is a unique run χ on $(\bar{a}^n)_n$. Note that

$$P = \Pi(\chi) = \bigcup_\eta \Pi_\chi(\eta) \cup \{\, \pi_\chi(\beta) \mid \beta \text{ an infinite branch} \,\} .$$

First, suppose that $s \in \Pi_\chi(\eta)$, for some $\eta \in \Lambda(\bar{a}^\square)$ of length k. Let β be a branch of $\Lambda(\bar{a}^\square)$ containing η. We claim that β has the desired properties. Let $\{(\langle\rangle, Q)\} \in \mathrm{Tr}_J(\beta)$. There exists a sequence $(\bar{b}^n)_n \in \mathrm{LC}_J(\bar{a}^\square)$ such that $\pi(\bar{c}^\square) = \{(\langle\rangle, Q)\}$, where $\bar{c}^\square$ is the trace of $(\bar{b}^n)_n$ along β. By Lemma 5.3 (c), there is a unique run χ' on $\bar{c}^\square$. Since $\Pi_\chi(\eta) \subseteq \Pi_{\chi'}(0^k)$, it follows that

$$\pi(\bar{c}^\square) = \{(\langle\rangle, R)\} \quad \text{where} \quad s \in \bigcup_{\zeta \in \Lambda(\bar{c}^\square)} \Pi_{\chi'}(\zeta) \subseteq R .$$

It remains to consider the case that $s = \pi_\chi(\beta)$, for some infinite branch β. We claim that β has the desired properties. Let $\{(\langle\rangle, Q)\} \in \mathrm{Tr}_J(\beta)$. There exists a sequence $(\bar{b}^n)_n \in \mathrm{LC}_J(\bar{a}^\square)$ such that $\pi(\bar{c}^\square) = \{(\langle\rangle, Q)\}$, where $\bar{c}^\square$ is the trace of $(\bar{b}^n)_n$ along β. By Lemma 5.3 (c), there is a unique run χ' on $\bar{c}^\square$. For the unique infinite branch $\gamma = 0^\omega$ of χ', we obtain

$$s = \pi_\chi(\beta) = \pi_{\chi'}(\gamma) \in \Pi(\chi') .$$

(c) Let $\mu : \Lambda(\bar{a}^\square) \to C$ be the function from Definition 3.1. If $\pi(\bar{a}^\square) = \emptyset$, there exists a vertex $\eta \in \Lambda(\bar{a}^\square)$ such that

$$\mu(\eta) = \emptyset \quad \text{and} \quad \mu(\zeta) \neq \emptyset, \quad \text{for all } \zeta \prec \eta .$$

Let β be a branch containing η. For every $(\bar{b}^n)_n \in \mathrm{LC}_J(\bar{a}^\square)$ with trace $\bar{c}^\square$ along β, it follows that $\pi(\bar{c}^\square) = \emptyset$. Hence, $\mathrm{Tr}_J(\beta) = \{\emptyset\}$.

(d) If $\pi(\bar{a}^\square) = \varnothing$, then (c) implies that there is some branch β with $\mathrm{Tr}_J(\beta) = \{\varnothing\}$. Conversely, if $\pi(\bar{a}^\square) = \{(\langle\rangle, P)\}$, then (a) implies that $\{(\langle\rangle, P)\} \in \mathrm{Tr}_J(\beta)$, for every branch β. Therefore, we have

$$\begin{aligned} \pi(\bar{a}^\square) = \varnothing \quad & \text{iff} \quad \text{there is some branch } \beta \text{ with } \mathrm{Tr}_J(\beta) = \{\varnothing\} \\ & \text{iff} \quad \{\varnothing\} \in \mathrm{BT}_J(\bar{a}^\square)\,. \end{aligned}$$

Suppose that $\pi(\bar{a}^\square) = \{(\langle\rangle, P)\} \neq \varnothing$. For a branch β of $\Lambda(\bar{a}^\square)$, we set

$$F(\beta) := \{\, Q \mid \{(\langle\rangle, Q)\} \in \mathrm{Tr}_J(\beta) \,\}\,.$$

We claim that

$$P = \bigcup \{\, \bigcap F(\beta) \mid \beta \text{ a branch} \,\}\,.$$

By (a), we have $\{(\langle\rangle, P)\} \in \mathrm{Tr}_J(\beta)$, for every β, i.e., $P \in F(\beta)$. Consequently, we have $P \supseteq \bigcap F(\beta)$, for every β, which implies that $P \supseteq \bigcup_\beta \bigcap F(\beta)$.

Conversely, it follows by (b) that, for every $s \in P$, there is some branch β with $s \in \bigcap F(\beta)$. Hence, $P \subseteq \bigcup_\beta \bigcap F(\beta)$. □

By the preceding lemma we know how to compute the product of a subminimal sequence. In the next lemma we reduce the computation of arbitrary products to this special case.

Lemma 5.5. *Let $\mathfrak{C}$ be the path-hyperclone associated with an ω-semigroup $\langle S, S_\omega\rangle$, and let $\bar{a}^\square$ be a sequence in separation normal form. Then*

$$\pi(\bar{a}^\square) = \sup\{\, \pi(\bar{b}^\square) \mid \bar{b}^\square \leq \bar{a}^\square \textit{ is subminimal} \,\}\,.$$

Proof. Given a run χ on $\bar{a}^\square$, we define a sequence $\bar{b}^\square_\chi \leq \bar{a}^\square$ such that

$$\pi(\bar{b}^\square_\chi) = \{(\langle\rangle, \Pi(\chi))\}\,.$$

Since each $\bar{a}^n$ is in separation normal form, it follows by induction on $n < \omega$ that, for all $k < \omega$, there is at most one element $\eta \in \Lambda(\bar{a}^\square)$ such that

$$|\eta| = n + 1 \quad \text{and} \quad \eta(n) = k\,.$$

If such an element exists, we denote it by η^n_k. Let

$$(b^n_\chi)_k := \begin{cases} \{\chi(\eta^n_k)\} & \text{if } \eta^n_k \text{ exists,} \\ \varnothing & \text{otherwise.} \end{cases}$$

Then $\bar{b}^\square_\chi \leq \bar{a}^\square$ is subminimal and χ is a run on $\bar{b}^\square_\chi$. By Lemma 5.3 (a), it follows that

$$\pi(\bar{b}^\square_\chi) = \{(\langle\rangle, \Pi(\chi))\} \subseteq \pi(\bar{a}^\square)\,.$$

Conversely, since every element of $\pi(\bar{a}^\square)$ is of the form $(\langle\rangle, \Pi(\chi))$, for some run χ, it follows that $\pi(\bar{a}^\square)$ is covered by the sets $\pi(\bar{b}^\square_\chi)$ corresponding to these runs. □

We use the two properties of path-hyperclones isolated in Lemmas 5.4 and 5.5 to define a subclass of ω-hyperclones that are sufficiently similar to path-hyperclones for our characterisation results to go through.

Definition 5.6. Let $\mathfrak{C}$ be an ω-hyperclone.

(a) An *ideal* of $\mathfrak{C}$ is a sub-ω-hyperclone $\mathfrak{J} \subseteq \mathfrak{C}$ where each domain $J_{\bar{u}}$ is downward closed, i.e., $\mathfrak{J}$ is a substructure of $\mathfrak{C}$ that is an ω-hyperclone such that $a \leq b \in J_{\bar{u}}$ implies $a \in J_{\bar{u}}$.

(b) An ω-hyperclone $\mathfrak{C}$ is *path-continuous* if there exists an ideal $\mathfrak{J}$ such that

- for every sequence $a^\square$ in separation normal form, we have

$$\pi(a^\square) = \sup\{\, \pi(b^\square) \mid b^\square \leq a^\square \text{ a sequence in } J \,\}\,,$$

- and the product $\pi(a^\square)$ of a sequence $a^\square$ in J is uniquely determined by the set $\mathrm{BT}_J(a^\square)$.

Proposition 5.7. *Every finitary path-hyperclone is path-continuous.*

Proof. By Lemmas 5.4 and 5.5, we can take for J the set of all subminimal elements. □

Example. Let us give an example of an ω-hyperclone that is *not* path-continuous. We set

$$C_{\bar{u}} := [4]^n\,, \quad \text{for each sort } \bar{u} \text{ of width } n := |\bar{u}|\,.$$

$\oplus$, o and λ_σ are defined canonically:

$$\bar{a} \oplus \bar{b} := \bar{a}\bar{b}\,, \quad \mathrm{o} = \langle\rangle\,, \quad \lambda_\sigma(\bar{a}) = \bar{a}^\sigma.$$

The substitution operation $:_{I,\sigma}$ is defined as follows. For $a \in C_{\langle u\rangle}$ of width 1 and $\bar{b} \in C_{\bar{v}}$, we set

$$a :_{I,\sigma} \bar{b} := \max\big(\{a\} \cup \{\, b_i \mid i \in I \cap \operatorname{supp}(u) \,\}\big)\,.$$

For a sequence $\bar{a}^{\square}$ where $\bar{a}^{\circ}$ has width 1, we define the infinite product by

$$\pi(\bar{a}^{\square}) := \max\big(\{x\} \cup \{\, \mu(\eta) \mid \eta \in \Lambda(\bar{a}^{\square}) \,\}\big)\,,$$

where $x \in [4]$ is defined as follows. Let us call a branch β of $\Lambda(\bar{a}^{\square})$ *heavy* if β is infinite and, for every prefix $\eta < \beta$, there is some $\zeta \geq \eta$ with $\mu(\zeta) > 0$. We set

$$x := \begin{cases} 3 & \text{if } \Lambda(\bar{a}^{\square}) \text{ has infinitely many heavy branches,} \\ 2 & \text{if } \Lambda(\bar{a}^{\square}) \text{ has heavy branches, but only finitely many,} \\ 0 & \text{if } \Lambda(\bar{a}^{\square}) \text{ does not have heavy branches.} \end{cases}$$

We equip $\mathfrak{C}$ with the natural order $0 < 1 < 2 < 3$. We claim that there is no ideal $(J_{\bar{u}})_{\bar{u}}$ with respect to which $\mathfrak{C}$ is path-continuous.

For a proof, we distinguish two cases. First, suppose that $1 \notin J_{\langle 0,1\rangle}$. Then $J_{\langle 0,1\rangle} = \{0\}$. Consider the sequences $\bar{a}^{\square}_0$ and $\bar{b}^{\square}_0$ where $\bar{a}^n_0 = \langle 0,0\rangle$ and $\bar{b}^n_0 = \langle 1,1\rangle$, are both elements of $C_{\langle 0,1\rangle,\langle 0,1\rangle}$. Let $\bar{a}^{\square}$ and $\bar{b}^{\square}$ be their respective unravellings. Since

$$\pi(\bar{b}^{\square}) = \langle 3,3\rangle > \langle 0,0\rangle = \pi(\bar{a}^{\square})\,,$$

it follows that

$$\pi(\bar{b}^{\square}) \neq \sup\{\, \bar{x}^{\square} \mid \bar{x}^{\square} \leq \bar{a}^{\square} \text{ in } J \,\}\,.$$

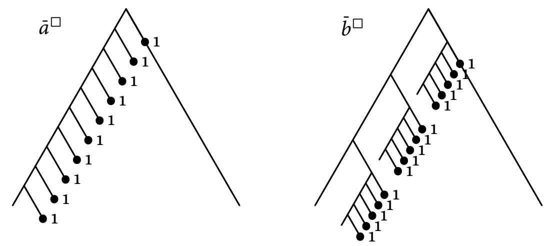
It remains to consider the case that $1 \in J_{\langle 0,1\rangle}$. Let $\bar{a}^{\square}_0$ be the sequence where

$$\bar{a}^0_0 = 0 \in C_{\langle 0,1\rangle} \quad \text{and} \quad \bar{a}^n_0 = \langle 0,1,0\rangle \in C_{\langle 0,1\rangle,\langle 2,2\rangle,\langle 2,2\rangle}\,, \qquad \text{for } n > 0\,,$$

and let $\bar{b}^{\square}_0$ be the sequence where

$$\begin{aligned} \bar{b}^0_0 &= 0 \in C_{\langle 0,1\rangle}\,, \\ \bar{b}^1_0 &= \langle 0,0\rangle \in C_{\langle 0,1\rangle,\langle 2,3\rangle}\,, \\ \bar{b}^n_0 &= \langle 0,0,1,0\rangle \in C_{\langle 0,1\rangle,\langle 1,2\rangle,\langle 3,3\rangle,\langle 3,3\rangle}\,, \qquad \text{for } n > 1\,. \end{aligned}$$

Again, let $\bar{a}^{\square}$ and $\bar{b}^{\square}$ be the unravellings of these sequences.



Then

$$\pi(\bar{a}^{\square}) = 2 \quad \text{and} \quad \pi(\bar{b}^{\square}) = 3\,,$$

but $\mathrm{BT}_J(\bar{a}^{\square}) = \big\{\{2,3\},\ \{1,2,3\}\big\} = \mathrm{BT}_J(\bar{b}^{\square})$.

From the results of the previous section it follows that we can describe each regular tree language via a morphism to some finitary path-continuous ω-hyperclone. To complete the picture it remains to prove the converse.

Theorem 5.8. *Let $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ be a morphism into a finitary path-continuous ω-hyperclone $\mathfrak{C}$. Every set recognised by φ is regular.*

Proof. Let $t \in F_{\langle\varnothing\rangle}[\Sigma]$. We fix enumerations $v^n_0, \dots, v^n_{m(n)-1}$ of all vertices $v \in \operatorname{dom}(t)$ of length $|v| = n$ from left to right. Setting $a^n := \operatorname{sep}(a^n_0 \oplus \dots \oplus a^n_{m(n)-1})$ with $a^n_i := \varphi(t(v^n_i))$, it follows that

$$\varphi(t) = \pi(a^{\square})\,.$$

Note that there is an isomorphism

$$\mu : \Lambda(a^{\square}) \to \operatorname{dom}(t)\,.$$

Using this isomorphism we can encode a sequence $\hat{a}^{\square} \leq a^{\square}$, a sequence $(b^n)_n \in \mathrm{LC}_J(\hat{a}^{\square})$, and a branch β as a labelling of $\operatorname{dom}(t)$. We will denote such labellings as tuples $\langle t, \hat{a}^{\square}\rangle$, $\langle t, \hat{a}^{\square}, \beta\rangle$, or $\langle t, \hat{a}^{\square}, \beta, (b^n)_n\rangle$, etc.

We construct an automaton recognising $\varphi^{-1}[P]$ in several steps. Let J be the ideal of $\mathfrak{C}$ witnessing path-continuity.

(1) For every element $c \in C$ of width 1, we first construct an automaton $\mathcal{B}_c$ accepting those tuples $\langle t, \hat{a}^{\square}, \beta, (b^n)_n\rangle$ such that

- $\hat{a}^{\square} \leq a^{\square}$ is in J;

- β is a branch of $\Lambda(a^\square)$;
- $(b^n)_n \in \mathrm{LC}_J(\hat{a}^\square)$;
- the product of the trace of $(b^n)_n$ along β evaluates to c.

Clearly, the first three conditions are strictly local and can be checked by an automaton. For the last one, we can use the automaton $\mathcal{A}_u$ from Lemma 4.4.

(2) For every set $Q \subseteq C$ of elements of width 1, we construct an automaton $\mathcal{C}_Q$ accepting those triples $\langle t, \hat{a}^\square, \beta \rangle$ such that

- $\hat{a}^\square \leq a^\square$ is in J;
- β is a branch of $\Lambda(a^\square)$;
- $\mathrm{Tr}_J(\beta) = Q$.

Using the automata $\mathcal{B}_c$ from (1), the automaton $\mathcal{C}_Q$ checks that

- for every $c \in Q$, there is some $(b^n)_n$ such that $\mathcal{B}_c$ accepts $\langle t, \hat{a}^\square, \beta, (b^n)_n \rangle$,
- for every $c \notin Q$, there is no $(b^n)_n$ such that $\mathcal{B}_c$ accepts $\langle t, \hat{a}^\square, \beta, (b^n)_n \rangle$.

(3) Having constructed the automata $\mathcal{C}_Q$ we can build an automaton $\mathcal{D}_c$, for $c \in C$, that accepts those pairs $\langle t, \hat{a}^\square \rangle$ such that

- $\hat{a}^\square \leq a^\square$ is a sequence in J;
- $\pi(\hat{a}^\square) = c$.

Since $\mathfrak{C}$ is path-continuous, there exists a set $F_c \subseteq \wp(\wp(C))$ such that

$$\pi(\hat{a}^\square) = c \quad \text{iff} \quad \mathrm{BT}_J(\hat{a}^\square) \in F_c\,.$$

Hence, $\mathcal{D}_c$ can guess some set $H \in F_c$ and check that

- for every branch β of $\Lambda(a^\square)$, there is some $Q \in H$ such that $\mathcal{C}_Q$ accepts $\langle t, \hat{a}^\square, \beta \rangle$,
- for every $Q \in H$, there is some branch β of $\Lambda(a^\square)$ such that $\mathcal{C}_Q$ accepts $\langle t, \hat{a}^\square, \beta \rangle$.

(4) After these preparations we can construct an automaton $\mathcal{E}_c$, for $c \in C$, accepting all trees t such that $\varphi(t) = c$. The automaton $\mathcal{E}_c$ has to check that

- for every sequence $\hat{a}^\square \leq a^\square$ in J, there is some $c_0 \leq c$ such that $\mathcal{D}_{c_0}$ accepts $\langle t, \hat{a}^\square \rangle$,
- for every c_0 with $c \nleq c_0$, there is some $c_1 \nleq c_0$ and a sequence $\hat{a}^\square \leq a^\square$ in J such that $\mathcal{D}_{c_1}$ accepts $\langle t, \hat{a}^\square \rangle$.

(5) Finally, the desired automaton accepting $\varphi^{-1}[P]$ guesses some $c \in P$ and checks that $t \in L(\mathcal{E}_c)$. □

Combining Theorems 4.3 and 5.8 and Proposition 5.7, we obtain one of the main theorems of this article.

Theorem 5.9. *Let $L \subseteq T^{\circ}_{\omega}[\Sigma]$. The following statements are equivalent:*

(1) *L is regular.*

(2) *L is recognised by a morphism $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ into a path-hyperclone $\mathfrak{C}$ associated with a finite ω-semigroup.*

(3) *L is recognised by a morphism $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ into a finitary path-continuous ω-hyperclone $\mathfrak{C}$.*

6 Closure properties

In this section we study closure properties of the class of path-continuous ω-hyperclones and of the class of languages recognised by them. We start with products.

Definition 6.1. Let $\mathfrak{C}^{(i)}$, $i \in I$, be a family of ω-hyperclones. The *product* $\prod_{i\in I} \mathfrak{C}^{(i)}$ is the ω-hyperclone $\mathfrak{D}$ where the domain of sort $\bar{u}$ is

$$D_{\bar{u}} := \prod_{i\in I} C^{(i)}_{\bar{u}}\,,$$

and where all operations are defined component-wise.

Lemma 6.2. *The product of ω-hyperclones is an ω-hyperclone.*

Proof. Except for axiom (s4), all of the ω-hyperclone axioms are either equations or implications between inequalities. Such axioms are preserved by products.

Axiom (s4) states that operations $\varrho_\sigma : C_{\bar{u}} \to C_{\sigma(\bar{u})}$ are bijective. This condition is also preserved by products since we can turn it into an equation by adding the inverse function ϱ_σ^{-1} to the structure. □

Lemma 6.3. *The product $\prod_{i \in I} \mathfrak{C}^{(i)}$ of path-continuous ω-hyperclones $\mathfrak{C}^{(i)}$ is path-continuous.*

Proof. Suppose that each $\mathfrak{C}^{(i)}$ is path-continuous and let $\mathfrak{J}_i$ be the corresponding ideal of $\mathfrak{C}^{(i)}$. To see that $\mathfrak{D} := \prod_i \mathfrak{C}^{(i)}$ is path-continuous, we check that the ideal $\mathfrak{H} := \prod_{i \in I} \mathfrak{J}_i$ satisfies the two conditions of Definition 5.6.

First, consider a sequence $a^\square$ in D in separation normal form where $a^n = (a_i^n)_{i \in I} \in D_{\bar{u}_n}$. Then

$$\begin{aligned}\pi(a^\square) = \big(\pi(a_i^\square)\big)_{i \in I} &= \big(\sup\{\,\pi(b_i^\square) \mid b_i^\square \leq a_i^\square \text{ in } J_i\,\}\big)_{i \in I} \\ &= \sup\{\,\pi(b^\square) \mid b^\square \leq a^\square \text{ in } H\,\}.\end{aligned}$$

The second condition we have to check is that, for a sequence $a^\square$ in H, the product $\pi(a^\square)$ is uniquely determined by the set $\mathrm{BT}_J(a^\square)$.

Note that, for $i \in I$, the components $\pi(a_i^\square)$ are determined by $\mathrm{BT}_J(a_i^\square)$. Hence, it is sufficient to show that the set $\mathrm{BT}_J(a^\square)$ determines all sets $\mathrm{BT}_J(a_i^\square)$, $i \in I$. The claim follows since $\mathrm{BT}_J(a_i^\square) = p_i(\mathrm{BT}_J(a^\square))$ where p_i is the projection to the i-th component. □

The closure of the class of path-continuous ω-hyperclones under products implies that the languages recognised by them are closed under boolean operations.

Theorem 6.4. *The class of languages recognised by finitary path-continuous ω-hyperclones is closed under boolean operations.*

Proof. If L and L' are recognised by, respectively, $\mathfrak{C}$ and $\mathfrak{C}'$, then the complement of L is also recognised by $\mathfrak{C}$, while $L \cap L'$ and $L \cup L'$ are recognised by $\mathfrak{C} \times \mathfrak{C}'$. □

Next we turn to closure under projections.

Definition 6.5. Let $\mathfrak{C}$ be an ω-hyperclone. We define an ω-hyperclone $\mathcal{P}(\mathfrak{C})$ as follows. The domain of sort $\bar{u} = \langle u_0, \dots, u_{n-1}\rangle$ is

$$\wp(C_{u_0}) \times \dots \times \wp(C_{u_{n-1}}).$$

(For $n = 0$, we take the empty product $\{\langle\rangle\}$.) To simplify notation we will identify an element $a = a_0 \oplus \dots \oplus a_{m-1} \in C_{\bar{u}}$ of width m with the m-tuple $\bar{a} = \langle a_0, \dots, a_{m-1}\rangle$, and we write $\bar{a} \in \bar{A}$, for $\bar{a} \in C_{\bar{u}}$ and $\bar{A} \in D_{\bar{u}}$, if we have $a_i \in A_i$, for all i. For $\bar{A}$ of width m and $\bar{B}$ of width n, we define the operations of the free monoid structure by

$$\bar{A} \oplus \bar{B} := \bar{A}\bar{B}, \quad 0 := \langle\rangle, \quad \text{and} \quad \lambda_\sigma(\bar{A}) := \bar{A}^\sigma.$$

For elements $\bar{A}$ is separation normal form, we define the substitution operation by

$$\bar{A} :_{I,\sigma} \bar{B} := \bar{D}, \qquad \text{where } D_i := \{\, a :_{I,\sigma} \bar{b} \mid a \in A_i,\ \bar{b} \in \bar{B}\,\}.$$

If $\bar{A}$ is not in separation normal form, say, $\bar{A} = \rho_\tau(\mathrm{sep}(\bar{A}))$, we set

$$\bar{A} :_{I,\sigma} \bar{B} := \mathrm{sep}(\bar{A}) :_{I,\sigma} \bar{B}^\tau.$$

Similarly, we first define the infinite product for sequences $\bar{A}^\square$ in separation normal by

$$\pi(\bar{A}^\square) := \bar{B}, \quad \text{where} \quad B_i := \{\, \pi(a^0, \bar{a}^1, \bar{a}^2 \dots) \mid a^0 \in A_i^0 \text{ and } \bar{a}^n \in \bar{A}^n,\ \text{for } n > 0\,\}.$$

For arbitrary sequences $\bar{A}^\square$, we then set $\pi(\bar{A}^\square) := \pi(\bar{B}^\square)$, where $\bar{B}^\square$ is the separation normal form of $\bar{A}^\square$. Finally, the ordering is defined by

$$\bar{A} \leq \bar{B} \quad :\text{iff} \quad \text{there exist injections } \varphi_i : A_i \to B_i \text{ such that } a \leq \varphi_i(a), \text{ for all } a \in A_i.$$

Lemma 6.6. *$\mathcal{P}(\mathfrak{C})$ is an ω-hyperclone.*

The proof is straightforward but tedious: axioms (M1)–(M5) and (S2), (S3) follow immediately from the definitions; each of the remaining axioms requires some small amount of calculations.

Proposition 6.7. *If $\mathfrak{C}$ is path-continuous then so is $\mathcal{P}(\mathfrak{C})$.*

Proof. Let $\mathfrak{J}$ be the ideal witnessing that $\mathfrak{C}$ is path-continuous. We claim that $H := \{\varnothing\} \cup \{\{a\} \mid a \in J\}$ is a witness of the path-continuity of $\mathfrak{D} := \mathcal{P}(\mathfrak{C})$. We have to check two conditions.

First, consider a sequence $\bar{A}^\square$ in D in separation normal form. We claim that

$$\pi(\bar{A}^\square) = \sup\{\,\pi(\bar{B}^\square) \mid \bar{B}^\square \leq \bar{A}^\square \text{ in } H\,\}.$$

W.l.o.g. assume that $\bar{A}^0 = A^0$ has width 1. Continuing our abuse of notation we write $\{\bar{a}\}$ for the tuple $\langle\{a_0\}, \dots, \{a_{n-1}\}\rangle$. We have

$$\begin{aligned}\pi(\bar{A}^\square) &= \bigcup\big\{\,\{\pi(\bar{a}^\square)\} \bigm| \bar{a}^n \in \bar{A}^n\,\big\} \\ &= \bigcup\big\{\,\{\sup\{\,\pi(\bar{b}^\square) \mid \bar{b}^\square \leq \bar{a}^\square \text{ in } J\,\}\} \bigm| \bar{a}^n \in \bar{A}^n\,\big\} \\ &= \bigcup\big\{\,\sup\{\,\pi(\{\bar{b}^0\}, \{\bar{b}^1\}, \dots) \bigm| \bar{b}^\square \leq \bar{a}^\square \text{ in } J\,\} \bigm| \bar{a}^n \in \bar{A}^n\,\big\} \\ &= \bigcup\big\{\,\sup\{\,\pi(\bar{B}^\square) \bigm| \bar{B}^\square \leq (\{\bar{a}^n\})_n \text{ in } H\,\} \bigm| \bar{a}^n \in \bar{A}^n\,\big\} \\ &= \sup\big\{\,\pi(\bar{B}^\square) \bigm| \bar{B}^\square \leq \bar{A}^\square \text{ in } H\,\big\}.\end{aligned}$$

The second condition we have to check is that, for a sequence $\bar{A}^\square$ in H, the product $\pi(\bar{A}^\square)$ is uniquely determined by the set $\mathrm{BT}_H(\bar{A}^\square)$. W.l.o.g. we may assume that $\bar{A}^\square$ is in separation normal form and that $\bar{A}^\circ$ has width 1. If $A^n_i = \varnothing$, for some n, i, then $\pi(\bar{A}^\square) = \varnothing$ and $\mathrm{BT}_H(\bar{A}^\square) = \{\{\varnothing\}\}$. Hence, we may assume that $A^n_i \neq \varnothing$, for all n, i. Since $\bar{A}^\square$ is in H, it follows that $A^n_i = \{a^n_i\}$, for suitable $a^n_i \in C$. Let β be a branch and $(\bar{B}^n)_n \in \mathrm{LC}_H(\bar{A}^\square)$. If $B^0_0 = \varnothing$, the trace of $(\bar{B}^n)_n$ along β evaluates to $\varnothing$. Otherwise, all B^n_i are nonempty and there are $b^n_i \in C$ such that $B^n_i = \{b^n_i\}$. If $\bar{C}^\square$ is the trace of $(\bar{B}^n)_n$ along β then $\bar{C}^n \in H$, for all n. Hence, $C^n_i = \{c^n_i\}$ and $\bar{c}^\square$ is the trace of $(\bar{b}^n)_n$ along β. It follows that

$$\mathrm{Tr}_H(\beta) = \{\varnothing\} \cup \{\, \{c\} \mid c \in \mathrm{Tr}_J(\beta) \,\}\,.$$

Since this set uniquely determines $\mathrm{Tr}_J(\beta)$, the product $\pi(\bar{A}^\square) = \{\pi(\bar{a}^\square)\}$ is uniquely determined by

$$\big\{\, \mathrm{Tr}_H(\beta) \bigm| \beta \text{ a branch} \,\big\} = \mathrm{BT}_H(\bar{A}^\square)\,.$$

This implies the claim. ☐

It follows that the class of recognisable languages is closed under projections.

Theorem 6.8. *Let $p : \Sigma \to \Gamma$ be a surjective, arity preserving function between functional signatures. If $L \subseteq F_{\langle\varnothing\rangle}[\Sigma]$ is recognised by a path-continuous ω-hyperclone $\mathfrak{C}$ then $p[L] \subseteq F_{\langle\varnothing\rangle}[\Gamma]$ is recognised by $\mathcal{P}(\mathfrak{C})$.*

Proof. Let $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ recognise L and set $P := \varphi[L]$. We claim that the function $\psi : \mathfrak{F}_\omega[\Gamma] \to \mathcal{P}(\mathfrak{C})$ with

$$\psi(\bar{t}) := (\varphi \circ p^{-1})(\bar{t})$$

recognises $p[L]$. More precisely, setting $Q := \{\, A \mid A \cap P \neq \varnothing \,\}$ we claim that $p[L] = \psi^{-1}[Q]$.

($\subseteq$) Let $t \in p[L]$. Then there is some $s \in L$ such that $t = p(s)$. Since $\varphi(s) \in P$ and $s \in p^{-1}(t)$ it follows that $\varphi(s) \in P \cap \varphi[p^{-1}(t)]$. Hence, $\psi(t) \cap P \neq \varnothing$ and $\psi(t) \in Q$.

($\supseteq$) Let $\psi(t) \in Q$. Then $\{\, \varphi(s) \mid s \in p^{-1}(t) \,\} \cap P = \psi(t) \cap P \neq \varnothing$. Hence, there is some $s \in p^{-1}(t)$ with $\varphi(s) \in P$. Consequently, $s \in L$ and $p(s) = t$. It follows that $t \in p[L]$. ☐

Together the closure properties established in this section imply that recognisability by finitary path-continuous ω-hyperclones is equivalent to axiomatisability in monadic second-order logic and, hence, to regularity. In particular, this yields an alternative proof of Theorem 5.9, a proof that does not make use of automata.

For the following theorem, we use the variant of monadic second-order logic (MSO) without first-order variables. The atomic formulae of this variant are of the form $X \subseteq Y$ and $R\bar{Z}$, for set variables $X, Y, Z_0, Z_1, \dots$ and relation symbols R. An atom of the form $R\bar{Z}$ is true if, and only if, there is some tuple $\bar{a} \in R$ with $a_i \in Z_i$, for all i. We regard a Σ-term t as a structure $\lceil t \rceil := \langle T, \preceq, (P_f)_{f\in\Sigma} \rangle$ where the universe $T := \mathrm{dom}(t)$ consists of all vertices of t, $\preceq$ is the tree ordering and, for every function symbol $f \in \Sigma$, we have a unary relation P_f containing all vertices of T labelled by f. Since in the inductive step below we will be dealing with formulae $\chi(\bar{X})$ with free set variables $\bar{X}$, we also have to consider expansions $\langle \lceil t \rceil, \bar{Q} \rangle := \langle T, \preceq, (P_f)_{f\in\Sigma}, \bar{Q} \rangle$ of such trees by additional unary predicates $\bar{Q}$ providing values for the variables $\bar{X}$. If $\bar{X} = \langle X_0, \dots, X_{n-1} \rangle$, we can encode such an expansion as a term $t_{\bar{Q}}$ over the signature

$$\Sigma_n := \big\{\, f_{\bar{b}} \bigm| f \in \Sigma,\ \bar{b} \in \{0,1\}^n \,\big\}$$

where, for $v \in \mathrm{dom}(t)$, we set

$$t_{\bar{Q}}(v) := f_{\bar{b}} \quad \text{with} \quad f := t(v) \quad \text{and} \quad b_i := \begin{cases} 1 & \text{if } v \in Q_i\,, \\ 0 & \text{if } v \notin Q_i\,. \end{cases}$$

In this way, we can associate with every MSO-formula $\chi(\bar{X})$ the language

$$L_\chi := \big\{\, t_{\bar{Q}} \in F_{\langle\varnothing\rangle}[\Sigma_n] \bigm| \lceil t \rceil \vDash \chi(\bar{Q}) \,\big\}\,.$$

Languages of this form are called MSO-*axiomatisable.*

Theorem 6.9. *Let Σ be a finite signature. A set $L \subseteq F_{\langle\varnothing\rangle}[\Sigma]$ of trees is recognised by a finitary path-continuous ω-hyperclone if, and only if, it is* MSO-*axiomatisable.*

Proof. ($\Rightarrow$) Let $\varphi : \mathfrak{F}_\omega[\Sigma] \to \mathfrak{C}$ be a homomorphism into a finitary path-continuous ω-hyperclone and let $\mathfrak{J}$ be the ideal witnessing that $\mathfrak{C}$ is path-continuous. For every $c \in C_{\langle\varnothing\rangle}$, we will construct a formula χ_c such that

$$\lceil t \rceil \vDash \chi_c \quad \text{iff} \quad \varphi(t) = c\,, \qquad \text{for all } t\,.$$

Given t, let $v^n_0, \dots, v^n_{m(n)-1}$ be an enumeration of all vertices $v \in \mathrm{dom}(t)$ of length $|v| = n$, and set

$$t^n_i := t(v^n_i)\,, \qquad t^n := \mathrm{sep}(t^n_0 \oplus \dots \oplus t^n_{m(n)-1})\,,$$
$$a^n_i := \varphi(t^n_i)\,, \qquad a^n := \mathrm{sep}(a^n_0 \oplus \dots \oplus a^n_{m(n)-1}) = \varphi(t^n)\,.$$

Then $t = \pi(t^\square)$ and $\varphi(t) = \pi(a^\square)$.

The formula χ_c has to check that $\pi(a^\square) = c$. To do so it has to consider a sequence $a_0^\square \leq a^\square$ in J and a sequence $(b^n)_n \in \mathrm{LC}_J(a_0^\square)$. We can encode such sequences a tuples $\bar{A}$ and $\bar{B}$ of set variables as follows. Let m be the maximal arity of a function in Σ and set

$$F_0 := \{ a \in C \mid a \leq \varphi(f) \text{ for some } f \in \Sigma \},$$
$$F_1 := C_{\langle\varnothing\rangle} \cup C_{\langle\varnothing,\varnothing\rangle} \cup \dots \cup C_{\varnothing^m}.$$

Note that the sets F_0 and F_1 are finite. To encode a sequence $a_0^\square$ where $a_0^n = \mathrm{sep}((a_0^n)_0 \oplus \dots \oplus (a_0^n)_{m(n)-1})$ we use a tuple $(A_x)_{x \in F_0}$ of set variables where

$$A_x := \{ u \in \mathrm{dom}(t) \mid (a_0^n)_i = x \text{ for } i \text{ and } n \text{ with } u = v_i^n \},$$

and we encode $(b^n)_{n<\omega}$ where $b^n = \mathrm{sep}(b_0^n \oplus \dots \oplus b_{m(n)-1}^n)$ by the set variables $(B_x)_{x \in F_1}$ where

$$B_x := \{ u \in \mathrm{dom}(t) \mid b_i^n = x \text{ for } i \text{ and } n \text{ with } u = v_i^n \}.$$

Using this encoding, we can write down MSO-formulae expressing the following facts:

- $\vartheta_\leq(\bar{A})$: "$a_0^\square \leq a^\square$ is in J."
- $\vartheta_{\mathrm{lc}}(\bar{A}, \bar{B})$: "$(b^n)_n \in \mathrm{LC}_J(a_0^\square)$."
- $\vartheta_{\mathrm{branch}}(Z)$: "$Z$ is a branch in $\mathrm{dom}(t)$."
- $\vartheta_{\mathrm{tr}}^c(\bar{A}, \bar{B}, Z)$: "The product of the trace of $\bar{B}$ along Z equals c."

For every set $Q \subseteq C_{\langle\varnothing\rangle}$, we set

$$\vartheta_{\mathrm{TR}}^Q(\bar{A}, Z) := \bigwedge_{c \in Q} \exists \bar{B}\big[\vartheta_{\mathrm{lc}}(\bar{A}, \bar{B}) \wedge \vartheta_{\mathrm{tr}}^c(\bar{A}, \bar{B}, Z)\big] \wedge \forall \bar{B}\big[\vartheta_{\mathrm{lc}}(\bar{A}, \bar{B}) \rightarrow \bigvee_{c \in Q} \vartheta_{\mathrm{tr}}^c(\bar{A}, \bar{B}, Z)\big].$$

Then we have

$$\lceil t \rceil \vDash \vartheta_{\mathrm{TR}}^Q(\bar{A}, Z) \quad \text{iff} \quad \mathrm{Tr}_J(\beta) = Q \quad \text{where } \beta \text{ is the branch encoded in } Z.$$

For every $c \in C_{\langle\varnothing\rangle}$, there exists a set $S_c \subseteq \wp(\wp(C_{\langle\varnothing\rangle}))$, such that

$$\pi(a_0^\square) = c \quad \text{iff} \quad \mathrm{BT}_J(a_0^\square) \in S_c.$$

It follows that the formula

$$\vartheta_\pi^c(\bar{A}) := \bigvee_{U \in S_c} \Big[\bigwedge_{Q \in U} \exists Z\big[\vartheta_{\mathrm{branch}}(Z) \wedge \vartheta_{\mathrm{TR}}^Q(\bar{A}, Z)\big] \wedge \forall Z\big[\vartheta_{\mathrm{branch}}(Z) \rightarrow \bigvee_{Q \in U} \vartheta_{\mathrm{TR}}^Q(\bar{A}, Z)\big]\Big]$$

satisfies

$$\lceil t \rceil \vDash \vartheta_\pi^c(\bar{A}) \quad \text{iff} \quad \pi(a_0^\square) = c,$$

for all sequences $a_0^\square$ in J. Setting

$$\vartheta_\pi^{\leq c}(\bar{A}) := \bigvee_{d \leq c} \vartheta_\pi^d(\bar{A})$$

we obtain the desired formula

$$\chi_c := \forall \bar{A}\big[\vartheta_\leq(\bar{A}) \rightarrow \vartheta_\pi^{\leq c}(\bar{A})\big] \wedge \bigwedge_{d \nleq c} \exists \bar{A}\big[\vartheta_\leq(\bar{A}) \wedge \neg\vartheta_\pi^{\leq d}(\bar{a})\big].$$

($\Leftarrow$) For simplicity, we will call a language L *recognisable* if it is recognisable by a finitary path-continuous ω-hyperclone. We will show that L_χ is recognisable by induction on the MSO-formula χ. By Theorems 6.4 and 6.8, the class of recognisable languages is closed under boolean operations and projections. It is therefore sufficient to show that every language L_χ axiomatised by an atomic MSO-formula χ is recognisable.

First, we consider a formula of the form $\chi = X_i \subseteq X_k$. Note that $L_\chi = F_{\langle\varnothing\rangle}[\Gamma]$ where

$$\Gamma := \{ f_{\bar{b}} \in \Sigma_n \mid b_i \leq b_k \}.$$

We have seen in Lemma 3.5, that the set $F_{\langle\varnothing\rangle}[\Gamma]$ is recognisable. Furthermore, the ω-hyperclone used in the proof of that lemma is finitary and path-continuous.

Next, suppose that $\chi = P_f X_i$. In this case we have $L_\chi = F_{\langle\varnothing\rangle}[\Sigma_n] \setminus F_{\langle\varnothing\rangle}[\Gamma]$ for

$$\Gamma := \{ g_{\bar{b}} \in \Sigma_n \mid g \neq f \text{ or } b_i = 0 \}.$$

As above $F_{\langle\varnothing\rangle}[\Gamma]$ is recognisable. The claim follows since the recognisable languages are closed under complement.

It remains to consider formulae of the form $\chi = X_i \leq X_k$. We use the path-hyperclone $\mathfrak{C}$ associated with the ω-semigroup $\langle S, S_\omega\rangle$ where $S := \{0,1\}^2 \cup \{0\}$ and $S_\omega := \{\bot, *, \top\}$. Intuitively, a pair $\langle b, c\rangle \in S$ records whether the corresponding term contains a vertex in X_i (in this case $b = 1$) or a vertex in X_k (in this case $c = 1$). The element $0 \in S$ indicates a term containing vertices $u \leq v$ with $u \in X_i$ and $v \in X_k$. Similarly, $\bot \in S_\omega$ represents a term without any vertex in X_k, $* \in S_\omega$ represents a term with some vertex in X_k, and $\top \in S_\omega$ represents a term with vertices $u \leq v$ with $u \in X_i$ and $v \in X_k$. The multiplication of $\langle S, S_\omega\rangle$ is defined as follows

$$\begin{aligned} 0 \cdot x &= 0 = x \cdot 0\,, \qquad \text{for all } x \in S\,,\\ \langle b, b'\rangle \cdot \langle c, c'\rangle &= \begin{cases} 0 & \text{if } b = 1 \text{ and } c' = 1\,,\\ \langle \max\{b,c\}, \max\{b',c'\}\rangle & \text{otherwise}\,,\end{cases}\\ x \cdot \top &= \top \qquad \text{for } x \in S\,,\\ 0 \cdot u &= \top \qquad \text{for } u \in S_\omega\,,\\ \langle 0, 0\rangle \cdot u &= u \qquad \text{for } u \in S_\omega\,,\\ \langle b, 1\rangle \cdot u &= * \qquad \text{for } u \in \{\bot, *\}\,,\\ \langle 1, b\rangle \cdot * &= \top\,,\\ \langle 1, 0\rangle \cdot \bot &= \bot\,. \end{aligned}$$

The infinite product $\pi(s_0, s_1, s_2, \dots)$ is defined as follows. If there is some $m < \omega$ with $s_m = 0$, or there are $l < m < \omega$ with $s_l = \langle 1, b\rangle$ and $s_m = \langle c, 1\rangle$, then

$$\pi(s_0, s_1, s_2, \dots) = \top\,.$$

Otherwise, if there is some $m < \omega$ with $s_m = \langle b, 1\rangle$ then

$$\pi(s_0, s_1, s_2, \dots) = *\,.$$

Finally, if $s_m \in \{\langle 0,0\rangle, \langle 1,0\rangle\}$ for all m, then we set

$$\pi(s_0, s_1, s_2, \dots) = \bot\,.$$

We define the homomorphism $\varphi : \mathfrak{F}_\omega[\Sigma_n] \to \mathfrak{C}$ as follows. For a function

symbol $f \in \Sigma$ of arity m and $\bar{b} \in \{0,1\}^n$, we set

$$\varphi(f_{\bar{b}}) := \begin{cases} \{\langle 0^m, \varnothing\rangle\} & \text{if } b_i = b_k = 1\,,\\ \{\langle\langle b_i, b_k\rangle^m, \varnothing\rangle\} & \text{otherwise}\,.\end{cases}$$

It follows that

$$L_\chi = \varphi^{-1}[P] \quad \text{where} \quad P := \big\{\, \{\langle\langle\rangle, X\rangle\} \in C_{\langle\varnothing\rangle} \;\big|\; \top \in X \,\big\}\,. \qquad \square$$

7 Conclusions

In this article we have developed the beginnings of a theory of recognisability for infinite trees, but there remains much to do. First of all, we do not believe that the framework we have set up is in its final form. The algebras we use (path-continuous ω-hyperclones) are far too complicated. In particular,

- we use infinitely many sorts,
- there are too many operations (in particular $:_{I,\sigma}$), and
- the definition of path-continuity is too complex.

Apart from simplifying the algebraic framework the logically next step consists in finding the right notion of a Wilke algebra. Our hope is that such algebras can be used in conjunction with Theorem 6.9 to give an alternative proof of Rabin's Tree Theorem. Finding such a proof has been an open problem for 30 years. The main missing ingredient seems to be the lack of a suitable Ramseyan factorisation theorem for infinite trees. It is even unclear what exactly such a theorem should state. Since the search for a Wilke algebra for ω-hyperclones seems to require exactly such a factorisation theorem, we hope that we can make some headway by approaching the problem from this direction.

Finally, a longterm goal would be the development of a theory of pseudo-varieties of path-continuous ω-hyperclones and a corresponding structure theory. But, before embarking upon such a project, it seems advisable to wait until we have found the 'right' definition for our algebras.

References

[1] P. Baltazar, *M-Solid Varieties of Languages*, Acta Cynbernetica, 18 (2008), pp. 719–731.

[2] 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.

[3] M. Bojańczyk and I. Walukiewicz, *Forest Algebras*, in Logic and Automata: History and Perspectives, J. Flum, E. Grädel, and T. Wilke, eds., Amsterdam University Press, 2007, pp. 107–132.

[4] Z. Ésik, *A variety theorem for trees and theories*, Publ. Math. Debrecen, 54 (1999), pp. 711–762.

[5] ———, *Characterizing* CTL*-like logics on finite trees*, Theoretical Computer Science, 356 (2006), pp. 136–152.

[6] Z. Ésik and P. Weil, *Algebraic recognizability of regular tree languages*, Theoretical Computer Science, 340 (2005), pp. 291–321.

[7] ———, *Algebraic Characterization of Logically Defined Tree Languages*, International Journal of Algebra and Computation, 20 (2010), pp. 195–239.

[8] D. Perrin and J.-E. Pin, *Infinite Words – Automata, Semigroups, Logic and Games*, Elsevier, 2004.

[9] T. Petkovic and S. Salehi, *Positive varieties of tree languages*, Theoretical Computer Science, 347 (2005), pp. 1–35.

[10] S. Salehi and M. Steinby, *Tree algebras and varieties of tree languages*, Theoretical Computer Science, 377 (2007), pp. 1–24.

[11] M. Steinby, *A theory of tree language varieties*, in Tree Automata and Languages, North-Holland, 1992, pp. 57–82.

[12] ———, *General Varieties of Tree Languages*, Theoretical Computer Science, 205 (1998), pp. 1–43.

[13] T. Wilke, *Algebras for Classifying Regular Tree Languages and an Application to Frontier Testability*, in Proc. 20th Int. Colloquium on Automata, Languages and Programming, ICALP 1993, LNCS 700, 1993, pp. 347–358.