



Saeed Salehi

Varieties of Tree Languages Definable by Syntactic Monoids

TURKU CENTRE *for* COMPUTER SCIENCE

TUCS Technical Report
No 619, August 2004



Varieties of Tree Languages Definable by Syntactic Monoids

Saeed Salehi

Abstract

An algebraic characterization of the families of tree languages definable by syntactic monoids is presented. This settles a question raised by several authors.

Keywords: Tree languages, Tree automata, Varieties of tree languages, Varieties of finite algebras, Syntactic monoids

TUCS Laboratory

Discrete Mathematics for Information Technology

1 Introduction

A Variety Theorem establishing a bijective correspondence between general varieties of tree languages definable by syntactic monoids and varieties of finite monoids, is proved. This has been a relatively long-standing open problem, the most recent references to which are made by Ésik [4] as “No variety theorem is known in the semigroup [monoid] approach” (page 759), and by Steinby [18] as “there are no general criteria for deciding whether or not a given GVTL [general variety of tree languages] can or cannot be defined by syntactic monoids” (page 41). The question was also mentioned in the last section of Wilke’s paper [21].

Most of the interesting classes of algebraic structures form varieties, and similarly, most of the interesting families of tree or string languages studied in the literature turn out to be varieties of some kind. The first Variety Theorem was proved by Eilenberg [3] who established a correspondence between varieties of finite monoids and varieties of regular (string) languages. It was motivated by characterizations of several families of languages by syntactic monoids or semigroups (see [3],[10]), above all by Schützenberger’s [15] theorem connecting star-free languages and aperiodic monoids.

Eilenberg’s theorem has since been extended in various directions. One could mention Pin’s [11] Variety Theorem for positive varieties of string languages and varieties of ordered monoids, or Thérien’s [19] extension that includes also varieties of congruences on free monoids. On the level of universal algebra, where tree automata and tree languages are studied, a Variety Theorem was proved by Steinby [16] for recognizable subsets of finitely generated free algebras. Both Eilenberg’s $*$ -varieties and $+$ -varieties, as well as varieties of regular tree languages (which was worked out in [17]), are special cases of the results of [16]. The correspondence to varieties of congruences, and some other generalizations, were added later by Almeida [1] and Steinby [17, 18]. Another example is Ésik’s [4] Variety Theorem between tree languages and theories (see also [5]). As Ésik observes in [4], page 758: “The crucial concept in any ‘Variety Theorem’ is that of the ‘syntactic structure’ or ‘syntactic algebra’.” For almost all those syntactic structures associated to tree languages in the literature, one (or some) variety theorem(s) have been proved. The most famous ‘syntactic structure’ for which a variety theorem was not known, is the syntactic semigroup/monoid of a tree language, introduced by Thomas [20], and further studied by Salomaa [14]. A different formalism, based on the essentially same concept, was brought up by Nivat and Podelski [6], [13].

To establish our correspondence between varieties of tree languages and varieties of finite monoids, we add three more closure properties to the definition of a general tree language variety introduced in [18]. One of them, that of being closed under inverse tree homomorphisms, is already investigated by Ésik [4], and the other two are stated in Theorem 4.14.

2 Notation and Preliminaries

Our notation is mainly based on [18]. However for understanding our results it is not necessary to read the whole of [18]. Here, we list the terminology used throughout the paper.

A finite set of function symbols is called a *ranked alphabet*. If Σ is a ranked alphabet, for every $m \geq 0$, the set of m -ary function symbols of Σ is denoted by Σ_m . In particular, Σ_0 is the set of constant symbols of Σ . For a ranked alphabet Σ and a *leaf alphabet* X , the set of ΣX -trees $T(\Sigma, X)$ is the smallest set satisfying

- (1) $\Sigma_0 \cup X \subseteq T(\Sigma, X)$, and
- (2) $f(t_1, \dots, t_m) \in T(\Sigma, X)$, for all $f \in \Sigma_m$ ($m > 0$) and $t_1, \dots, t_m \in T(\Sigma, X)$.

Any subset of $T(\Sigma, X)$ is called a *tree language*.

The ΣX -term algebra $\mathcal{T}(\Sigma, X) = (T(\Sigma, X), \Sigma)$ is defined by setting

- (1) $c^{\mathcal{T}(\Sigma, X)} = c$ for each $c \in \Sigma_0$, and
- (2) $f^{\mathcal{T}(\Sigma, X)}(t_1, \dots, t_m) = f(t_1, \dots, t_m)$ for all $m > 0$, $f \in \Sigma_m$, and $t_1, \dots, t_m \in T(\Sigma, X)$.

Let ξ be a (special) symbol which does not appear in any ranked alphabet or leaf alphabet considered here. The set of ΣX -contexts, denoted by $C(\Sigma, X)$, consists of the $\Sigma(X \cup \{\xi\})$ -trees in which ξ appears exactly once. For $P, Q \in C(\Sigma, X)$ and $t \in T(\Sigma, X)$ the context $Q \cdot P$, the composite of P and Q , results from P by replacing the special leaf ξ with Q , and the term $t \cdot P$ results from P by replacing ξ with t . Note that $C(\Sigma, X)$ is a monoid with composition as the operation and ξ as the unit element, and that $t \cdot (Q \cdot P) = (t \cdot Q) \cdot P$ holds for all $P, Q \in C(\Sigma, X)$, $t \in T(\Sigma, X)$. For a tree language $T \subseteq T(\Sigma, X)$ and context P , the *inverse translation* of T under P is $P^{-1}(T) = \{t \in T(\Sigma, X) \mid t \cdot P \in T\}$. Also the *inverse morphism* of T under a homomorphism $\varphi : T(\Sigma, Y) \rightarrow T(\Sigma, X)$ is $T\varphi^{-1} = \{t \in T(\Sigma, Y) \mid t\varphi \in T\}$.

A ΣX -recognizer (\mathcal{A}, α, F) consists of a finite Σ -algebra $\mathcal{A} = (A, \Sigma)$, an *initial assignment* $\alpha : X \rightarrow A$, and a *set of final states* $F \subseteq A$. The function α can uniquely be extended to a homomorphism $\alpha^{\mathcal{A}} : \mathcal{T}(\Sigma, X) \rightarrow \mathcal{A}$, and the tree language recognized by (\mathcal{A}, α, F) is $\{t \in T(\Sigma, X) \mid t\alpha^{\mathcal{A}} \in F\}$. In that case we also simply say that T is recognized by the algebra \mathcal{A} .

All algebras considered in this paper, except for term algebras, are finite, and the tree languages studied here are recognizable by finite algebras. A class of finite algebras of a fixed type is called a *variety of finite algebras* if it is closed under subalgebras, homomorphic images, and finite products. They are sometimes called *pseudo-varieties*, to be differentiated from real varieties whose members need not to be finite. Birkhoff's variety theorem [2] provides a logical characterization of those "original" varieties. In particular, a variety of finite monoids, abbreviated by VFM, is a class of finite monoids closed under submonoids, homomorphic images, and finite monoid products. A family $\mathcal{V} = \{\mathcal{V}(X)\}$ of tree languages of a fixed type Σ is a mapping which assigns to every finite leaf alphabet a collection $\mathcal{V} = \{\mathcal{V}(X)\}$ of recognizable ΣX -tree languages. A family \mathcal{V} is called a *variety of tree*

languages if each $\mathcal{V}(X)$ is closed under Boolean operations and inverse translations, and the whole collection is closed under the inverse homomorphisms between term algebras (see [17]; below we will consider generalized varieties of tree languages).

Let $\mathcal{A} = (A, \Sigma)$ be an algebra. Every elementary context

$$P = f(a_1, \dots, \xi, \dots, a_m) \in C(\Sigma, A),$$

where $f \in \Sigma_m$ and $a_1, \dots, a_m \in A$, induces a unary function on A defined by $P^{\mathcal{A}}(a) = f^{\mathcal{A}}(a_1, \dots, a, \dots, a_m)$ for each $a \in A$. Such functions are called *elementary translations* of \mathcal{A} . The functions induced by compositions of such elementary contexts are defined by setting $(Q \cdot P)^{\mathcal{A}}(a) = P^{\mathcal{A}}(Q^{\mathcal{A}}(a))$ for any two contexts P and Q and any $a \in A$. These functions constitute the set of *translations* of \mathcal{A} denoted by $\text{Tr}(\mathcal{A})$. Note that two different contexts may induce the same translation.

The set $\text{Tr}(\mathcal{A})$ is a monoid with composition as the operation, called the *translation monoid* of \mathcal{A} , which is also denoted by $\text{Tr}(\mathcal{A})$. We note that $\text{Tr}(\mathcal{A})$ includes the identity translation $\xi^{\mathcal{A}} = 1_A$. The composition of translations p and q is denoted by $q \cdot p$, that is $(q \cdot p)(a) = p(q(a))$ for all $a \in A$ (cf. Section 5 of [18]).

For a tree language $T \subseteq T(\Sigma, X)$, the *syntactic congruence* θ_T of T is defined by

$$t \theta_T s \iff \forall P \in C(\Sigma, X) (t \cdot P \in T \leftrightarrow s \cdot P \in T),$$

for $t, s \in T(\Sigma, X)$, and the *syntactic algebra* $\text{SA}(T)$ of T is the quotient Σ -algebra $T(\Sigma, X)/\theta_T$ (see Definition 5.9 of [18]).

Also, the *m-congruence* μ_T of T on the monoid $C(\Sigma, X)$ is defined by

$$P \mu_T Q \iff \forall R \in C(\Sigma, X) \forall t \in T(\Sigma, X) (t \cdot P \cdot R \in T \leftrightarrow t \cdot Q \cdot R \in T),$$

for $P, Q \in C(\Sigma, X)$, and the *syntactic monoid* $\text{SM}(T)$ of T is the quotient monoid $C(\Sigma, X)/\mu_T$ (cf. [20] or Definition 10.1 of [18]).

Remark 2.1 It was shown in [14] that the translation monoid of the syntactic algebra of a tree language is isomorphic to the syntactic monoid of the tree language, i.e., $\text{Tr}(\text{SA}(T)) \cong \text{SM}(T)$ for every tree language T .

A *tree homomorphism* is a mapping $\varphi : T(\Sigma, X) \rightarrow T(\Omega, Y)$ for ranked alphabets Σ and Ω , and leaf alphabets X and Y , determined by some mappings $\varphi_X : X \rightarrow T(\Omega, Y)$, and $\varphi_m : \Sigma_m \rightarrow T(\Omega, Y \cup \{\xi_1, \dots, \xi_m\})$, where $\Sigma_m \neq \emptyset$ and the ξ_i 's are new variables, inductively as follows

(1) $x\varphi = \varphi_X(x)$ for $x \in X$, $c\varphi = \varphi_0(c)$ for $c \in \Sigma_0$, and

(2) $f(t_1, \dots, t_n)\varphi = \varphi_n(f)[\xi_1 \leftarrow t_1\varphi, \dots, \xi_n \leftarrow t_n\varphi]$ that is ξ_i is replaced with $t_i\varphi$ for all i (cf. [18], page 7).

A tree homomorphism $\varphi : T(\Sigma, X) \rightarrow T(\Omega, Y)$ is called *regular* if for every $f \in \Sigma_m$ ($m \geq 1$), each ξ_1, \dots, ξ_m appears exactly once in $\varphi_m(f)$.

The unique extension $\varphi_* : C(\Sigma, X) \rightarrow C(\Omega, Y)$ of a regular tree homomorphism φ to contexts is obtained by setting $\varphi_*(\xi) = \xi$ (cf. [18], Proposition 10.3).¹ We note that the identities $(Q \cdot P)\varphi_* = Q\varphi_* \cdot P\varphi_*$ and $(t \cdot Q \cdot P)\varphi = t\varphi \cdot Q\varphi_* \cdot P\varphi_*$ hold for all $P, Q \in C(\Sigma, X)$ and $t \in T(\Sigma, X)$.

¹Indeed any tree homomorphism $\varphi : T(\Sigma, X) \rightarrow T(\Omega, Y)$ can be extended to $\tilde{\varphi} : C(\Sigma, X) \rightarrow T(\Omega, Y \cup \{\xi\})$ by setting $\xi\tilde{\varphi} = \xi$, but if φ is not regular the range of $\tilde{\varphi}$ may

3 Algebras Definable by Translation Monoids

The notions of *subalgebra*, *homomorphism*, and *direct product* are defined as usual in Universal Algebra, whereas for their generalizations, *g-subalgebra*, *g-homomorphism*, and *generalized product*, are defined for algebras which are not necessarily of the same type. We recall the following definitions from [18] (Definitions 3.1, 3.2, 3.3, 3.14).

Definition 3.1 Let $\mathcal{A} = (A, \Sigma)$ and $\mathcal{B} = (B, \Omega)$ be finite algebras.

The algebra \mathcal{B} is a *g-subalgebra* of \mathcal{A} , in notation $\mathcal{B} \subseteq_g \mathcal{A}$, if $B \subseteq A$, $\Omega_m \subseteq \Sigma_m$ for all $m \geq 0$, and for every $g \in \Omega_m$, $g^{\mathcal{B}}$ is the restriction of $g^{\mathcal{A}}$ to B .

An *assignment* is a mapping $\kappa : \Sigma \rightarrow \Omega$ such that $\kappa(\Sigma_m) \subseteq \Omega_m$ for all $m \geq 0$. A *g-morphism* from \mathcal{A} to \mathcal{B} is a pair (κ, φ) , where $\kappa : \Sigma \rightarrow \Omega$ is an assignment and $\varphi : A \rightarrow B$ is a mapping satisfying $f^{\mathcal{A}}(a_1, \dots, a_m)\varphi = (f\kappa)^{\mathcal{B}}(a_1\varphi, \dots, a_m\varphi)$ for any $m \geq 0$, $f \in \Sigma_m$, and $a_1, \dots, a_m \in A$. If both κ and φ are surjective, then (κ, φ) is called a *g-epimorphism*, and in that case we write $\mathcal{B} \leftarrow_g \mathcal{A}$ (\mathcal{B} is a g-epimorphic image of \mathcal{A}). When \mathcal{B} is a g-epimorphic image of a g-subalgebra of \mathcal{A} , we write $\mathcal{B} \prec_g \mathcal{A}$. When both κ and φ are bijective, (κ, φ) is called a *g-isomorphism*, and $\mathcal{B} \cong_g \mathcal{A}$ means that \mathcal{B} and \mathcal{A} are g-isomorphic.

Let $\Sigma^1, \dots, \Sigma^n$ and Γ be ranked alphabets. The product $\Sigma^1 \times \dots \times \Sigma^n$ is a ranked alphabet such that $(\Sigma^1 \times \dots \times \Sigma^n)_m = \Sigma_m^1 \times \dots \times \Sigma_m^n$ for every $m \geq 0$. For any assignment $\kappa : \Gamma \rightarrow \Sigma^1 \times \dots \times \Sigma^n$, and any algebras $\mathcal{A}_1 = (A_1, \Sigma^1), \dots, \mathcal{A}_n = (A_n, \Sigma^n)$, the κ -*product* of $\mathcal{A}_1, \dots, \mathcal{A}_n$ is the Γ -algebra $\kappa(\mathcal{A}_1, \dots, \mathcal{A}_n) = (A_1 \times \dots \times A_n, \Gamma)$ defined by

- (1) $c^{\kappa(\mathcal{A}_1, \dots, \mathcal{A}_n)} = (c_1^{\mathcal{A}_1}, \dots, c_n^{\mathcal{A}_n})$ for $c \in \Gamma_0$, where $c\kappa = (c_1, \dots, c_n)$, and
- (2) $f^{\kappa(\mathcal{A}_1, \dots, \mathcal{A}_n)}(\mathbf{a}_1, \dots, \mathbf{a}_m) = (f_1^{\mathcal{A}_1}(a_{11}, \dots, a_{m1}), \dots, f_n^{\mathcal{A}_n}(a_{1n}, \dots, a_{mn}))$ for $f \in \Gamma_m$ ($m > 0$) and $\mathbf{a}_i = (a_{i1}, \dots, a_{in}) \in A_1 \times \dots \times A_n$, where $f\kappa = (f_1, \dots, f_n)$.

Without specifying the assignment κ , such algebras are called *g-products*. In the notations $\subseteq_g, \leftarrow_g, \prec_g$, and \cong_g , the subscript g is dropped when \mathcal{A} and \mathcal{B} are of the same type, say Σ , and the assignment $\kappa : \Sigma \rightarrow \Sigma$ is the identity mapping.

The abbreviation GVFA stands for *general variety of finite algebras* which is a class of finite algebras, of all finite types, closed under g-subalgebras, g-epimorphic images, and g-products (Definition 4.3 of [18]). It is easy to see that a class of algebras \mathbf{K} is a GVFA, if for any $\mathcal{A}_1, \dots, \mathcal{A}_n \in \mathbf{K}$, any g-product $\kappa(\mathcal{A}_1, \dots, \mathcal{A}_n)$, and any algebra \mathcal{A} , if $\mathcal{A} \prec_g \kappa(\mathcal{A}_1, \dots, \mathcal{A}_n)$ then $\mathcal{A} \in \mathbf{K}$ (cf. Corollary 4.8 of [18]).

Definition 3.2 For a VFM \mathbf{M} , \mathbf{M}^a is the class of all finite algebras whose translation monoids are in \mathbf{M} , i.e., $\mathcal{A} \in \mathbf{M}^a \Leftrightarrow \text{Tr}(\mathcal{A}) \in \mathbf{M}$ for any finite algebra \mathcal{A} .

not be $C(\Omega, Y)$. Hence the regularity of φ is needed for the existence of the extension φ_* , see also Example 4.8.

A class of finite algebras \mathbf{K} is said to be *definable by translation monoids*, if there is a VFM \mathbf{M} such that $\mathbf{M}^{\mathbf{a}} = \mathbf{K}$.

By Proposition 10.8 of [18], a class of finite algebras definable by translation monoids is a GVFA. In fact, any such class can be proved to be a *d-variety of finite algebras* (see page 758 of [4]). An algebraic characterization of the classes of finite algebras definable by translation monoids is given in the main theorem of this section.

Definition 3.3 Let \mathcal{A} be a finite algebra. With each translation $p \in \text{Tr}(\mathcal{A})$ we associate a unary function symbol \bar{p} . Let $\Lambda_{\mathcal{A}} = \{\bar{p} \mid p \in \text{Tr}(\mathcal{A})\}$ be the unary ranked alphabet formed by these symbols and let the $\Lambda_{\mathcal{A}}$ -algebra $\mathcal{A}^e = (\text{Tr}(\mathcal{A}), \Lambda_{\mathcal{A}})$ be defined by $\bar{p}^{\mathcal{A}^e}(q) = q \cdot p$ for all $p, q \in \text{Tr}(\mathcal{A})$.

The proof of the main theorem of this section is based on the following lemmas (cf. [8, 9] for similar results for unary algebras).

Lemma 3.4 For any finite algebra \mathcal{A} , $\text{Tr}(\mathcal{A}) \cong \text{Tr}(\mathcal{A}^e)$.

Proof. The elementary translations of \mathcal{A}^e are of the form $\bar{p}^{\mathcal{A}^e}(\xi)$ where $p \in \text{Tr}(\mathcal{A})$, and clearly $\bar{q}^{\mathcal{A}^e}(\xi) \cdot \bar{p}^{\mathcal{A}^e}(\xi) = \overline{q \cdot p}^{\mathcal{A}^e}(\xi)$ for all $q, p \in \text{Tr}(\mathcal{A})$. For the identity translation 1_A of \mathcal{A} the translation $\overline{1_A}^{\mathcal{A}^e}(\xi)$ is the identity translation of \mathcal{A}^e . This means that $\text{Tr}(\mathcal{A}^e) = \{\overline{p}^{\mathcal{A}^e}(\xi) \mid p \in \text{Tr}(\mathcal{A})\}$. Moreover, $\overline{p}^{\mathcal{A}^e}(\xi) \neq \overline{q}^{\mathcal{A}^e}(\xi)$ whenever $p \neq q$, since $\overline{p}^{\mathcal{A}^e}(\xi) = \overline{q}^{\mathcal{A}^e}(\xi)$ implies $p = 1_A \cdot p = \overline{p}^{\mathcal{A}^e}(1_A) = \overline{q}^{\mathcal{A}^e}(1_A) = 1_A \cdot q = q$. Hence, the mapping $\text{Tr}(\mathcal{A}) \rightarrow \text{Tr}(\mathcal{A}^e)$, $p \mapsto \overline{p}^{\mathcal{A}^e}(\xi)$ is a monoid isomorphism. \square

Lemma 3.5 Let $\mathcal{A} = (A, \Sigma)$ and $\mathcal{B} = (B, \Omega)$ be two finite algebras.

1. If $\text{Tr}(\mathcal{A}) \prec \text{Tr}(\mathcal{B})$, then $\mathcal{A}^e \prec_g \mathcal{B}^e$.
2. $\text{Tr}(\mathcal{A}) \times \text{Tr}(\mathcal{B}) \cong \text{Tr}(\kappa(\mathcal{A}^e, \mathcal{B}^e))$ for some g-product $\kappa(\mathcal{A}^e, \mathcal{B}^e)$.

Proof. 1. Suppose $\text{Tr}(\mathcal{A}) \leftarrow M \subseteq \text{Tr}(\mathcal{B})$ for some monoid M . Let $\Lambda_M = \{\bar{p} \in \Lambda_{\mathcal{B}} \mid p \in M\}$. Then clearly $\mathcal{M} = (M, \Lambda_M) \subseteq_g \mathcal{B}^e$, where \mathcal{M} is defined by $\bar{p}^{\mathcal{M}}(q) = q \cdot p$ ($p, q \in M$). Let $\varphi : M \rightarrow \text{Tr}(\mathcal{A})$ be a monoid epimorphism. Define the assignment $\kappa : \Lambda_M \rightarrow \Lambda_{\mathcal{A}}$ by $\bar{q}\kappa = \overline{q\varphi}$ for all $q \in M$. It is clear that κ is surjective and for all $q, r \in M \subseteq \text{Tr}(\mathcal{B})$, $(\overline{q}^{\mathcal{B}^e}(r))\varphi = (r \cdot q)\varphi = r\varphi \cdot q\varphi = \overline{q\varphi}^{\mathcal{A}^e}(r\varphi) = (\overline{q\kappa})^{\mathcal{A}^e}(r\varphi)$. Hence $(\kappa, \varphi) : \mathcal{M} \rightarrow \mathcal{A}^e$ is a g-epimorphism. Thus $\mathcal{A}^e \leftarrow_g \mathcal{M} \subseteq_g \mathcal{B}^e$.

2. Let $\Gamma = \{\overline{\langle p, q \rangle} \mid p \in \text{Tr}(\mathcal{A}), q \in \text{Tr}(\mathcal{B})\}$ be a set of unary function symbols, and define the assignment $\kappa : \Gamma \rightarrow \Lambda_{\mathcal{A}} \times \Lambda_{\mathcal{B}}$ by $\overline{\langle p, q \rangle}\kappa = (\overline{p}, \overline{q})$. Let $\mathcal{P} = \kappa(\mathcal{A}^e, \mathcal{B}^e)$ be the corresponding g-product of \mathcal{A}^e and \mathcal{B}^e . We show that $\text{Tr}(\mathcal{P}) = \{\overline{\langle p, q \rangle}^{\mathcal{P}}(\xi) \mid p \in \text{Tr}(\mathcal{A}), q \in \text{Tr}(\mathcal{B})\}$. Firstly, we note that if 1_A and 1_B are the identity translations of \mathcal{A} and \mathcal{B} respectively, then $\overline{\langle 1_A, 1_B \rangle}^{\mathcal{P}}(\xi)$

is the identity translation of \mathcal{P} . Secondly, by the definition of κ -products, for all $p, p' \in \text{Tr}(\mathcal{A})$, $q, q' \in \text{Tr}(\mathcal{B})$,

$$\overline{\langle p, q \rangle}^{\mathcal{P}}(p', q') = (\overline{p}^{A^e}(p'), \overline{q}^{B^e}(q')) = (p' \cdot p, q' \cdot q).$$

So, if $\overline{\langle p, q \rangle}^{\mathcal{P}}(\xi) = \overline{\langle p', q' \rangle}^{\mathcal{P}}(\xi)$, then $(p, q) = (1_A \cdot p, 1_B \cdot q) = \overline{\langle p, q \rangle}^{\mathcal{P}}(1_A, 1_B) = \overline{\langle p', q' \rangle}^{\mathcal{P}}(1_A, 1_B) = (1_A \cdot p', 1_B \cdot q') = (p', q')$. So, $\overline{\langle p, q \rangle}^{\mathcal{P}}(\xi) \neq \overline{\langle p', q' \rangle}^{\mathcal{P}}(\xi)$, when $p \neq p'$ or $q \neq q'$. Finally, we show that the set $\{\overline{\langle p, q \rangle}^{\mathcal{P}}(\xi) \mid p \in \text{Tr}(\mathcal{A}), q \in \text{Tr}(\mathcal{B})\}$ is closed under the composition of translations.

For all $p, p', p'' \in \text{Tr}(\mathcal{A})$, $q, q', q'' \in \text{Tr}(\mathcal{B})$,

$$\begin{aligned} \overline{\langle p', q' \rangle}^{\mathcal{P}} \cdot \overline{\langle p, q \rangle}^{\mathcal{P}}(p'', q'') &= \overline{\langle p, q \rangle}^{\mathcal{P}}(p'' \cdot p', q'' \cdot q') \\ &= ((p'' \cdot p') \cdot p, (q'' \cdot q') \cdot q) \\ &= (p'' \cdot (p' \cdot p), q'' \cdot (q' \cdot q)) \\ &= \overline{\langle p' \cdot p, q' \cdot q \rangle}^{\mathcal{P}}(p'', q''). \end{aligned}$$

Hence, $\overline{\langle p', q' \rangle}^{\mathcal{P}}(\xi) \cdot \overline{\langle p, q \rangle}^{\mathcal{P}}(\xi) = \overline{\langle p' \cdot p, q' \cdot q \rangle}^{\mathcal{P}}(\xi)$. It follows that the mapping $\text{Tr}(\mathcal{A}) \times \text{Tr}(\mathcal{B}) \rightarrow \text{Tr}(\mathcal{P})$, $(p, q) \mapsto \overline{\langle p, q \rangle}^{\mathcal{P}}(\xi)$, is a monoid isomorphism. \square

Since g-products of g-products are g-isomorphic to a g-product of the original algebras (Lemma 4.2 of [18]), Lemma 3.5(2) can be generalized as follows.

Lemma 3.6 For any algebras $\mathcal{A}_1, \dots, \mathcal{A}_n$ ($n \geq 1$) there is a g-product $\kappa(\mathcal{A}_1^e, \dots, \mathcal{A}_n^e)$ such that $\text{Tr}(\mathcal{A}_1) \times \dots \times \text{Tr}(\mathcal{A}_n) \cong \text{Tr}(\kappa(\mathcal{A}_1^e, \dots, \mathcal{A}_n^e))$.

Now we are ready to prove the main theorem.

Theorem 3.7 Any class of finite algebras \mathbf{K} is definable by translation monoids iff it is a GVFA such that $\mathcal{A} \in \mathbf{K}$ iff $\mathcal{A}^e \in \mathbf{K}$, for any \mathcal{A} .

Proof. Suppose $\mathbf{K} = \mathbf{M}^a$ for a VFM \mathbf{M} . Then by Lemma 3.4, $\text{Tr}(\mathcal{A}) \cong \text{Tr}(\mathcal{A}^e)$, so $\mathcal{A} \in \mathbf{K} \Leftrightarrow \text{Tr}(\mathcal{A}) \in \mathbf{M} \Leftrightarrow \text{Tr}(\mathcal{A}^e) \in \mathbf{M} \Leftrightarrow \mathcal{A}^e \in \mathbf{K}$. For the converse, suppose the GVFA \mathbf{K} satisfies the equivalence $\mathcal{A} \in \mathbf{K} \Leftrightarrow \mathcal{A}^e \in \mathbf{K}$ for any finite algebra \mathcal{A} . Let \mathbf{M} be the VFM generated by $\{\text{Tr}(\mathcal{A}) \mid \mathcal{A} \in \mathbf{K}\}$. We show that $\mathbf{K} = \mathbf{M}^a$. Obviously $\mathbf{K} \subseteq \mathbf{M}^a$. For the opposite inclusion, let $\mathcal{B} \in \mathbf{M}^a$. So, there are $\mathcal{A}_1, \dots, \mathcal{A}_m \in \mathbf{K}$ such that $\text{Tr}(\mathcal{B}) \prec \text{Tr}(\mathcal{A}_1) \times \dots \times \text{Tr}(\mathcal{A}_m)$. By Lemma 3.6, $\text{Tr}(\mathcal{B}) \prec \text{Tr}(\mathcal{P})$ for some g-product \mathcal{P} of $\mathcal{A}_1^e, \dots, \mathcal{A}_m^e$. By the property of \mathbf{K} , $\mathcal{A}_1^e, \dots, \mathcal{A}_m^e \in \mathbf{K}$, and so $\mathcal{P} \in \mathbf{K}$, hence $\mathcal{P}^e \in \mathbf{K}$. By Lemma 3.5 (1) from $\text{Tr}(\mathcal{B}) \prec \text{Tr}(\mathcal{P})$ we get $\mathcal{B}^e \prec_g \mathcal{P}^e$, and since $\mathcal{P}^e \in \mathbf{K}$, also $\mathcal{B}^e \in \mathbf{K}$, which implies that $\mathcal{B} \in \mathbf{K}$. Thus $\mathbf{M}^a \subseteq \mathbf{K}$. \square

Remark 3.8 The proof of Theorem 3.7 also yields the fact that for any GVFA \mathbf{K} definable by translation monoids, the class $\{\text{Tr}(\mathcal{A}) \mid \mathcal{A} \in \mathbf{K}\}$ is a variety of finite monoids.

Another characterization of the classes of finite algebras definable by translation monoids which follows from Lemmas 3.4 and 3.5 is the following.

Theorem 3.9 Any class of finite algebras \mathbf{K} is definable by translation monoids iff it is a GVFA such that for all finite algebras \mathcal{A} and \mathcal{B} , if $\text{Tr}(\mathcal{A}) \cong \text{Tr}(\mathcal{B})$ and $\mathcal{A} \in \mathbf{K}$, then $\mathcal{B} \in \mathbf{K}$.

4 Families of Tree Languages Definable by Syntactic Monoids

A *general variety of tree languages* (GVTL) is a family $\mathcal{V} = \{\mathcal{V}(\Sigma, X)\}$ which assigns to every ranked alphabet Σ and leaf alphabet X , a set $\mathcal{V}(\Sigma, X)$ of recognizable ΣX -tree languages, and is closed under all Boolean operations, inverse translations, and inverse g-morphisms. That is to say, for any ranked alphabets Σ, Ω , leaf alphabets X, Y , context $P \in \mathbf{C}(\Sigma, X)$, and g-morphism $\varphi : \mathbf{T}(\Omega, Y) \rightarrow \mathbf{T}(\Sigma, X)$ (see Definition 3.1), if $T, T' \in \mathcal{V}(\Sigma, X)$, then $\mathbf{T}(\Sigma, X) \setminus T, T \cap T', P^{-1}(T) \in \mathcal{V}(\Sigma, X)$, and $T\varphi^{-1} \in \mathcal{V}(\Omega, Y)$ (Definition 7.1 of [18]).

For a family of recognizable tree languages \mathcal{V} , $\mathcal{V}^{\mathbf{a}}$ is the GVFA generated by the class $\{\text{SA}(T) \mid T \in \mathcal{V}(\Sigma, X), \text{ for some } \Sigma, X\}$.

Remark 4.1 The General Variety Theorem in [18], Proposition 9.15, implies that:

- (1) For any GVTL \mathcal{V} , the class $\mathcal{V}^{\mathbf{a}}$ satisfies the following equivalence for any tree language $T \subseteq \mathbf{T}(\Sigma, X)$: $T \in \mathcal{V}(\Sigma, X) \Leftrightarrow \text{SA}(T) \in \mathcal{V}^{\mathbf{a}}$.
- (2) For any GVFA \mathbf{K} there is a unique GVTL \mathcal{V} such that $\mathcal{V}^{\mathbf{a}} = \mathbf{K}$.

Definition 4.2 For a VFM \mathbf{M} , let $\mathbf{M}^{\mathbf{t}}$ be the family of all recognizable tree languages whose syntactic monoids are in \mathbf{M} , that is to say for any tree language $T \subseteq \mathbf{T}(\Sigma, X)$, $T \in \mathbf{M}^{\mathbf{t}}(\Sigma, X) \Leftrightarrow \text{SM}(T) \in \mathbf{M}$ holds.

A family of recognizable tree languages \mathcal{V} is said to be *definable by syntactic monoids* if there is a VFM \mathbf{M} such that $\mathbf{M}^{\mathbf{t}} = \mathcal{V}$.

Steinby has shown that for any VFM \mathbf{M} , $\mathbf{M}^{\mathbf{t}}$ is a GVTL ([18], Proposition 10.3). His proof can be applied to show that $\mathbf{M}^{\mathbf{t}}$ is also closed under inverse of regular tree homomorphisms. The general varieties of tree languages closed under inverse (arbitrary) tree homomorphisms are studied by Ésik [4] who characterized them by their *syntactic theories*. Theorem 14.2 of [4] establishes a correspondence between d-varieties of finite algebras and general tree language varieties closed under inverse tree homomorphisms. However, those varieties may not be definable by syntactic monoids, as the following example shows.

Example 4.3 Let $\text{Def}_1 = \{\text{Def}_1(\Sigma, X)\}$ be the family of 1-definite tree languages, i.e., $T \in \text{Def}_1(\Sigma, X)$ iff for all ΣX -trees t and s , $\text{root}(t) = \text{root}(s)$ and $t \in T$ imply $s \in T$, where $\text{root}(t)$ is the root symbol of t . It is a

GVTL ([18]) which can be shown to be closed under inverse *strict* regular tree homomorphisms (see [4] Subsection 11.1 and Section 5 below). Let $\Sigma = \Sigma_2 = \{f, g\}$, $X = \{x, y\}$, and $T = \{x\} \cup \{f(t_1, t_2) \mid t_1, t_2 \in T(\Sigma, X)\}$. Clearly $T \in \text{Def}_1(\Sigma, X)$. It can be easily shown that the syntactic monoid of T consists of an identity element and two right zeros. This is also the syntactic monoid of the language T' of the ΣX -trees whose leftmost leaves are x , by Example 10.4 of [18]. Since $T' \notin \text{Def}_1(\Sigma, X)$, then Def_1 is not definable by syntactic monoids.

This actually shows that the GVTL of all definite tree languages is not definable by syntactic monoids, since T' is not k -definite for any $k \geq 1$.

Remark 4.4 In [7] it is claimed that the variety of definite tree languages can be characterized by the property that all the non-identity idempotents of their syntactic monoids are right zeros (left zeros in the formalism of [7]). This clearly stands in conflict with the above Example 4.3.

Indeed, it can be shown that Theorem 1 of [7] does not hold. When the syntactic semigroup of a tree language is defined as the syntactic monoid with the identity element removed, the authors clearly overlook the possibility that the identity element may be obtained also as the product of some non-identity elements, and the proof of the theorem of [7] holds in just one direction. A concrete example showing that the equality between lines 9 and 10 on page 189 does not necessarily hold, can be obtained by considering the tree language T' of our Example 4.3.

It can also be noted that finite monoids whose non-identity idempotents are right zeros, do not form a VFM. Finally, in Section 5 we shall see that a more appropriate definition of the syntactic semigroup and omitting trees that in a sense correspond to the empty word, does not save the result of [7].

We shall characterize the general varieties of tree languages that are definable by syntactic monoids by requiring them to satisfy two more conditions in addition to being closed under inverse regular tree homomorphisms.

Definition 4.5 A regular tree homomorphism $\varphi : T(\Sigma, X) \rightarrow T(\Omega, Y)$ is said to be *full with respect to* a tree language $T \subseteq T(\Omega, Y)$, if for every $Q \in C(\Omega, Y)$ and every $s \in T(\Omega, Y)$, there are $P \in C(\Sigma, X)$ and $t \in T(\Sigma, X)$, such that $Q \mu_T P \varphi_*$ and $s \theta_T t \varphi$ hold.

Remark 4.6 At first glance it seems that verifying fullness of φ with respect to T requires checking the existence of $P \in C(\Sigma, X)$ and $t \in T(\Sigma, X)$ for all (infinitely many) $Q \in C(\Omega, Y)$ and $s \in T(\Omega, Y)$ such that $Q \mu_T P \varphi_*$ and $s \theta_T t \varphi$ hold. In fact it is decidable for a recognizable T to check whether or not φ is full with respect to T : let $\varphi^T : T(\Omega, Y) \rightarrow T(\Omega, Y)/\theta_T$, $t\varphi^T = t/\theta_T$ and $\lambda^T : C(\Omega, Y) \rightarrow C(\Omega, Y)/\mu_T$, $P\lambda^T = P/\mu_T$ be the natural morphisms. Then the tree homomorphism $\varphi : T(\Sigma, X) \rightarrow T(\Omega, Y)$ is full with respect to T iff both the mappings $\varphi\varphi^T : T(\Sigma, X) \rightarrow T(\Omega, Y)/\theta_T$ and $\varphi_*\lambda^T : C(\Sigma, X) \rightarrow C(\Omega, Y)/\mu_T$ are surjective.

Recall that for an equivalence relation θ on a set A , the quotient set of A under θ is denoted by A/θ , and $a\theta$ is the equivalence θ -class containing $a \in A$.

Lemma 4.7 If $\varphi : T(\Sigma, X) \rightarrow T(\Omega, Y)$ is a regular tree homomorphism and $T \subseteq T(\Omega, Y)$, then $\text{SM}(T\varphi^{-1}) \prec \text{SM}(T)$, and if φ is full with respect to T , then $\text{SM}(T\varphi^{-1}) \cong \text{SM}(T)$.

Proof. We note that $\varphi_* : C(\Sigma, X) \rightarrow C(\Omega, Y)$ is a monoid homomorphism. Let $S \subseteq C(\Omega, Y)$ be the image of φ_* , and let μ be the restriction of μ_T to S . Then S/μ is a submonoid of $C(\Omega, Y)/\mu_T$. We show that $P\varphi_*\mu Q\varphi_*$ implies $P\mu_{T\varphi^{-1}}Q$ for all $P, Q \in C(\Sigma, X)$.

Suppose $P\varphi_*\mu Q\varphi_*$ and take arbitrary $t \in T(\Omega, Y)$ and $R \in C(\Omega, Y)$. Then

$$\begin{aligned} t \cdot P \cdot R \in T\varphi^{-1} &\Leftrightarrow t\varphi \cdot P\varphi_* \cdot R\varphi_* \in T \\ &\Leftrightarrow t\varphi \cdot Q\varphi_* \cdot R\varphi_* \in T \\ &\Leftrightarrow t \cdot Q \cdot R \in T\varphi^{-1}, \end{aligned}$$

that is $P\mu_{T\varphi^{-1}}Q$. So the mapping $\psi : S/\mu \rightarrow C(\Sigma, X)/\mu_{T\varphi^{-1}}$ defined by $((P\varphi_*)\mu)\psi = P\mu_{T\varphi^{-1}}$ is well-defined and surjective. It is also a monoid homomorphism, since $((P\varphi_*)\mu \cdot (Q\varphi_*)\mu)\psi = ((P \cdot Q)\varphi_*\mu)\psi = (P \cdot Q)\mu_{T\varphi^{-1}} = P\mu_{T\varphi^{-1}} \cdot Q\mu_{T\varphi^{-1}} = ((P\varphi_*)\mu)\psi \cdot ((Q\varphi_*)\mu)\psi$ for all $P, Q \in C(\Sigma, X)$. Hence $\text{SM}(T\varphi^{-1}) \leftarrow S/\mu \subseteq \text{SM}(T)$, so $\text{SM}(T\varphi^{-1}) \prec \text{SM}(T)$.

Now, suppose φ is full with respect to T . We show $P\mu_{T\varphi^{-1}}Q$ iff $P\varphi_*\mu_T Q\varphi_*$ for any $P, Q \in C(\Sigma, X)$. Clearly, $P\varphi_*\mu_T Q\varphi_*$ implies $P\mu_{T\varphi^{-1}}Q$. For the converse, suppose $P\mu_{T\varphi^{-1}}Q$, and take arbitrary $R' \in C(\Omega, Y)$, and $t' \in T(\Omega, Y)$. There are $R \in C(\Sigma, X)$ and $t \in T(\Sigma, X)$ such that $R\varphi_*\mu_T R'$ and $t\varphi\theta_T t'$. Hence

$$\begin{aligned} t' \cdot P\varphi_* \cdot R' \in T &\Leftrightarrow t\varphi \cdot P\varphi_* \cdot R\varphi_* \in T \\ &\Leftrightarrow (t \cdot P \cdot R)\varphi \in T \\ &\Leftrightarrow t \cdot P \cdot R \in T\varphi^{-1} \\ &\Leftrightarrow t \cdot Q \cdot R \in T\varphi^{-1} \\ &\Leftrightarrow t\varphi \cdot Q\varphi_* \cdot R\varphi_* \in T \\ &\Leftrightarrow t' \cdot Q\varphi_* \cdot R' \in T, \end{aligned}$$

which shows that $P\varphi_*\mu_T Q\varphi_*$. Hence $P\mu_{T\varphi^{-1}}Q$ iff $P\varphi_*\mu_T Q\varphi_*$, and since the function $\varphi_* : C(\Sigma, X) \rightarrow C(\Omega, Y)$ is a monoid homomorphism, the mapping $C(\Sigma, X)/\mu_{T\varphi^{-1}} \rightarrow C(\Omega, Y)/\mu_T$, $P\mu_{T\varphi^{-1}} \mapsto (P\varphi_*)\mu_T$ is a monoid isomorphism between $\text{SM}(T\varphi^{-1})$ and $\text{SM}(T)$. \square

In the following example we show that the regularity condition on φ in the previous lemma can not be relaxed.

Example 4.8 Define the ranked alphabets $\Omega = \Omega_2 = \{f\}$ and $\Sigma = \Sigma_1 = \{g, h\}$, and the leaf alphabet $X = \{u, v, w\}$. Let $(\mathbb{Z}_3, +)$ be the cyclic group of order 3. Define $\chi : T(\Omega, X) \rightarrow \mathbb{Z}_3$ inductively by $u\chi = 0$, $v\chi = 1$, $w\chi = 2$, and $f(t, s)\chi = t\chi + s\chi$. Let $T = \{0\}\chi^{-1}$. It is easy to see that the

syntactic monoid of T consists of the μ_T -classes of the elementary contexts $f(u, \xi), f(v, \xi), f(w, \xi)$, and in fact $\text{SM}(T) \simeq (\mathbb{Z}_3, +)$.

Define the tree homomorphisms $\varphi, \psi : \text{T}(\Sigma, X) \rightarrow \text{T}(\Omega, X)$ by $\varphi_X(x) = \psi_X(x) = x$ for $x \in X$, and $\varphi_1(g) = \psi_1(g) = f(v, \xi)$, $\varphi_1(h) = f(\xi, \xi)$, and $\psi_1(h) = u$. These tree homomorphisms are not regular: ξ appears twice in $\varphi_1(h)$ and does not appear at all in $\psi_1(h)$.

We show that neither $\text{SM}(T\varphi^{-1})$ nor $\text{SM}(T\psi^{-1})$ can divide $\text{SM}(T)$. The following identities can be verified by straightforward computations:

$$\begin{aligned} - (v \cdot h(\xi) \cdot g(\xi))\varphi\chi &= 0, & (v \cdot g(\xi) \cdot h(\xi))\varphi\chi &= 1, \text{ and} \\ - (v \cdot h(\xi) \cdot g(\xi))\psi\chi &= 1, & (v \cdot g(\xi) \cdot h(\xi))\psi\chi &= 0. \end{aligned}$$

So, $(h(\xi) \cdot g(\xi), g(\xi) \cdot h(\xi)) \notin \mu_{T\varphi^{-1}}, \mu_{T\psi^{-1}}$ which proves that $\text{SM}(T\varphi^{-1})$ and $\text{SM}(T\psi^{-1})$ are not commutative.

Remark 4.9 Let \mathbf{C} be the variety of finite commutative monoids. By Example 4.8, the GVTL \mathbf{C}^\dagger is not closed under inverse non-regular tree homomorphisms; cf. Theorem 4.14. So, \mathbf{C}^\dagger is not definable by syntactic theories in the sense of [4]. On the other hand, by Example 4.3, the family of definite tree languages is not definable by syntactic monoids, even though it is definable by syntactic theories, cf. [4] Subsection 11.1.

Thus, the concepts of “definability by syntactic theories” and of “definability by syntactic monoids” are not comparable to each other, though they are both weaker than “definability by syntactic algebras”.

Lemma 4.10 Let $\mathcal{A} = (A, \Sigma)$ be a finite algebra, and X be a leaf alphabet disjoint from A . For any tree language $L \subseteq \text{T}(\Lambda_{\mathcal{A}}, X)$ recognized by \mathcal{A}^e , there exists a regular tree homomorphism $\varphi : \text{T}(\Lambda_{\mathcal{A}}, X) \rightarrow \text{T}(\Sigma, X \cup A)$, and a tree language $T \subseteq \text{T}(\Sigma, X \cup A)$ such that $L = T\varphi^{-1}$, and T can be recognized by a finite power \mathcal{A}^n where $n = |A|$.

Proof. Let $\alpha : X \rightarrow \text{Tr}(\mathcal{A})$ be an initial assignment for \mathcal{A}^e and $F \subseteq \text{Tr}(\mathcal{A})$ be a subset such that $L = \{t \in \text{T}(\Lambda_{\mathcal{A}}, X) \mid t\alpha^{\mathcal{A}^e} \in F\}$. Define the tree homomorphism $\varphi : \text{T}(\Lambda_{\mathcal{A}}, X) \rightarrow \text{T}(\Sigma, X \cup A)$ by $\varphi_X(x) = x$ for all $x \in X$, and for every $p \in \text{Tr}(\mathcal{A})$ choose a $\varphi_1(\bar{p}) \in \mathbf{C}(\Sigma, A)$ such that $\varphi_1(\bar{p})^{\mathcal{A}} = p$. Obviously φ is a regular tree homomorphism. Suppose that $A = \{a_1, \dots, a_n\}$. Let $F' = \{(p(a_1), \dots, p(a_n)) \in A^n \mid p \in F\}$, and define the initial assignment $\beta : X \cup A \rightarrow A^n$ for \mathcal{A}^n by $x\beta = ((x\alpha)(a_1), \dots, (x\alpha)(a_n))$ for all $x \in X$, and $a\beta = (a, \dots, a) \in A^n$ for all $a \in A$. Let T be the subset of $\text{T}(\Sigma, X \cup A)$ recognized by $(\mathcal{A}^n, \beta, F')$. We show that $L = T\varphi^{-1}$. Every tree w in $\text{T}(\Lambda_{\mathcal{A}}, X)$ is of the form $w = \bar{p}_1(\bar{p}_2(\dots \bar{p}_k(x) \dots))$ for some $p_1, \dots, p_k \in \text{Tr}(\mathcal{A})$ ($k \geq 0$) and $x \in X$. For such a tree w ,

$$\begin{aligned} w\alpha^{\mathcal{A}^e} &= x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1, \text{ and} \\ (w\varphi)\beta^{\mathcal{A}^n} &= (x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1(a_1), \dots, x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1(a_n)). \text{ So,} \\ w\varphi \in T &\Leftrightarrow (w\varphi)\beta^{\mathcal{A}^n} \in F' \\ &\Leftrightarrow \text{for some } p \in F, p(a) = x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1(a) \text{ for all } a \in A \\ &\Leftrightarrow x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1 \in F \\ &\Leftrightarrow w\alpha^{\mathcal{A}^e} \in F \\ &\Leftrightarrow w \in L. \end{aligned}$$

□

Lemma 4.11 Let $\mathcal{A} = (A, \Sigma)$ be a finite algebra and X be a leaf alphabet disjoint from $A \cup \Sigma$. For any tree language $T \subseteq \mathsf{T}(\Sigma, X)$ recognized by \mathcal{A} there exists a unary ranked alphabet Λ , and a regular tree homomorphism $\varphi : \mathsf{T}(\Lambda, X \cup \Sigma_0) \rightarrow \mathsf{T}(\Sigma, X)$ such that φ is full with respect to T , and for every $z \in X \cup \Sigma_0$, $T\varphi^{-1} \cap \mathsf{T}(\Lambda, \{z\})$ can be recognized as a subset of $\mathsf{T}(\Lambda, \{z\})$ by \mathcal{A}^e .

Proof. Let $\mathcal{B} = (B, \Sigma)$ be the syntactic algebra of T . Then $\mathcal{B} \prec \mathcal{A}$. Suppose $T = \{t \in \mathsf{T}(\Sigma, X) \mid t\beta^{\mathcal{B}} \in F\}$, where $\beta : X \rightarrow B$ is an initial assignment for \mathcal{B} and $F \subseteq B$. Since \mathcal{B} is the minimal tree automaton recognizing T , the set B is generated by $\beta(X)$. The mapping $\beta : X \rightarrow B$ can be uniquely extended to a monoid homomorphism $\beta_c : \mathsf{C}(\Sigma, X) \rightarrow \mathsf{C}(\Sigma, B)$. Since B is generated by $\beta(X)$, the mapping $\beta_c^{\mathcal{B}} : \mathsf{C}(\Sigma, X) \rightarrow \mathsf{Tr}(\mathcal{B})$, $\beta_c^{\mathcal{B}}(Q) = \beta_c(Q)^{\mathcal{B}}$ is surjective. Define the tree homomorphism $\varphi : \mathsf{T}(\Lambda_{\mathcal{B}}, X \cup \Sigma_0) \rightarrow \mathsf{T}(\Sigma, X)$ by $\varphi_X(x) = x$ for all $x \in X \cup \Sigma_0$, and for every $q \in \mathsf{Tr}(\mathcal{B})$ choose a $\varphi_1(\bar{q}) = Q \in \mathsf{C}(\Sigma, X)$ such that $\beta_c(Q)^{\mathcal{B}} = q$. Note that φ is a regular tree homomorphism. It remains to show that φ is full with respect to T and that for every $z \in X \cup \Sigma_0$, $L_z = T\varphi^{-1} \cap \mathsf{T}(\Lambda, \{z\})$ can be recognized as a subset of $\mathsf{T}(\Lambda, \{z\})$ by \mathcal{B}^e . This will finish the proof since $\mathsf{Tr}(\mathcal{B}) \prec \mathsf{Tr}(\mathcal{A})$ follows from $\mathcal{B} \prec \mathcal{A}$ by Lemma 10.7 of [18], and so $\mathcal{B}^e \prec \mathcal{A}^e$ by Lemma 3.5, which implies that L_z can also be recognized by \mathcal{A}^e .

Firstly, we show that φ is full with respect to T . Let $Q \in \mathsf{C}(\Sigma, X)$ be a context. For $q = \beta_c(Q)^{\mathcal{B}} \in \mathsf{Tr}(\mathcal{B})$, $\bar{q}(\xi)\varphi_*\mu_T Q$ holds. By induction on the height of t we show that for any $t \in \mathsf{T}(\Sigma, X)$ there is an $s \in \mathsf{T}(\Lambda_{\mathcal{B}}, X \cup \Sigma_0)$ such that $t\theta_T s\varphi$. If $t = x \in X \cup \Sigma_0$, then $s\varphi\theta_T t$ for $s = t$. If $t = t' \cdot P$ for some $P \in \mathsf{C}(\Sigma, X)$ and $t' \in \mathsf{T}(\Sigma, X)$ such that the height of t' is less than the height of t , then by the induction hypothesis there is an $s' \in \mathsf{T}(\Lambda_{\mathcal{B}}, X \cup \Sigma_0)$ such that $t'\theta_T s'\varphi$. Also, for some $p \in \mathsf{Tr}(\mathcal{B})$, $\bar{p}(\xi)\varphi_*\mu_T P$ holds. Let $s = \bar{p}(s')$. Then

$$s\varphi = s'\varphi \cdot \bar{p}(\xi)\varphi_*\theta_T t' \cdot P = t.$$

Secondly, we show that L_z can be recognized by \mathcal{B}^e for a fixed $z \in X \cup \Sigma_0$. Let $1_{\mathcal{B}}$ be the identity translation of \mathcal{B} . Define the initial assignment $\alpha : \{z\} \rightarrow \mathsf{Tr}(\mathcal{B})$ for \mathcal{B}^e by $z\alpha = 1_{\mathcal{B}}$, and let $F_z = \{q \in \mathsf{Tr}(\mathcal{B}) \mid q(z\beta^{\mathcal{B}}) \in F\}$. We show that L_z is recognized by $(\mathcal{B}^e, \alpha, F_z)$. Every $w \in \mathsf{T}(\Lambda_{\mathcal{B}}, \{z\})$ can be written in the form

$$w = \bar{q}_1 \left(\bar{q}_2 \left(\dots \bar{q}_h(z) \dots \right) \right)$$

for some $q_1, \dots, q_h \in \mathsf{Tr}(\mathcal{B})$ ($h \geq 0$). For such a tree w ,

$$w\alpha^{\mathcal{B}^e} = 1_{\mathcal{B}} \cdot q_h \cdot \dots \cdot q_2 \cdot q_1, \text{ and } (w\varphi)\beta^{\mathcal{B}} = q_h \cdot \dots \cdot q_2 \cdot q_1(z\beta^{\mathcal{B}}). \text{ Thus,}$$

$$\begin{aligned} w \in L_z &\Leftrightarrow w\varphi \in T &\Leftrightarrow (w\varphi)\beta^{\mathcal{B}} \in F \\ &&\Leftrightarrow q_h \cdot \dots \cdot q_2 \cdot q_1(z\beta^{\mathcal{B}}) \in F \\ &&\Leftrightarrow q_h \cdot \dots \cdot q_2 \cdot q_1 \in F_z \\ &&\Leftrightarrow w\alpha^{\mathcal{B}^e} \in F_z. \end{aligned}$$

So, $L_z = \{w \in \mathsf{T}(\Lambda, \{z\}) \mid w\alpha^{\mathcal{B}^e} \in F_z\}$. □

We end the section by proving a Variety Theorem for tree languages and syntactic monoids, and presenting some examples that justify the theorem (another interesting example is presented in [12]).

Before presenting the main theorem we note two remarks.

Remark 4.12 Let Λ be a unary ranked alphabet. For every leaf alphabet X and every subset $Y \subseteq X$, $C(\Lambda, Y) = C(\Lambda, X)$, and the relation μ_T for a tree language $T \subseteq T(\Lambda, Y)$ on $C(\Lambda, Y)$ is the same relation μ_T on $C(\Lambda, X)$ when T is viewed as a subset of $T(\Lambda, X)$.

So, if a family of tree languages $\mathcal{V} = \{\mathcal{V}(\Sigma, X)\}$ is definable by syntactic monoids, then for every unary ranked alphabet Λ , and any leaf alphabets X and Y , if $Y \subseteq X$ then $\mathcal{V}(\Lambda, Y) \subseteq \mathcal{V}(\Lambda, X)$.

Recall the notion of \mathcal{V}^a at the beginning of the section.

Remark 4.13 By Propositions 6.13 and 5.8(b) of [18] it follows that every finite algebra can be represented as a subdirect product of the syntactic algebras of some tree languages that are recognizable by the algebra. This implies that for any GVTL \mathcal{V} and any finite algebra \mathcal{A} , if every tree language recognizable by \mathcal{A} belongs to \mathcal{V} , then $\mathcal{A} \in \mathcal{V}^a$.

Theorem 4.14 A family of recognizable tree languages \mathcal{V} is definable by syntactic monoids iff \mathcal{V} is a GVTL that is closed under inverse regular tree homomorphisms and satisfies the following conditions:

- (1) For every unary ranked alphabet Λ , and any leaf alphabets X and Y , if $Y \subseteq X$ then $\mathcal{V}(\Lambda, Y) \subseteq \mathcal{V}(\Lambda, X)$.
- (2) For any regular tree homomorphism $\varphi : T(\Sigma, X) \rightarrow T(\Omega, Y)$ which is full with respect to a tree language $T \subseteq T(\Omega, Y)$, if $T\varphi^{-1} \in \mathcal{V}(\Sigma, X)$ then $T \in \mathcal{V}(\Omega, Y)$.

Proof. That for any VFM \mathbf{M} , \mathbf{M}^t satisfies the conditions of Theorem 4.14 follows from Lemma 4.7, Remark 4.12, and the facts mentioned at the beginning of the section. For the converse, suppose the GVTL \mathcal{V} satisfies the conditions presented in the theorem. We complete the proof of the theorem by showing that \mathcal{V}^a satisfies the condition of Theorem 3.7. Indeed, Theorem 3.7 implies then that there is a VFM \mathbf{M} such that $\mathcal{V}^a = \mathbf{M}^a$, and

$$T \in \mathcal{V} \Leftrightarrow \text{SA}(T) \in \mathcal{V}^a \Leftrightarrow \text{Tr}(\text{SA}(T)) \in \mathbf{M} \Leftrightarrow \text{SM}(T) \in \mathbf{M}$$

holds for every tree language T by Remarks 4.1 and 2.1, which proves that $\mathcal{V} = \mathbf{M}^t$. So, all we have to show is that $\mathcal{A} \in \mathcal{V}^a$ iff $\mathcal{A}^\ell \in \mathcal{V}^a$ for any \mathcal{A} .

Let $\mathcal{A} = (A, \Sigma)$ be a finite algebra in \mathcal{V}^a . By Lemma 4.10, any tree language $L \subseteq T(\Lambda_{\mathcal{A}}, X)$ recognized by \mathcal{A}^ℓ can be written as $L = T\varphi^{-1}$, where $\varphi : T(\Lambda_{\mathcal{A}}, X) \rightarrow T(\Sigma, X \cup A)$ is a regular tree homomorphism, and T is a tree language recognized by some power \mathcal{A}^n of \mathcal{A} . Then $\mathcal{A}^n \in \mathcal{V}^a$ implies that $T \in \mathcal{V}(\Sigma, X \cup A)$, and hence $L = T\varphi^{-1} \in \mathcal{V}(\Lambda_{\mathcal{A}}, X)$. This holds for every tree language L recognizable by \mathcal{A}^ℓ , so $\mathcal{A}^\ell \in \mathcal{V}^a$ by Remark 4.13.

Now, suppose $\mathcal{A}^e \in \mathcal{V}^a$ for a finite algebra $\mathcal{A} = (A, \Sigma)$. Let $T \subseteq \mathsf{T}(\Sigma, X)$ be a tree language recognizable by \mathcal{A} . By Lemma 4.11, there is a unary ranked alphabet Λ and a regular tree homomorphism $\varphi : \mathsf{T}(\Lambda, X \cup \Sigma_0) \rightarrow \mathsf{T}(\Sigma, X)$ full with respect to T such that for every $z \in X \cup \Sigma_0$, $L_z = T\varphi^{-1} \cap \mathsf{T}(\Lambda, \{z\})$ can be recognized by \mathcal{A}^e as a subset of $\mathsf{T}(\Lambda, \{z\})$. So, $L_z \in \mathcal{V}(\Lambda, \{z\})$, thus $L_z \in \mathcal{V}(\Lambda, X \cup \Sigma_0)$. Hence $T\varphi^{-1} = \bigcup_{z \in X \cup \Sigma_0} L_z \in \mathcal{V}(\Lambda, X \cup \Sigma_0)$. Since φ is full with respect to T , then $T \in \mathcal{V}(\Sigma, X)$. This holds for every tree language T recognizable by \mathcal{A} , hence $\mathcal{A} \in \mathcal{V}^a$ by Remark 4.13. \square

Example 4.15 It was shown in Example 4.3 that Def_1 is not definable by syntactic monoids. Here we show that it does not satisfy condition (2) of Theorem 4.14. Let Σ, X, T, T' be as in Example 4.3. Define the regular tree homomorphism $\varphi : \mathsf{T}(\Sigma, X) \rightarrow \mathsf{T}(\Sigma, X)$, by $\varphi_X(x) = x$, $\varphi_X(y) = y$, and $\varphi_2(f) = f(x, f(\xi_1, \xi_2))$, $\varphi_2(g) = g(y, g(\xi_1, \xi_2))$. Now φ is full with respect to T' since for any $t \in \mathsf{T}(\Sigma, X)$, if $t \in T'$ then $f(y, x)\varphi\theta_{T'}t$, and if $t \notin T'$ then $g(y, x)\varphi\theta_{T'}t$. Similarly, for $P \in \mathsf{C}(\Sigma, X)$, if the leftmost leaf of P is x then $f(y, \xi)\varphi_*\mu_{T'}P$, if the leftmost leaf of P is y then $g(y, \xi)\varphi_*\mu_{T'}P$, and if the leftmost leaf of P is ξ then $\xi\varphi_*\mu_{T'}P$. Clearly $T'\varphi^{-1} = T$, since for any $t \in \mathsf{T}(\Sigma, X)$, the leftmost leaf of $t\varphi$ is x iff either $t = x$ or the root of t is f . By Example 4.3, $T'\varphi^{-1} = T \in \text{Def}_1$, but $T' \notin \text{Def}_1$.

Example 4.16 Let $\text{Ap} = \{\text{Ap}(\Sigma, X)\}$ be the family of aperiodic tree languages. It was shown to be a GVTL in Example 7.8 of [18]. It is also known that Ap is definable by the variety of aperiodic (syntactic) monoids, see [20]. The argument of Example 7.8 in [18] showing that Ap is closed under inverse g-morphisms can be applied to show that Ap is in fact closed under inverse regular tree homomorphisms. It is also straightforward to see that Ap satisfies condition (1) of Theorem 4.14. We show that it also satisfies condition (2). Suppose $\varphi : \mathsf{T}(\Sigma, X) \rightarrow \mathsf{T}(\Omega, Y)$ is a regular tree homomorphism full with respect to $T \subseteq \mathsf{T}(\Omega, Y)$, and $T\varphi^{-1} \in \text{Ap}(\Sigma, X)$. There is an n such that for all $t \in \mathsf{T}(\Sigma, X)$ and all $P, Q \in \mathsf{C}(\Sigma, X)$, $t \cdot P^n \cdot Q \in T\varphi^{-1} \Leftrightarrow t \cdot P^{n+1} \cdot Q \in T\varphi^{-1}$. For any $s \in \mathsf{T}(\Omega, Y)$ and any $R, U \in \mathsf{C}(\Omega, Y)$, there are $t \in \mathsf{T}(\Sigma, X)$ and $P, Q \in \mathsf{C}(\Sigma, X)$ such that $t\varphi\theta_Ts$, $P\varphi_*\mu_T R$, and $Q\varphi_*\mu_T U$. So, $s \cdot R^n \cdot U \in T \Leftrightarrow t\varphi \cdot P^n\varphi_* \cdot Q\varphi_* \in T \Leftrightarrow t \cdot P^n \cdot Q \in T\varphi^{-1} \Leftrightarrow t \cdot P^{n+1} \cdot Q \in T\varphi^{-1} \Leftrightarrow t\varphi \cdot P^{n+1}\varphi_* \cdot Q\varphi_* \in T \Leftrightarrow s \cdot R^{n+1} \cdot U \in T$, which shows that $T \in \text{Ap}(\Omega, Y)$.

Example 4.17 The family of nilpotent tree languages $\text{Nil} = \{\text{Nil}(\Sigma, X)\}$ which consists of finite and cofinite tree languages is a GVFA (see [18], Example 7.5). Let $\Lambda = \Lambda_1 = \{\alpha\}$ be a unary ranked alphabet and $X = \{x, y\}$ be a leaf alphabet. Let $T = \{\alpha(y), \alpha(\alpha(y)), \alpha(\alpha(\alpha(y))), \dots\}$. Clearly $T \in \text{Nil}(\Lambda, \{y\})$, but $T \notin \text{Nil}(\Lambda, X)$. Hence, Nil does not satisfy the condition (1) of Theorem 4.14, so it is not definable by syntactic monoids.

5 Definability by Semigroups

In this section, we show how to modify the above results as to yield characterizations of varieties of finite algebras definable by translation semigroups and of varieties of tree languages definable by syntactic semigroups.

5.1 Algebras Definable by Translation Semigroups

The difference between the translation monoid and the translation semigroup of an algebra is that the latter does not automatically contain the identity translation, although it may be included as an elementary translation or as a composition of some elementary translations.

Denote the translation semigroup of an algebra $\mathcal{A} = (A, \Sigma)$ by $\text{TrS}(\mathcal{A})$ and let $\Lambda_{\mathcal{A}}$ be as in Definition 3.3 except that $\text{Tr}(\mathcal{A})$ is replaced with $\text{TrS}(\mathcal{A})$. We associate with \mathcal{A} a new symbol $\mathbf{I}_{\mathcal{A}}$ that does not appear in $A \cup \Sigma \cup \text{TrS}(\mathcal{A})$. Define the $\Lambda_{\mathcal{A}}$ -algebra $\mathcal{A}^{\mathcal{S}} = (\text{TrS}(\mathcal{A}) \cup \{\mathbf{I}_{\mathcal{A}}\}, \Lambda_{\mathcal{A}})$ by $\bar{p}^{\mathcal{A}^{\mathcal{S}}}(q) = q \cdot p$ and $\bar{p}^{\mathcal{A}^{\mathcal{S}}}(\mathbf{I}_{\mathcal{A}}) = p$ for all $p, q \in \text{TrS}(\mathcal{A})$.

Lemma 5.1 For any finite algebras $\mathcal{A} = (A, \Sigma)$ and $\mathcal{B} = (B, \Omega)$,

- (1) $\text{TrS}(\mathcal{A}) \cong \text{TrS}(\mathcal{A}^{\mathcal{S}})$;
- (2) If $\text{TrS}(\mathcal{A}) \prec \text{TrS}(\mathcal{B})$, then $\mathcal{A}^{\mathcal{S}} \prec_g \mathcal{B}^{\mathcal{S}}$; and
- (3) $\text{TrS}(\mathcal{A}) \times \text{TrS}(\mathcal{B}) \cong \text{Tr}(\kappa(\mathcal{A}^{\mathcal{S}}, \mathcal{B}^{\mathcal{S}}))$ for some g-product $\kappa(\mathcal{A}^{\mathcal{S}}, \mathcal{B}^{\mathcal{S}})$.

Moreover, for any $k \geq 1$, and algebras $\mathcal{A}_1, \dots, \mathcal{A}_k$, there is a g-product \mathcal{P} of $\mathcal{A}_1^{\mathcal{S}}, \dots, \mathcal{A}_k^{\mathcal{S}}$ such that $\text{TrS}(\mathcal{A}_1) \times \dots \times \text{TrS}(\mathcal{A}_k) \cong \text{TrS}(\mathcal{P})$.

Proof. The statements (1) and (3) can be proved similarly as their counterparts in Lemmas 3.4, 3.5, and 3.6 just by replacing the identity translation 1_A (and 1_B) with $\mathbf{I}_{\mathcal{A}}$ (with $\mathbf{I}_{\mathcal{B}}$). We prove (2):

For a semigroup S that satisfies $\text{TrS}(\mathcal{A}) \leftarrow S \subseteq \text{TrS}(\mathcal{B})$, let $\Lambda_S = \{\bar{p} \in \Lambda_{\mathcal{B}} \mid p \in S\}$. Then clearly $\mathcal{S} = (S \cup \{\mathbf{I}_{\mathcal{B}}\}, \Lambda_S) \subseteq_g \mathcal{B}^{\mathcal{S}}$ where the interpretation of $\bar{p} \in \Lambda_S$ in \mathcal{S} is defined by $\bar{p}^{\mathcal{S}}(q) = q \cdot p$ and $\bar{p}^{\mathcal{S}}(\mathbf{I}_{\mathcal{B}}) = p$ for $p, q \in S$. Suppose $\varphi : S \rightarrow \text{TrS}(\mathcal{A})$ is a semigroup epimorphism. Define the assignment $\kappa : \Lambda_S \rightarrow \Lambda_{\mathcal{A}}$ by $\bar{q}\kappa = \bar{q}\varphi$ for all $q \in S$. It is clear that κ is surjective and for all $q, r \in S \subseteq \text{TrS}(\mathcal{B})$, $(\bar{q}^{\mathcal{B}^{\mathcal{S}}}(r))\varphi = (r \cdot q)\varphi = r\varphi \cdot q\varphi = \bar{q}\varphi^{\mathcal{A}^{\mathcal{S}}}(r\varphi) = (q\kappa)^{\mathcal{A}^{\mathcal{S}}}(r\varphi)$. Hence $(\kappa, \tilde{\varphi}) : \mathcal{S} \rightarrow \mathcal{A}^{\mathcal{S}}$ defined by $s\tilde{\varphi} = s\varphi$ for $s \in S$ and $\mathbf{I}_{\mathcal{B}}\tilde{\varphi} = \mathbf{I}_{\mathcal{A}}$, is a g-epimorphism. Thus $\mathcal{A}^{\mathcal{S}} \leftarrow_g \mathcal{S} \subseteq_g \mathcal{B}^{\mathcal{S}}$. \square

The following characterization of the class of finite algebras definable by translation semigroups can be proved similarly as Theorem 3.7.

Theorem 5.2 A class of finite algebras \mathbf{K} is definable by translation semigroups iff it is a GVFA such that the equivalence $\mathcal{A} \in \mathbf{K}$ iff $\mathcal{A}^{\mathcal{S}} \in \mathbf{K}$ holds for any finite algebra \mathcal{A} .

5.2 Languages Definable by Syntactic Semigroups

Let X be a leaf alphabet and Σ be a ranked alphabet such that $\Sigma \neq \Sigma_0$. A *trivial tree language* T consists of constant or leaf symbols only, i.e., $T \subseteq$

$\Sigma_0 \cup X$. For such a tree language T , the syntactic semigroup of T is the trivial semigroup consisting of a zero element, while its syntactic monoid consists of a zero element and an identity element. Since the trivial semigroup belongs to every variety of finite semigroups, any family of tree languages definable by syntactic semigroups should contain all these trivial tree languages. So, it is reasonable to consider $+$ -varieties of tree languages (cf. [4] Section 11).

The sets of *non-trivial ΣX -trees* and *non-trivial ΣX -contexts* are defined by $T^+(\Sigma, X) = T(\Sigma, X) \setminus (\Sigma_0 \cup X)$ and $C^+(\Sigma, X) = C(\Sigma, X) \setminus \{\xi\}$, respectively. Any subset of $T^+(\Sigma, X)$ is called a *trivial-free tree language*.

For a trivial-free tree language $T \subseteq T^+(\Sigma, X)$ the syntactic semigroup of T is the quotient semigroup $C^+(\Sigma, X)/\mu_T$ where μ_T is restricted to $C^+(\Sigma, X)$.

A regular tree homomorphism $\varphi : T(\Sigma, X) \rightarrow T(\Omega, Y)$ is called *strict*, if $\varphi_m(f)$ is not trivial for any $f \in \Sigma_m$ with $m > 0$, and $\varphi_X(X), \varphi_0(\Sigma_0) \subseteq Y \cup \Omega_0$ (cf. Definition 11.1 of [4]). We note that if φ is strict and regular, then $T^+(\Sigma, X)\varphi^{-1} = T^+(\Omega, Y)$.

A family of regular trivial-free tree languages $\{\mathcal{V}(\Sigma, X)\} \subseteq \{T^+(\Sigma, X)\}$ is called a $+$ -GVTL if it is closed under Boolean operations, inverse translations and inverse strict regular tree homomorphisms, and moreover satisfies the following conditions:

- (1) For every unary ranked alphabet Λ , and any leaf alphabets X and Y , if $Y \subseteq X$ then $\mathcal{V}(\Lambda, Y) \subseteq \mathcal{V}(\Lambda, X)$.
- (2) For any strict regular tree homomorphism $\varphi : T(\Sigma, X) \rightarrow T(\Omega, Y)$ full with respect to $T \subseteq T^+(\Omega, Y)$, if $T\varphi^{-1} \in \mathcal{V}(\Sigma, X)$ then $T \in \mathcal{V}(\Omega, Y)$.

That any variety of trivial-free tree languages definable by syntactic semigroups is a $+$ -GVTL can be proved analogously to that of the monoid case. We claim the converse in the following theorem.

Theorem 5.3 A family of trivial-free tree languages is definable by syntactic semigroups iff it is a $+$ -GVTL of tree languages.

The proof, once we have proved the following semigroup counterparts of Lemmas 4.10 and 4.11, is very similar to that of Theorem 4.14.

Lemma 5.4 Let $\mathcal{A} = (A, \Sigma)$ be a finite algebra, and X be a leaf alphabet disjoint from $A \cup \Sigma$.

(1) For any trivial-free tree language $L \subseteq T^+(\Lambda_{\mathcal{A}}, X)$ recognized by $\mathcal{A}^{\mathcal{S}}$, there exists a strict regular tree homomorphism $\varphi : T(\Lambda_{\mathcal{A}}, X) \rightarrow T(\Sigma, X \cup A)$, and a trivial-free tree language $T \subseteq T^+(\Sigma, X \cup A)$, such that $L = T\varphi^{-1}$, and T can be recognized by a finite power of \mathcal{A} .

(2) For any trivial-free tree language $T \subseteq T^+(\Sigma, X)$ recognized by \mathcal{A} there exists a unary ranked alphabet Λ and a strict regular tree homomorphism $\varphi : T(\Lambda, X \cup \Sigma_0) \rightarrow T(\Sigma, X)$ such that φ is full with respect to T , and for every $z \in X \cup \Sigma_0$, $T\varphi^{-1} \cap T(\Lambda, \{z\})$ can be recognized by $\mathcal{A}^{\mathcal{S}}$ as a subset of $T(\Lambda, \{z\})$.

Proof. (1) Suppose for an initial assignment $\alpha : X \rightarrow \text{Tr}(\mathcal{A}) \cup \{\mathbf{I}_{\mathcal{A}}\}$ and a subset $F \subseteq \text{Tr}(\mathcal{A}) \cup \{\mathbf{I}_{\mathcal{A}}\}$, $L = \{t \in \text{T}(\Lambda_{\mathcal{A}}, X) \mid t\alpha^{A^e} \in F\}$ holds. Since L is trivial-free, we can assume that $\mathbf{I}_{\mathcal{A}} \notin F$, or equivalently $F \subseteq \text{Tr}(\mathcal{A})$. Let $Y = \{x \in X \mid x\alpha = \mathbf{I}_{\mathcal{A}}\}$. Define the tree homomorphism $\varphi : \text{T}(\Lambda_{\mathcal{A}}, X) \rightarrow \text{T}(\Sigma, X \cup A)$ by $\varphi_X(x) = x$ for all $x \in X$, and for every $p \in \text{Tr}(\mathcal{A})$ choose a $\varphi_1(\bar{p}) \in \text{C}(\Sigma, A)$ such that $\varphi_1(\bar{p})^A = p$. Obviously φ is a strict regular tree homomorphism. Suppose that $A = \{a_1, \dots, a_m\}$. Let $F' = \{(p(a_1), \dots, p(a_m)) \in A^m \mid p \in F\}$, and define the initial assignment $\beta : X \cup A \rightarrow A^m$ by $x\beta = ((x\alpha)(a_1), \dots, (x\alpha)(a_m))$ for all $x \in X \setminus Y$, $y\beta = (a_1, \dots, a_m)$ for all $y \in Y$, and $a\beta = (a, \dots, a) \in A^m$ for all $a \in A$. Let T be the subset of $\text{T}(\Sigma, X \cup A)$ recognized by $(\mathcal{A}^m, \beta, F')$. We show $L = T\varphi^{-1}$. Every trivial-free tree w in $\text{T}^+(\Lambda_{\mathcal{A}}, X)$ is of the form $w = \bar{p}_1(\bar{p}_2(\dots \bar{p}_k(x)\dots))$ for some $p_1, \dots, p_k \in \text{Tr}(\mathcal{A})$ ($k > 0$) and $x \in X$. For such a tree w , $w\alpha^{A^e} = x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1$ if $x \in X \setminus Y$, and $w\alpha^{A^e} = p_k \cdot \dots \cdot p_2 \cdot p_1$ if $x \in Y$; also $(w\varphi)\beta^{A^m} = (x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1(a_1), \dots, x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1(a_m))$ holds. So, for $x \in X \setminus Y$ we have $w\varphi \in T$ iff $(w\varphi)\beta^{A^m} \in F'$ iff “for some $p \in F$, $p(a) = x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1(a)$, for all $a \in A$ ” iff $x\alpha \cdot p_k \cdot \dots \cdot p_2 \cdot p_1 \in F$ iff $w\alpha^{A^e} \in F$ iff $w \in L$. Similarly, for $x \in Y$ we have $w\varphi \in T$ iff $(w\varphi)\beta^{A^m} \in F'$ iff “for some $p \in F$, $p(a) = p_k \cdot \dots \cdot p_2 \cdot p_1(a)$, for all $a \in A$ ” iff $p_k \cdot \dots \cdot p_2 \cdot p_1 \in F$ iff $w\alpha^{A^e} \in F$ iff $w \in L$.

(2) The proof is almost identical to that of Lemma 4.11, only 1_A is replaced with $\mathbf{I}_{\mathcal{A}}$. □

It was shown in Example 4.3 that the variety of 1-definite tree languages is not definable by syntactic monoids. In the following example we show that its trivial-free counterpart is not definable by syntactic semigroups.

Example 5.5 The syntactic semigroup of the trivial-free 1-definite tree language $T \setminus \{x\}$ where T is defined in Example 4.3, consists of two elements both of which are right zeros. Let $\Lambda = \Lambda_1 = \{\alpha, \beta\}$ and $X = \{x, y\}$. Let T'' be the set of all ΛX -trees which either have root label α and leaf label x or have root label β and leaf label y , i.e., $T'' = \{\alpha(p(x)) \mid p \in \text{C}(\Lambda, X)\} \cup \{\beta(p(y)) \mid p \in \text{C}(\Lambda, X)\}$. It is easy to see that the syntactic semigroup of T'' consists of two right zero elements, but clearly T'' is not 1-definite. So, the trivial-free 1-definite tree languages are not definable by syntactic semigroups.

Indeed, T'' is not k -definite for any $k \geq 1$, thus the trivial-free definite tree languages are not definable by syntactic semigroups.

5.3 Monoids vs. Semigroups

In this subsection we show that the concepts of “definability by semigroups” and “definability by monoids” are not comparable to each other.

The abbreviation VFS stands for variety of finite semigroups. For a VFS \mathbf{S} , let \mathbf{S}^a be the class of all finite algebras whose translation semigroups are

in \mathbf{S} , and \mathbf{S}^t be the family of all recognizable trivial-free tree languages whose syntactic semigroups are in \mathbf{S} (cf. Definitions 3.2 and 4.2).

We recall Proposition 10.9 of [18] which can be extended to VFS's.

Theorem 5.6 For any VFM \mathbf{M} and VFS \mathbf{S} , the identities $\mathbf{M}^{\text{at}} = \mathbf{M}^t$, $\mathbf{M}^{\text{ta}} = \mathbf{M}^a$, $\mathbf{S}^{\text{at}} = \mathbf{S}^t$ and $\mathbf{S}^{\text{ta}} = \mathbf{S}^a$ hold.

Theorem 5.7 (1) There is a VFM \mathbf{M} for which no VFS \mathbf{S} , satisfies $\mathbf{M}^a = \mathbf{S}^a$ or $\mathbf{M}^t = \mathbf{S}^t$.

(2) There is a VFS \mathbf{S} such that for no VFM \mathbf{M} , $\mathbf{M}^a = \mathbf{S}^a$ or $\mathbf{M}^t = \mathbf{S}^t$ holds.

Proof. (1) Let \mathbf{M} be the class of all finite monoids which satisfy the equation $y \cdot x \cdot x = y$. Obviously, \mathbf{M} is a VFM. Let $\Sigma = \Sigma_1 = \{f\}$ and put the algebras $\mathcal{A} = (A, \Sigma)$ and $\mathcal{B} = (B, \Sigma)$ be defined by $A = \{a\}$, $f^{\mathcal{A}}(a) = a$, and $B = \{a, b\}$, $f^{\mathcal{B}}(a) = f^{\mathcal{B}} = a$. Then $\text{Tr}(\mathcal{A}) \cong \text{TrS}(\mathcal{A}) \cong \text{TrS}(\mathcal{B})$ is the trivial semigroup, but the monoid $\text{Tr}(\mathcal{B})$ consists of a zero element (0) and a unit (1). Now, $\mathcal{A} \in \mathbf{M}^a$, but $\mathcal{B} \notin \mathbf{M}^a$ since $\text{Tr}(\mathcal{B})$ does not satisfy the equation $y \cdot x \cdot x = y$: $1 \cdot 0 \cdot 0 = 0 \neq 1$. Hence, \mathbf{M}^a is not definable by translation semigroups. Now if $\mathbf{M}^t = \mathbf{S}^t$ hold for a VFS \mathbf{S} , then by Theorem 5.6 we would have $\mathbf{M}^a = \mathbf{M}^{\text{ta}} = \mathbf{S}^{\text{ta}} = \mathbf{S}^a$, contradiction.

(2) Let \mathbf{S} be the variety of finite right zero semigroups, i.e., the class of all semigroups that satisfy the equation $y \cdot x = x$. It can be easily seen that if T and T' are the tree languages of Example 4.3, then $T \setminus \{x\} \in \mathbf{S}^t(\Sigma, X)$ since the syntactic semigroup of $T \setminus \{x\}$ has two elements both of which are right zeros. On the other hand, the syntactic semigroup of T' consists of an identity element and two right zeros (like its syntactic monoid). Thus $T' \notin \mathbf{S}^t(\Sigma, X)$. This shows that \mathbf{S}^t is not definable by syntactic monoids (since $T \setminus \{x\}$ and T' have isomorphic syntactic monoids) whence $\mathbf{M}^t = \mathbf{S}^t$ does not hold for any VFM \mathbf{M} . On the other hand, if $\mathbf{M}^a = \mathbf{S}^a$ holds for some VFM \mathbf{M} , then by Theorem 5.6, $\mathbf{M}^t = \mathbf{M}^{\text{at}} = \mathbf{S}^{\text{at}} = \mathbf{S}^t$, contradiction. \square

Theorems 5.7 justifies the task of studying the definability by semigroup separately from the monoid case.

6 String languages definable by translation monoids

In this final section, we present for strings the results corresponding to those of the previous sections. Familiarity with the basic notions of string languages and automata are presumed.

Let X be a finite alphabet, and X^* be the set of words over X . A string language over X is any subset of X^* . In the literature the syntactic monoid $\text{SM}(L)$ of a string language $L \subseteq X^*$ is defined to be the quotient monoid X^*/θ_L where $w \theta_L w' \iff \forall u, v \in X^* (uvw \in L \leftrightarrow uw'v \in L)$.

For a monoid $\mathcal{M} = (M, \cdot)$ the translations of \mathcal{M} are the unary functions on M defined by $x \mapsto m \cdot x \cdot m'$ for some $m, m' \in M$. Denote the composition of the translations p and q by $p \circ q$, that is $p \circ q(m) = p(q(m))$ for all

$m \in M$. We note that the set of translations of \mathcal{M} is a monoid with respect to composition operation. Denote the translation monoid of \mathcal{M} by $\text{Tr}(\mathcal{M})$. For a string language L , let the *translation monoid* $\text{TM}(L)$ of L be the translation monoid of the syntactic monoid of L , i.e., $\text{TM}(L) = \text{Tr}(\text{SM}(L))$.

Note that by necessity the terms ‘syntactic monoid’ and ‘translation monoid’ have different meanings and interpretations in this section.

Eilenberg’s [3] variety theorem establishes a correspondence between a variety of finite monoids \mathbf{M} and a variety of string languages $\mathcal{L} = \{\mathcal{L}(X)\}$ such that for any $L \subseteq X^*$, $L \in \mathcal{L}(X) \iff \text{SM}(L) \in \mathbf{M}$.

A variety of string languages $\mathcal{V} = \{\mathcal{V}(X)\}$ is definable by translation monoids if there exists a variety of finite monoids \mathbf{M} such that for any $L \subseteq X^*$, $L \in \mathcal{V}(X) \iff \text{TM}(L) \in \mathbf{M}$. We shall characterize these varieties of string languages in Theorem 6.6 below.

It is known that not any variety of string languages can be defined by translation monoids (one example is the class of reverse definite, or frontier testable, string languages, cf. [21]).

For a monoid $\mathcal{M} = (M, \cdot)$, the *reverse* of \mathcal{M} is the monoid $\mathcal{M}^{\text{R}} = (M, \cdot_{\text{R}})$ where $m \cdot_{\text{R}} m' = m' \cdot m$ for $m, m' \in M$. Clearly $(\mathcal{M}^{\text{R}})^{\text{R}} = \mathcal{M}$. We show that a variety of finite monoids is definable by translation monoids (see Definition 3.2) iff it is closed under the reversing operation.

First, we show that the reverse of a monoid and the original monoid have isomorphic translation monoids.

Lemma 6.1 For a monoid $\mathcal{M} = (M, \cdot)$, $\text{Tr}(\mathcal{M}) \cong \text{Tr}(\mathcal{M}^{\text{R}})$.

Proof. For any translation $p(x) = m \cdot x \cdot m'$ ($m, m' \in M$) of \mathcal{M} let $p^{\text{R}}(x) = m' \cdot x \cdot m$. The mapping $\text{Tr}(\mathcal{M}) \rightarrow \text{Tr}(\mathcal{M}^{\text{R}})$, $p \mapsto p^{\text{R}}$ is an isomorphism. \square

Next, we present some connections between the translation monoid of a monoid and the original monoid.

Lemma 6.2 For any monoid \mathcal{M} , (1) $\mathcal{M} \subseteq \text{Tr}(\mathcal{M})$, and (2) $\text{Tr}(\mathcal{M}) \leftarrow \mathcal{M} \times \mathcal{M}^{\text{R}}$.

Proof. Let $\mathcal{M} = (M, \cdot)$. (1) For any $m \in M$, let p_m be the translation defined by $p_m(x) = m \cdot x$ on M . It is easy to see that the mapping $m \mapsto p_m$ is a monoid monomorphism that embeds \mathcal{M} into $\text{Tr}(\mathcal{M})$. (2) For any $m, n \in M$, let $q_{(m,n)}$ be the translation of \mathcal{M} defined by $q_{(m,n)}(x) = m \cdot x \cdot n$. It can be easily seen that $(m, n) \mapsto m \cdot x \cdot n$ yields an epimorphism $\mathcal{M} \times \mathcal{M}^{\text{R}} \rightarrow \text{Tr}(\mathcal{M})$. \square

Finally, we characterize the varieties of finite monoids definable by translation monoids.

Theorem 6.3 A variety of finite monoids \mathbf{M} is definable by translation monoids iff it is closed under the reversing operation, i.e., $\mathcal{M} \in \mathbf{M} \Rightarrow \mathcal{M}^{\text{R}} \in \mathbf{M}$ for any monoid \mathcal{M} .

Proof. By Lemma 6.1, every variety of finite monoids definable by translation monoids is closed under the reversing operation. Now suppose a variety of finite monoids \mathbf{M} is closed under the reversing operation. We show that $\mathcal{M} \in \mathbf{M} \iff \text{Tr}(\mathcal{M}) \in \mathbf{M}$ for any monoid \mathcal{M} . The implication $\text{Tr}(\mathcal{M}) \in \mathbf{M} \Rightarrow \mathcal{M} \in \mathbf{M}$ follows from Lemma 6.2(1). For the converse, let $\mathcal{M} \in \mathbf{M}$. Then also $\mathcal{M}^R \in \mathbf{M}$, and hence $\text{Tr}(\mathcal{M}) \in \mathbf{M}$ by Lemma 6.2(2). \square

The proof also implies that:

Corollary 6.4 If a variety of finite monoids \mathbf{M} is definable by translation monoids, then \mathbf{M} is generated by the translation monoids of its members.

In the sequel we characterize the varieties of string languages definable by translation monoids.

For a string $w = x_1x_2 \dots x_n \in X^*$ define the reverse of w as $w^R = x_n \dots x_2x_1$. We note that $u^Rv^R = (vu)^R$ holds for all $u, v \in X^*$. For a string language $L \subseteq X^*$, $L^R = \{w^R \in X^* \mid w \in L\}$.

The following lemma is a known fact (see e.g. [3]).

Lemma 6.5 For any string language $L \subseteq X^*$, $\text{SM}(L^R) \cong \text{SM}(L)^R$.

Our characterization of the varieties of string languages definable by translation monoids is the following.

Theorem 6.6 A class of string languages \mathcal{V} is definable by translation monoids iff it is a variety of string languages closed under the reversing operation, i.e., $L \in \mathcal{V}(X) \Rightarrow L^R \in \mathcal{V}(X)$ for any string language $L \subseteq X^*$.

Proof. Since Lemmas 6.5 and 6.1 imply that $\text{TM}(L) \cong \text{TM}(L^R)$ for any string language L , any variety of string languages definable by translation monoids is closed under the reversing operation. Now, suppose \mathcal{V} is a variety of string languages closed under the reversing operation. By Eilenberger's variety theorem there is a variety of finite monoids \mathbf{M} such that for any string language $L \subseteq X^*$, $L \in \mathcal{V}(X) \Leftrightarrow \text{SM}(L) \in \mathbf{M}$. We show that the class \mathbf{M} also defines the translation monoids of \mathcal{V} , that is to say, for any $L \subseteq X^*$, $L \in \mathcal{V}(X) \Leftrightarrow \text{TM}(L) \in \mathbf{M}$. First, suppose L is in $\mathcal{V}(X)$. Then also $L^R \in \mathcal{V}(X)$, so $\text{SM}(L) \in \mathbf{M}$ and $\text{SM}(L^R) \in \mathbf{M}$. By Lemma 6.5, $\text{SM}(L)^R \in \mathbf{M}$, and since $\text{TM}(L)$ is an epimorphic image of $\text{SM}(L) \times \text{SM}(L)^R$ by Lemma 6.2, $\text{TM}(L) \in \mathbf{M}$. Next, suppose $\text{TM}(L) \in \mathbf{M}$ for a string language $L \subseteq X^*$. Since by Lemma 6.2, $\text{SM}(L)$ is isomorphic to a submonoid of $\text{TM}(L)$, then $\text{SM}(L) \in \mathbf{M}$, and hence $L \in \mathcal{V}(X)$. \square

Corollary 6.7 Let \mathcal{V} be a variety of string languages definable by translation monoids. Then the variety generated by the translation monoids of \mathcal{V} is equal to the variety generated by the syntactic monoids of \mathcal{V} .

An analogue of Theorem 6.6 can be proved for translation semigroups.

Unlike Theorems 4.14 and 5.3 for tree languages, by Theorem 6.6 checking whether or not a variety of string languages is definable by translation monoids or semigroups is rather easy. For example the variety of definite string languages and the variety of reverse definite string languages are not definable by translation semigroups, while the variety of aperiodic string languages and the variety of commutative string languages (i.e., having commutative syntactic monoids) are definable by translation monoids.

Acknowledgements

I would like to thank Magnus Steinby, Tatjana Petković, Ville Piirainen, and Zoltán Ésik for helpful comments and stimulating discussions.

References

- [1] Almeida J., On pseudovarieties, varieties of languages, filters of congruences, pseudoidentities and related topics, *Algebra Universalis* **27** (1990) 333–350.
- [2] Birkhoff G., On the structure of abstract algebras, *Proc. Cambridge Phil. Soc.* **31** (1935) 433–454.
- [3] Eilenberg S., *Automata, Languages, and Machines*, Vol. **B**. Pure and Applied Mathematics, Vol. 59, Academic Press, New York – London (1976).
- [4] Ésik Z., A variety theorem for trees and theories, Automata and formal languages VIII (Salgótarján, 1996), *Publ. Math. Debrecen* **54** (1999), suppl., 711–762.
- [5] Ésik Z. & Weil P., On Logically Defined Recognizable Tree Languages, *Proceedings of FSTTCS'03*, Lect. Notes in Comp. Sci. **2914**, Springer-Verlag (2003), 195–207.
- [6] Nivat M. & Podelski A., Tree monoids and recognizability of sets of finite trees, *Resolution of Equations in Algebraic Structures*, Vol. **1**, Academic Press, Boston MA (1989) 351–367.
- [7] Nivat M. & Podelski A., Definite tree languages (cont'd), *Bull. EATCS* **38** (1989) 186–190.
- [8] Petković T., Ćirić M. & Bogdanović S., Unary algebras, semigroups and congruences on free semigroups, *Theoret. Comput. Sci.* **324** (2004) 87–105.
- [9] Petković T., Ćirić M. & Bogdanović S., Eilenberg type theorems for automata, submitted.

- [10] Pin J.E., *Varieties of formal languages*, Foundations of Computer Science, Plenum Publishing Corp., New York (1986).
- [11] Pin J.E., A variety theorem without complementation, *Izvestiya VUZ Matematika* **39** (1995) 80–90.
- [12] Piirainen V., Monotone algebras, R -trivial monoids and a variety of tree languages, *Bull. EATCS* **84** (2004) 189–194.
- [13] Podelski A., A monoid approach to tree languages, in: Nivat M. & Podelski A. (ed.) *Tree Automata and Languages*, Elsevier-Amsterdam (1992) 41–56.
- [14] Salomaa K., *Syntactic monoids of regular forests* (Finnish), M.Sc. Thesis, Department of Mathematics, Turku University (1983).
- [15] Schützenberger M. P., On finite monoids having only trivial subgroups, *Information and Control* **8** (1965) 190–194.
- [16] Steinby M., Syntactic algebras and varieties of recognizable sets, in: *Proc. CAAP’79* (University of Lille 1979) 226–240.
- [17] Steinby M., A theory of tree language varieties, in: Nivat M. & Podelski A. (ed.) *Tree Automata and Languages*, Elsevier-Amsterdam (1992) 57–81.
- [18] Steinby M., General varieties of tree languages, *Theoret. Comput. Sci.* **205** (1998) 1–43.
- [19] Thérien D., Recognizable languages and congruences, *Semigroup Forum* **23** (1981) 371–373.
- [20] Thomas W., Logical aspects in the study of tree languages, *Ninth Colloquium on Trees in Algebra and in Programming* (Proc. CAAP’84), Cambridge University Press (1984) 31–51.
- [21] Wilke T., An algebraic characterization of frontier testable tree languages, *Theoret. Comput. Sci.* **154** (1996) 85–106.

TURKU
CENTRE *for*
COMPUTER
SCIENCE

Lemminkäisenkatu 14 A, 20520 Turku, Finland | www.tucs.fi



University of Turku

- Department of Information Technology
- Department of Mathematics



Åbo Akademi University

- Department of Computer Science
- Institute for Advanced Management Systems Research



Turku School of Economics and Business Administration

- Institute of Information Systems Sciences

ISBN 952-12-1394-9

ISSN 1239-1891