EF+EX Forest Algebras

Andreas Krebs[1(✉)] and Howard Straubing[2]

[1] Wilhelm-Schickard-Institut at Eberhard-Karls, Universität Tübingen, Tübingen, Germany
krebs@informatik.uni-tuebingen.de
[2] Boston College, Chestnut Hill, USA

Abstract. We examine languages of unranked forests definable using the temporal operators EF and EX. We characterize the languages definable in this logic, and various fragments thereof, using the syntactic forest algebras introduced by Bojanczyk and Walukiewicz. Our algebraic characterizations yield efficient algorithms for deciding when a given language of forests is definable in this logic. The proofs are based on understanding the wreath product closures of a few small algebras, for which we introduce a general ideal theory for forest algebras. This combines ideas from the work of Bojanczyk and Walukiewicz for the analogous logics on binary trees and from early work of Stiffler on wreath product of finite semigroups.

1 Overview

Understanding the expressive power of temporal and first-order logic on trees is important in several areas of computer science, for example in formal verification. Using algebraic methods, in particular, finite monoids, to understand the power of subclasses of the regular languages of finite words has proven to be extremely successful, especially in the characterization of regular languages definable in various fragments of first-order and temporal logics ([CPP93,TW96,Str94]). Here we are interested in sets of of finite trees (or, more precisely, sets of finite forests), where the analogous algebraic structures are forest algebras.

Bojanczyk *et. al.* [BW08,BSW12] introduced forest algebras, and underscored the importance of the wreath product decomposition theory of these algebras in the study of the expressive power of temporal and first-order logics on finite unranked trees. For languages inside of CTL the associated forest algebras can be built completely via the wreath product of copies of the forest algebra $\mathcal{U}_2 = (\{0,\infty\},\{1,0,c_0\})$, where the vertical element 0 is the constant map to ∞, and the vertical element c_0 is the constant map to 0 ([BSW12]). The problem of effectively characterizing the wreath product closure of $\mathcal{U}_2$ is thus an important open problem, equivalent to characterization of CTL. Note that if one strips away the additive structure of $\mathcal{U}_2$, the wreath product closure is the family of all finite aperiodic semigroups (the Krohn-Rhodes Theorem). Forest algebras have been successfully applied to the obtain characterization of other logics on trees; see, for example [BSS12,BS09].

A. Maletti (Ed.): CAI 2015, LNCS 9270, pp. 128–139, 2015.
DOI: 10.1007/978-3-319-23021-4_12

Here we study in detail the wreath product closures of proper subalgebras of $\mathcal{U}_2$. In one sense, this generalizes early work of Stiffler [Sti73], who carried out an analogous program for wreath products of semigroups. Along the way, we develop the outlines of a general ideal theory for forest algebras, which we believe will be useful in subsequent work. After developing the underlying algebraic theory, we give an application to logic, obtaining a characterization of the languages of unranked forests definable with the temporal operators EF and EX.

Bojanczyk and Walukiewicz [BW06] obtained similar results for binary trees, using methods quite different from ours. Esik [Ési05] considered the analogous logics for ranked trees, and proved similar decidability results with techniques very much in the same spirit as ours, relying on a version of the wreath product tree automata acting on ranked trees.

Much of our goal in presenting these results in the context of unranked forest algebras is to develop the outlines of a general ideal theory for these algebras, and to show its connection with wreath product decompositions. We believe this approach will prove useful in subsequent work.

2 Forest Algebras

2.1 Preliminaries

We refer the reader to [BW08,BSW12] for the definitions of abstract forest algebra, free forest algebra, and syntactic forest algebra. We denote the free forest algebra over a finite alphabet A by $A^{\Delta} = (H_A, V_A)$, where H_A denotes the monoid of forests over A, with concatenation as the operation, and V_A denotes the monoid of contexts over A, with composition as the operation. A subset L of H_A is called a forest language over A. We denote its syntactic forest algebra by (H_L, V_L), and its syntactic morphism by $\mu_L : A^{\Delta} \to (H_L, V_L)$.

For the most part, our principal objects of study are not the forest algebras themselves, but homomorphisms $\alpha : A^{\Delta} \to (H, V)$. It is important to bear in mind that each such homomorphism is actually a pair of monoid homomorphisms, one mapping H_A to H and the other mapping V_A to V. It should usually be clear from the context which of the two component homomorphisms we mean, and thus we denote them both by α. The 'freeness' of A^{Δ} is the fact that a homomorphism α into (H, V) is completely determined by giving its value, in V, at each $a \in A$.

A homomorphism α as above *recognizes* a language $L \subseteq H_A$ if there exists $X \subseteq H$ such that $\alpha^{-1}(X) = L$.

If $\alpha : A^{\Delta} \to (H, V)$ and $\beta : A^{\Delta} \to (H', V')$, are homomorphisms, we say that β *factors through* α if for all $s, s' \in H_A$, $\alpha(s) = \alpha(s')$ implies $\beta(s) = \beta(s')$. This is equivalent to the existence of a homomorphism ρ from the image of α into (H', V') such that $\beta = \rho\alpha$. A homomorphism α recognizes $L \subseteq H_A$ if and only if μ_L factors through α.([BW08]).

In the course of the paper we will see several *congruences* defined on free forest algebras. Such a congruence is determined by an equivalence relation $\sim$ on H_A

such that for any $p \in V_A$, $s \sim s'$ implies $ps \sim ps'$. This gives a well-defined action of V_A on the set of $\sim$-classes of H_A. We define an equivalence relation (also denoted $\sim$) on V_A by setting $p \sim p'$ if for all $s \in H_A$, $ps \sim p's$. The result is a quotient forest algebra $(H_A/{\sim}, V_A/{\sim})$. In order to prove that an equivalence relation $\sim$ on H_A is a congruence, it is sufficient to verify that $s \sim s'$ and $t \sim t'$ implies $s + t \sim s' + t'$ and $as \sim as'$ for all $s, s', t, t' \in H_A$ and $a \in A$.

2.2 Horizontally Idempotent and Commutative Algebras

We now introduce an important restriction. Throughout the rest of the paper, we will assume that all of our finite forest algebras (H, V) have H idempotent and commutative; that is $h + h' = h' + h$ and $h + h = h$ for all $h, h' \in H$. This is a natural restriction when talking about classes of forest algebras arising in temporal logics, which is the principal application motivating this study.

When H is horizontally idempotent and commutative, the sum of all its elements is an absorbing element for the monoid. While an absorbing element in a monoid is ordinarily written 0, since we use additive notation for H, its identity is denoted 0, and accordingly we denote the absorbing element, which is necessarily unique, by ∞.

We say that two forests $s_1, s_2 \in H_A$ are *idempotent-and-commutative equivalent* if s can be transformed into t by a sequence of operations of the following three types: *(i)* interchange the order of two adjacent subtrees (that is, if $s = p(t_1+t_2)$ for some context p and trees t_1, t_2, then we transform s to $p(t_2+t_1)$); *(ii)* replace a subtree t by two adjacent copies (that is, transform pt to $p(t+t)$); *(iii)* replace two identical adjacent subtrees by a single copy (transform $p(t+t)$ to pt). Since operations *(ii)* and *(iii)* are inverses of one another, and operation *(i)* is its own inverse, this is indeed an equivalence relation.

We have the following obvious lemma:

Lemma 1. *Let $\alpha : A^\Delta \to (H, V)$ be a homomorphism, where H is horizontally idempotent and commutative. If $s, t \in H_A$ are idempotent-and-commutative equivalent, then $\alpha(s) = \alpha(t)$.*

There is a smallest nontrivial idempotent and commutative forest algebra, $\mathcal{U}_1 = (\{0, \infty\}, \{1, 0\})$. The horizontal and vertical monoids of $\mathcal{U}_1$ are isomorphic, but we use different names for the elements because of the additive notation for the operation in one of these monoids, and multiplicative notation in the other. We have not completely specified how the vertical monoid acts on the horizontal monoid—this is done by setting $0 \cdot x = \infty$ for $x \in \{0, \infty\}$.

2.3 1-Definiteness

In Section 5 we will discuss in detail the notion of definiteness in forest algebras; for this preliminary section, we will only need to consider a special case. A forest algebra homomorphism $\alpha : A^\Delta \to (H, V)$ is said to be *1-definite* if for $s \in H_A$, the value of $\alpha(s)$ depends only on the set of labels of the root nodes of s.

We define an equivalence relation $\sim_1$ on H_A by setting $s \sim_1 s'$ if and only if the sets of labels of root nodes of s and s' are equal. This defines a congruence on A^Δ. We denote the homomorphism from A^Δ onto the quotient under $\sim_1$ by $\alpha_{A,1}$. It is easy to show that a homomorphism $\alpha : A^\Delta \to (H,V)$ is 1-definite if and only if it factors through $\alpha_{A,1}$.

2.4 Wreath Products

We summarize the discussion of wreath products given in [BSW12]. The *wreath product* of two forest algebras $(H_1,V_1),(H_2,V_2)$ is $(H_1,V_1)\circ(H_2,V_2) = (H_1 \times H_2, V_1 \times V_2^{H_1})$, where the monoid structure of $H_1 \times H_2$ is the ordinary direct product, and the action is given by $(v_1,f)(h_1,h_2) = (v_1h_1, f(h_1)h_2)$, for all $h_1 \in H_1$, $h_2 \in H_2$, $v_1 \in V_1$, and $f : H_1 \to V_2$. It is straightforward to verify that the resulting structure satisfies the axioms for a forest algebra. Note that if one forgets about the monoid structure on H_1 and H_2, this is just the ordinary wreath product of left transformation monoids. Because we use left actions rather than the right actions that are traditional in the study of monoid decompositions, we reverse the usual order of the factors. The projection maps $\pi : (h_1,h_2) \mapsto h_1, (v,f) \mapsto v$, define a homomorphism from the wreath product onto the left-hand factor.

2.5 Reachability

Let (H,V) be a finite forest algebra. For $h,h' \in H$ we write $h \leq h'$ if $h = vh'$ for some $v \in V$, and say that h is *reachable* from h'. This gives a preorder on H. We set $h \cong h'$ if both $h \leq h'$ and $h' \leq h$. An equivalence class of $\cong$ is called a *reachability class*. The preorder consequently results in a partial order on the set of reachability classes of H. We always have $h+h' \leq h$, because $h+h' = (1+h')h$. If $h \in H$ and Γ is a reachability class of H then we write, for example, $h \geq \Gamma$ to mean that $\Gamma \leq \Gamma'$, where Γ' is the class of h.

A *reachability ideal* in (H,V) is a subset I of H such that $h \in I$ and $h' \leq h$ implies $h' \in I$. If we have a homomorphism $\alpha : A^\Delta \to (H,V)$ and a reachability ideal $I \subseteq H$, we define an equivalence relation $\sim_I$ on H_A by setting $s \sim_I s'$ if $\alpha(s) = \alpha(s') \notin I$, or if $\alpha(s), \alpha(s') \in I$. Easily $s \sim_I s'$ implies $ps \sim_I ps'$ for any $p \in V_A$. We thus obtain a homomorphism onto the quotient algebra $\alpha_I : A^\Delta \to (H/{\sim_I}, V/{\sim_I})$ which factors through α. Note that I is, in particular, a two-sided ideal in the monoid H, and $H/{\sim_I}$ is identical to the usual quotient monoid $H/I = (H-I)\cup\{\infty\}$. We will thus use the notation $(H/I, V/I)$ for the quotient algebra, instead of $(H/{\sim_I}, V/{\sim_I})$. If $\Gamma \subseteq H$ is a reachability class, then both $I_\Gamma = \{h \in H : h \not> \Gamma\}$ and $I_{\geq\Gamma} = \{h \in H : h \not\geq \Gamma\}$ are reachability ideals. We denote the associated quotients and projection homomorphisms by (H_Γ, V_Γ), α_Γ, $(H_{\geq\Gamma}, V_{\geq\Gamma})$, $\alpha_{\geq\Gamma}$.

Given the restriction that H is idempotent and commutative, the absorbing element ∞ is reachable from every element. The reachability class of ∞ is accordingly the unique minimal class, which we denote Γ_{min}. A reachability class

Γ is *subminimal* if $\Gamma_{\mathsf{min}} < \Gamma$, but there is no class Λ with $\Gamma_{\mathsf{min}} < \Lambda < \Gamma$. The following lemma will be used several times.

Lemma 2. *Let $\alpha : A^{\Delta} \to (H, V)$, and let $\Gamma_1, \ldots, \Gamma_r$ be the subminimal reachability classes of (H, V). Then*

$$\alpha_{\Gamma_{\mathsf{min}}} : A^{\Delta} \to (H_{\Gamma_{\mathsf{min}}}, V_{\Gamma_{\mathsf{min}}})$$

factors through the direct product

$$\left(\prod_{j=1}^{r} \alpha_{\geq \Gamma_j}\right) : A^{\Delta} \to \prod_{j=1}^{r} (H_{\geq \Gamma_j}, V_{\geq \Gamma_j}).$$

Further each of the algebras $(H_{\geq \Gamma_j}, V_{\geq \Gamma_j})$ has a unique subminimal reachability class.

We will also need the following lemma, which concerns the behavior of reachability classes under homomorphisms.

Lemma 3. *Let $\beta : (H_1, V_1) \to (H_2, V_2)$ be a homomorphism of finite forest algebras. Let $\Lambda \subseteq H_1$ be a reachability class. There is a reachability class Γ of (H_2, V_2) such that $\beta(\Lambda) \subseteq \Gamma$. If Λ is a minimal class of (H_1, V_1) satisfying $\beta(\Lambda) \subseteq \Gamma$, and β is onto, then $\beta(\Lambda) = \Gamma$. If, further, H_2 is idempotent and commutative, then there is only one such minimal class Λ.*

3 Connections to Logic

For the definition of temporal logic and especially the temporal operators EF and EX we refer to [BSW12] as our approach closely follows the one given there.

Intuitively, when we interpret formulas in trees, $\mathsf{EF}\phi$ means 'at some time in the future ϕ' and $\mathsf{EX}\phi$ means 'at some next time ϕ'. When we interpret such formulas in forests, we are in a sense treating the forest as though it were a tree with a phantom root node. Observe that if $a \in A$, we do not interpret the formula a in forests at all. Thus a formula can have different interpretations depending on whether we view it as a tree or a forest formula. For example, as a forest formula $\mathsf{EX}a$ means 'there is a root node labeled a' while as a tree formula it means 'some child of the root is labeled a'. If ϕ is a forest formula, then we denote by L_ϕ the set of all $s \in H_A$ such that $s \models \phi$. L_ϕ is the *language defined by ϕ*.

Example 4. Consider the following property of forests over $\{a, b\}$: There is a tree component containing only as, and another tree component that contains at least one b. Now consider the set L of forests s that either have this property, or in which for some node x, the forest of strict descendants of x has the property. The property itself is defined by the forest formula

$$\psi : \mathsf{EX}(a \wedge \neg\mathsf{EF}b) \wedge \mathsf{EX}(b \vee \mathsf{EF}b)$$

and L is defined by $\psi \vee \mathsf{EF}\psi$. In Example 9, we discuss the syntactic forest algebra of L.

3.1 Correspondence of Operators with Wreath Products

The principal result of this paper is the algebraic characterization of the forest languages using the operators EF and EX, either separately or in combination. It will require some algebraic preparation, in Sections 4, 5 and 6 before we can give the precise statement of this theorem. The bridge between the logic and the algebra is provided by the next two propositions.

Let ϕ be a tree formula. Then ϕ can be written as a disjunction $\bigvee_{a \in A}(a \wedge \psi_a)$, where each ψ_a is a forest formula. Let $\Psi = \{\psi_a : a \in A\}$. We'll call Ψ the set of forest formulas of ϕ. We say that a homomorphism $\beta : A^\Delta \to (H, V)$ recognizes Ψ if the value of $\beta(s)$ determines exactly which formulas of Ψ are satisfied by s. To construct such a homomorphism, we can take the direct product of the syntactic algebras of L_ψ for $\psi \in \Psi$, and set β to be the product of the syntactic morphisms.

The following theorem, adapted from [BSW12], gives the connection between the EF operator and wreath products with $\mathcal{U}_1$:

Proposition 5. *(a) Suppose that ϕ is a tree formula, Ψ is the set of forest formulas of ϕ, and that Ψ is recognized by $\alpha : A^\Delta \to (H, V)$. Then $\mathsf{EF}\phi$ is recognized by a homomorphism $\beta : A^\Delta \to (H, V) \circ \mathcal{U}_1$, where $\pi\beta = \alpha$.*
(b) Suppose that $L \subseteq H_A$ is recognized by a homomorphism $\beta : A^\Delta \to (H, V) \circ \mathcal{U}_1$. Then L is a boolean combination of languages of the form $\mathsf{EF}(a \wedge \phi)$, where L_ϕ is recognized by $\pi\beta$.

Here we prove an analogous result for the temporal operator EX.

Proposition 6. *(a) Suppose that ϕ is a tree formula, Ψ is the set of forest formulas of ϕ, and that Ψ is recognized by $\alpha : A^\Delta \to (H, V)$. Then $\mathsf{EX}\phi$ is recognized by a homomorphism $\alpha \otimes \beta : A^\Delta \to (H, V) \circ (H', V')$, where $\beta : (A \times H)^\Delta \to (H', V')$ is 1-definite.*
(b) Suppose that $L \subseteq H_A$ is recognized by a homomorphism $\alpha \otimes \beta : A^\Delta \to (H, V) \circ (H', V')$, Suppose further that every language recognized by α is defined by a formula in some set Ψ of formulas. If $\beta : (A \times H)^\Delta \to (H', V')$ is 1-definite, then L is a boolean combination of languages of the form L_ψ and $\mathsf{EX}(a \wedge \psi)$, where $\psi \in \Psi$.

4 EF-algebras

Following [BSW12], we define:

Definition 7. *A finite forest algebra (H, V) is an EF-algebra if it satisfies the identities $h + h' = h' + h, vh + h = vh$ for all $h, h' \in H$ and $v \in V$. The second identity with $v = 1$ gives $h + h = h$. Thus every EF-algebra is horizontally idempotent and commutative.*

The following result is proved in [BSW12], and is the key element in the characterization of languages definable in one of the temporal logics we consider in Section 3. We will give a new proof, as it provides a good first illustration of how we use the reachability ideal theory introduced above in decomposition arguments.

Theorem 8. *Let $\alpha : A^\Delta \to (H, V)$ be a homomorphism onto a forest algebra. (H, V) is an* EF*-algebra if and only if α factors through a homomorphism $\beta : A^\Delta \to \mathcal{U}_1 \circ \cdots \circ \mathcal{U}_1$.*

A classic result of Stiffler [Sti73] shows that a right transformation monoid (Q, M) divides an iterated wreath product of copies of the transformation monoid $U_1 = (\{0,1\}, \{0,1\})$ if and only if M is $\mathcal{R}$-trivial. In terms of transformation monoids this means there is no pair of distinct states $q \neq q' \in Q$ such that $qm = q', q'm' = q$ for some $m, m' \in M$. Since forest algebras are left transformation monoids, the analogous result would suggest that a forest algebra (H, V) divides an iterated wreath product of copies of $\mathcal{U}_1$ if and only if V is $\mathcal{L}$-trivial—that is, if and only if (H, V) has trivial reachability classes. We have already seen that this condition is necessary.

However, the following example shows that it is not sufficient.

Example 9. Figure 1 below defines the syntactic forest algebra of the language L of Example 4. The nodes in the diagram represent the elements of the horizontal monoid, and the arrows give the action of a generating set of letters $A = \{a, b\}$ on the horizontal monoid. The letter transitions, together with the conventions about idempotence and commutativity, and the meaning of 0 and ∞, completely determine the addition and the action. Since $\infty = a + b = a + ba \neq ba = b$, this is not an EF-algebra, but the reachability classes are singletons.

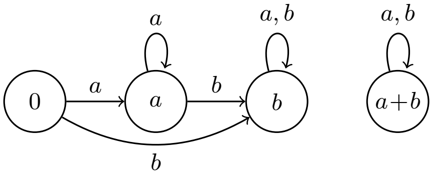

Fig. 1. An algebra with trivial reachability classes that is not an EF-algebra

5 Definiteness

5.1 Definite Homomorphisms

Let $k > 0$. A finite semigroup S is said to be *reverse k-definite* if it satisfies the identity $x_1 x_2 \cdots x_k y = x_1 \cdots x_k$. The reason for the word 'reverse' is that definiteness of semigroups was originally formulated in terms of right transformation monoids, so the natural analogue of definiteness in the setting of forest algebras corresponds to reverse definiteness in semigroups. Observe that the notions of definiteness and reverse definiteness in semigroups do not really make sense for monoids, since only the trivial monoid can satisfy the underlying identities. For much the same reason, we define definiteness for forest algebras not as a property of the algebras themselves, but of homomorphisms $\alpha : A^{\boldsymbol{\Delta}} \to (H, V)$.

The *depth* of a context $p \in V_A$ is defined to be the depth of its hole; so for instance a context with its hole at a root node has depth 0. We say that the homomorphism α is k-definite, where $k > 0$, if for every $p \in V_A$ of depth at least k, and for all $s, s' \in H_A$, $\alpha(ps) = \alpha(ps')$. Easily, if α_1, α_2 are k-definite homomorphisms, then so are $\alpha_1 \times \alpha_2$ and $\psi\alpha_1$, where $\psi : (H, V) \to (H', V')$ is a homomorphism of forest algebras.

A context is *guarded* if it has depth at least 1, that is, if the hole is not at the root. We denote by V_A^{gu} the subsemigroup of V_A consisting of the guarded contexts.

Lemma 10. *Let $k > 0$. A homomorphism $\alpha : A^{\boldsymbol{\Delta}} \to (H, V)$ is k-definite if and only if $\alpha(V_A^{\mathsf{gu}})$ is a reverse k-definite semigroup.*

Definition 11. *An* EX*-homomorphism is a homomorphism that is k-definite for some $k \in \mathbb{N}$.*

5.2 Free *k*-definite Algebra

We construct what we will call *free k-definite algebra* over an alphabet A. This is a slight abuse of terminology, since as we noted above, it is the homomorphism into this algebra, and not the algebra itself, that is k-definite. We do this by recursively defining a sequence of congruences $\sim_k$ on A^{Δ}. If $k = 0$, then $\sim_0$ is just the trivial congruence that identifies all forests. If $k \geq 0$ and $\sim_k$ ha been defined then we associate to each forest $s = a_1 s_1 + \cdots a_r s_r$, where each $a_i \in A$, $s_i \in H_A$, the set

$$T_s^{k+1} = \{(a_i, [s_i]_{\sim_k}) : 1 \leq i \leq r\},$$

where $[]_{\sim_k}$ denotes the $\sim_k$-class of a forest. We then define $s \sim_{k+1} s'$ if and only if $T_s^{k+1} = T_{s'}^{k+1}$.

Proposition 12. *Let $k \geq 0$. Then $\sim_{k+1}$ refines $\sim_k$. $\sim_k$ is a congruence of finite index on A^{Δ}, with a horizontally idempotent and commutative quotient.*

Intuitively, $s \sim_k s'$ means that the forests s and s' are identical at the k levels closest to the root, up to idempotent and commutative equivalence. In fact, this intuition provides an equivalent characterization of $\sim_k$, which we give below. We omit the simple proof.

Lemma 13. *Let $s, s' \in H_A$ and $k > 0$. Let $\bar{s}$, $\bar{s'}$, denote, respectively, the forests obtained from s and s' by removing all the nodes at depth k or more. Then $s \sim_k s'$ if and only if $\bar{s}$ and $\bar{s'}$ are idempotent-and-commutative equivalent.*

Let us denote by $\alpha_{A,k}$ the homomorphism from A^{Δ} onto its quotient by $\sim_k$. In the case where $k = 1$, we will identify $H_A/{\sim_1}$ with the monoid $(\mathcal{P}(A), \cup)$, and the horizontal component of $\alpha_{A,1}$ with the map that sends each forest to the set of its root nodes.

The following theorem gives both the precise sense in which this is the 'free k-definite forest algebra', as well as the wreath product decomposition of k-definite homomorphisms into 1-definite homomorphisms into a forest algebra with horizontal monoid $\{0, \infty\}$.

Theorem 14. *Let $\alpha : A^{\Delta} \rightarrow (H, V)$ be a homomorphism onto a finite forest algebra. Let $k > 0$. The following are equivalent.*

(a) α is k-definite.
(b) α factors through $\alpha_{A,k}$.
(c) α factors through

$$\beta_1 \otimes \cdots \otimes \beta_k : A^{\Delta} \rightarrow (H_1, V_1) \circ \cdots \circ (H_k, V_k),$$

where each $\beta_i : (A \times H_1 \times \cdots \times H_{i-1})^{\Delta} \rightarrow (H_i, V_i)$ is 1-definite.
(d) α factors through an iterated wreath product of k 1-definite homomorphisms into $\mathcal{U}_2$.

6 $(\mathsf{EF}, \mathsf{EX})$-algebras

6.1 The Principal Result

Definition 15. *An $(\mathsf{EF}, \mathsf{EX})$-homomorphism $\alpha : A^{\Delta} \rightarrow (H, V)$ is one that factors through an iterated wreath product*

$$\beta_1 \otimes \cdots \otimes \beta_k,$$

where each β_i either maps into $\mathcal{U}_1$ or is 1-definite. By Theorem 14 we can suppose that each 1-definite β_i maps into $\mathcal{U}_2$.

The principal result of this paper is an effective necessary and sufficient condition for a homomorphism to be a $(\mathsf{EF}, \mathsf{EX})$-homomorphism.

Definition 16. *Suppose $\alpha : A^{\Delta} \rightarrow (H, V)$. Let $s_1, s_2 \in H_A$, $k > 0$, and $\Gamma \subseteq H$ a reachability class for (H, V). We say that s_1, s_2 are (α, k, Γ)-confused, and write $s_1 \equiv_{\alpha,k,\Gamma} s_2$, if*

$$(s_1)^{\alpha_\Gamma} \sim_k (s_2)^{\alpha_\Gamma}, \quad \alpha(s_1), \alpha(s_2) \in \Gamma.$$

Observe that the equivalence relation $\sim_k$ in the first item is over the extended alphabet $A \times H_\Gamma$. It is worth emphasizing what $(s)^{\alpha_\Gamma}$ is when $\alpha(s) \in \Gamma$: We are tagging each node of x of s with the value $\alpha(t) \in H$ if the tree rooted at x is *at* and $\alpha(t) > \Gamma$, but we are tagging the node by ∞–effectively leaving the node untagged–if $\alpha(t) \in \Gamma$. Since $\alpha(s) \in \Gamma$, every node is of one of these two types.

Definition 17. *A homomorphism α is nonconfusing if and only if there exists $k > 0$ such that $\equiv_{\alpha,k,\Gamma}$ is equality for reachability classes Γ.*

In the full version of the paper [KS14] we show that it can be effectively determined if a forest algebra morphism is nonconfusing.

It follows from Proposition 12 that $\equiv_{\alpha,k+1,\Gamma}$ refines $\equiv_{\alpha,k,\Gamma}$, so that if α is nonconfusing with associated parameter k, then it is nonconfusing for all $m > k$.

Our main result is:

Theorem 18. *Let $\alpha : A^{\Delta} \to (H, V)$ be a homomorphism into a finite forest algebra. Then α is a* $(\mathsf{EF}, \mathsf{EX})$ *homomorphism if and only if it is nonconfusing.*

The proof of Theorem 18 will be given in the next two subsections.

Example 19. Consider once again the algebra of Examples 4 and 9 and the associated homomorphism α from $\{a, b\}^{\Delta}$. Since the algebra has trivial reachability classes, α is nonconfusing for all k, so Theorem 18 implies that α is an $(\mathsf{EF}, \mathsf{EX})$-homomorphism. We will see in the course of the proof of the main theorem how the wreath product decomposition is obtained.

Example 20. Consider again the forest algebra $\mathcal{U}_2 = (\{0, \infty\}, \{1, c_\infty, c_0\})$, and the homomorphism α from $\{a, b, c\}^{\Delta}$ onto $\mathcal{U}_2$ that maps a to 1, b to c_0 and c to c_∞. There is a unique reachability class Γ, so for any forest s, s^{α_Γ} is identical to s. Now observe that $a^k b \sim_k a^k c$, but that these are mapped to different elements under α. So by our main theorem, α is not an $(\mathsf{EF}, \mathsf{EX})$-homomorphism.

6.2 Sufficiency of the Condition

We will use the ideal theory developed earlier to prove that every nonconfusing homomorphism factors through a wreath product decomposition of the required kind. The structure of our proof resembles the one given for Theorem 8. Once again, we proceed by induction on $|H|$. The base of the induction is the trivial case $|H| = 1$. Let us suppose that $\alpha : A^{\Delta} \to (H, V)$ is nonconfusing with parameter k, that $|H| > 1$, and that every nonconfusing homomorphism into a forest algebra with a smaller horizontal monoid factors through a wreath product of the required kind.

Let $\Gamma = \Gamma_{\mathsf{min}}$. Suppose first that $|\Gamma| > 1$. We claim that α factors through

$$\beta = \alpha_\Gamma \otimes \alpha_{B,k} : A^{\Delta} \to (H_\Gamma, V_\Gamma) \circ B^{\Delta}/{\sim_k}$$

where $B = A \times H_\Gamma$. Since $|H_\Gamma| < |H|$ and α_Γ is also nonconfusing, the induction hypothesis gives the desired decomposition of α. To establish the claim, let $s \in H_A$. Then

$$\beta(s) = (\alpha_\Gamma(s), [s^{\alpha_\Gamma}]_{\sim_k}).$$

If $s \notin \Gamma$, then the value of the left-hand coordinate determines $\alpha(s)$. If $s \in \Gamma$, then by the nonconfusion condition, the value of the right-hand coordinate determines $\alpha(s)$. Thus α factors through β as required.

So let $|\Gamma| = 1$. Then $\Gamma = \{\infty\}$ and $(H_\Gamma, V_\Gamma) = (H, V)$. Lemma 2 implies that we can suppose (H, V) has a single subminimal reachability class, because each of the component homomorphisms in the direct product is nonconfusing, and the direct product factors through the wreath product.

Thus we have a unique minimal element ∞, and a unique subminimal ideal Γ'. We claim that α factors through

$$\beta = \alpha_1 \otimes \alpha_2 \otimes \alpha_3 : A^{\Delta} \to (H_{\Gamma'}, V_{\Gamma'}) \circ B^{\Delta}/{\sim_k} \circ \mathcal{U}_1,$$

where $\alpha_1 = \alpha_{\Gamma'}$ and $\alpha_2 = \alpha_{B,k}$, where $B = A \times H_{\Gamma'}$, and $\alpha_3 : (B \times 2^B)^{\Delta} \to \mathcal{U}_1$ will be defined below. To see how α_3 should be defined, let us consider what this homomorphism needs to tell us. If $\alpha(s) > \Gamma'$, then the first coordinate of $\beta(s)$ determines $\alpha(s)$. If $\alpha(s) \in \Gamma'$, then the first two components of $\beta(s)$ determine $\alpha(s)$, by nonconfusion. So we will use the third component to distinguish between $\alpha(s) \in \Gamma'$ and $\alpha(s) = \infty$. The value of the first component already determines whether or not $\alpha(s) \in \Gamma' \cup \{\infty\}$, so we really just need to be able to tell when $\alpha(s) = \infty$. There are several cases to consider, depending on whether or not s contains a tree t such that $\alpha(t) = \infty$. If not, then $s = t_1 + \cdots + t_r$, where $\alpha(t_i) \geq \Gamma'$ for all i. Observe that if this is the case, then the set of values $\{\alpha(t_1), \cdots, \alpha(t_r)\}$ is determined by the second component $\{[t_1^{\alpha_1}]_{\sim_k}, \ldots, [t_r^{\alpha_1}]_{\sim_k}\}$ of $\beta(s)$. If s contains a tree t such that $\alpha(t) = \infty$, pick such a tree at maximal depth. Then $t = a(t_1 + \cdots + t_r)$, where once again $\alpha(t_i) \geq \Gamma'$ for all i, and the set of values $\{\alpha(t_1), \cdots, \alpha(t_r)\}$ is determined by the second component of $\beta(s)$. We now specify the value of $\alpha_3(a, h, Q)$. As remarked above, Q determines a set of values all in Γ' or strictly higher. Let $h_Q \in H_A$ be the sum of these values. If either $h_Q = \infty$, or $ah_Q = \infty$, set $\alpha_3(a, h, Q) = 0$. Otherwise, $\alpha_3(a, h, q) = 1$.

The third component of $\beta(s)$ will be ∞ if and only if there is some subtree $a(t_1 + \cdots + t_r)$ such that

$$\alpha_3(a, \alpha_1(t_1 + \cdots + t_r), \{[t_1^{\alpha_1}]_{\sim_k}, \ldots, [t_r^{\alpha_1}]_{\sim_k}\}) = 0.$$

If we pick the subtree of maximal depth at which this occurs, then as argued above, $\alpha(s) = \infty$. The only other way we can have $\alpha(s) = \infty$ is if there is no such subtree, but $s = t_1 + \cdots + t_r$ where each $\alpha(t_i) \geq \Gamma'$ and the sum of these values is ∞. In this case, the fact that no such subtree exists is determined by the third coordinate of $\beta(s)$ being 1, and the set of $\alpha(t_i) \geq \Gamma'$ is determined by the second coordinate of $\beta(s)$. So in all cases $\beta(s)$ determines $\alpha(s)$.

6.3 Necessity of the Condition

To prove the converse, we have to show preservation of nonconfusion under quotients and wreath products with the allowable factors. This is carried out in the following three lemmas. Preservation under quotients (Lemma 21) is the most difficult of the three to show.

Lemma 21. *Let $\alpha : A^{\Delta} \to (H_1, V_1), \beta : A^{\Delta} \to (H_2, V_2)$, be homomorphisms onto finite forest algebras such that β factors through α. If α is nonconfusing then so is β.*

Lemma 22. *Suppose that $\alpha : A^{\Delta} \to (H, V) \circ \mathcal{U}_1$ is a homomorphism, and that $\beta = \pi\alpha$, where π is the projection morphism onto (H, V), is nonconfusing. Then α is nonconfusing.*

Lemma 23. *Suppose that $\alpha = \beta \otimes \gamma : A^{\Delta} \to (H,V) \circ (H',V')$ is a homomorphism, that β is nonconfusing, and that $\gamma : (A \times H)^{\Delta} \to (H',V')$ is 1-definite. Then α is nonconfusing.*

7 Results

Using the wreath product characterizations of EF-algebras, EX-homomorphisms, and $(\mathsf{EF},\mathsf{EX})$-homomorphisms of the previous three sections, we get:

Theorem 24. *Let A be a finite alphabet, and let $L \subseteq H_A$.*

(a) L is defined by an EF-formula if and only if (H_L, V_L) is an EF-algebra.
(b) L is defined by an EX-formula if and only if μ_L is an EX-homomorphism.
(c) L is defined by an $\mathsf{EF}+\mathsf{EX}$-formula if and only if μ_L is an $(\mathsf{EF},\mathsf{EX})$-homomorphism.
(d) There are effective procedures for determining, given a finite tree automaton recognizing L, whether L is definable by an EF-, EX-, or $\mathsf{EF}+\mathsf{EX}$-formula, and for producing a defining formula in case one exists.

References

[BS09] Benedikt, M., Segoufin, L.: Regular tree languages definable in FO and in $\mathrm{FO}_{\mathrm{mod}}$. ACM Trans. Comput. Log. **11**(1) (2009)

[BSS12] Bojanczyk, M., Segoufin, L., Straubing, H.: Piecewise testable tree languages. Logical Methods in Computer Science **8**(3) (2012)

[BSW12] Bojanczyk, M., Straubing, H., Walukiewicz, I.: Wreath products of forest algebras with applications to tree logics. Logical Methods in Computer Science **8**(3) (2012)

[BW06] Bojanczyk, M., Walukiewicz, I.: Characterizing EF and EX tree logics. Theor. Comput. Sci. **358**(2–3), 255–272 (2006)

[BW08] Bojanczyk, M., Walukiewicz, I.: Forest algebras. In: Flum, J., Grädel, E., Wilke, T. (eds.) Logic and Automata. Texts in Logic and Games, vol. 2, pp. 107–132. Amsterdam University Press (2008)

[CPP93] Cohen, J., Perrin, D., Pin, J.-E.: On the expressive power of temporal logic. J. Comput. Syst. Sci. **46**(3), 271–294 (1993)

[Ési05] Ésik, Z.: An algebraic characterization of the expressive power of temporal logics on finite trees. In: 1st Int. Conf. Algebraic Informatics. Aristotle Univ. of Thessaloniki, pp. 53–110 (2005)

[KS14] Krebs, A., Straubing, H.: EF+EX forest algebras. CoRR, abs/1408.0809 (2014)

[Sti73] Stiffler, P.E.: Extension of the fundamental theorem of finite semigroups. Advances in Mathematics **11**(2), 159–209 (1973)

[Str94] Straubing, H.: Finite Automata, Formal Logic, and Circuit Complexity. Birkhäuser, Boston (1994)

[TW96] Thérien, D., Wilke, T.: Temporal logic and semidirect products: An effective characterization of the until hierarchy. In: FOCS, pp. 256–263. IEEE Computer Society (1996)