Annales Academiæ Scientiarum Fennicæ
Mathematica
Volumen 25, 2000, 325–336

ON SOME CLASSES OF TREE AUTOMATA AND TREE LANGUAGES

Ferenc Gécseg

József Attila University, Department of Informatics
Aradi vértanúk tere 1, H-6720 Szeged, Hungary; gecseg@inf.u-szeged.hu

Abstract. In this paper we give a structural characterization of three classes of tree automata. Namely, we shall homomorphically represent the classes of nilpotent, definite, and monotone tree automata by means of quasi-cascade-products of unary nilpotent and unary definite tree automata in the first two cases, and by means of products of simpler tree automata in the third case.

Introduction

In the sixties, an intensive axiomatic investigation of the algebra of regular languages and some important special language classes was carried out. One of the best results for the axiomatic characterizations of regular languages can be found in [11]. The studies of special languages concerned mainly the classes of nilpotent and definite languages (see [8], [12] and [1], [6], [10], [9], respectively). The concepts of nilpotent languages, nilpotent automata, definite languages, and definite automata, together with the basic results characterizing them, have been generalized to trees. In [14] a nice survey is given of the structural properties of the classes of nilpotent and definite tree automata.

1. Notions and notation

Sets of operational symbols will be denoted by Σ with or without superscripts. If Σ is finite, then it is called a *ranked alphabet*. For the subset of Σ consisting of all m-ary operational symbols from Σ we shall use the notation Σ_m $(m \geq 0)$. By a Σ-*algebra* we mean a pair $\mathfrak{A} = (A, \{\sigma^{\mathfrak{A}} \mid \sigma \in \Sigma\})$, where $\sigma^{\mathfrak{A}}$ is an m-ary operation on A if $\sigma \in \Sigma_m$. If there will be no danger of confusion then we omit the superscript $\mathfrak{A}$ in $\sigma^{\mathfrak{A}}$ and simply write $\mathfrak{A} = (A, \Sigma)$. Finally, all algebras considered in this paper will be finite, i.e. A is finite and Σ is a ranked alphabet.

A *rank type* is a nonvoid set R of nonnegative integers. A set Σ of operational symbols is *of rank type* R if $\{m \mid \Sigma_m \neq \emptyset\} = R$. An algebra $\mathfrak{A} = (A, \Sigma)$ is *of rank type* R if Σ is of rank type R. To avoid technical difficulties, we shall suppose

---

1991 Mathematics Subject Classification: Primary 68Q70.

Supported by the Hungarian National Foundation for Scientific Research under Grant T 014888 and by the Ministry of Culture and Education of Hungary under Grant FKFP 0704/1977.

that $0 \notin R$ for all rank types R considered in this paper. It is also assumed that none of the sets of operational symbols contains any 0-ary symbol.

Let Ξ be a set of variables. The set $T_\Sigma(\Xi)$ of $\Sigma\Xi$*-trees* is defined as follows:

(i) $\Xi \subseteq T_\Sigma(\Xi)$,

(ii) $\sigma(p_1, \ldots, p_m) \in T_\Sigma(\Xi)$ whenever $m > 0$, $\sigma \in \Sigma_m$ and $p_1, \ldots, p_m \in T_\Sigma(\Xi)$,

and

(iii) every $\Sigma\Xi$-tree can be obtained by applying the rules (i) and (ii) a finite number of times.

In the sequel Ξ will stand for the countable set $\{\xi_1, \xi_2, \ldots\}$ and for every $m > 0$, Ξ_m will denote the subset $\{\xi_1, \ldots, \xi_m\}$ of Ξ. The subset of $T_\Sigma(\Xi_m)$ consisting of all trees in which each ξ_i $(1 \leq \xi_i \leq m)$ occurs exactly once will be denoted by $\widehat{T}_\Sigma(\Xi_m)$. Moreover, if $\Sigma = \Sigma_1$ then the elements of $T_\Sigma(\Xi_1)$ will be called also *words* in analogy with classical automata theory.

If $p \in T_\Sigma(\Xi_l)$ and $p_1, \ldots, p_l \in T_\Sigma(\Xi_m)$ are trees, then $p(p_1, \ldots, p_l) \in T_\Sigma(\Xi_m)$ is the tree obtained by replacing each occurrence of ξ_i $(i = 1, \ldots, l)$ in p by p_i. Moreover, if for $\Sigma = \Sigma_1$, $p \in T_\Sigma(\Xi_1)$ can be given in the form $p = q(r)$ $(q, r \in T_\Sigma(\Xi_1))$ then q is a *suffix* of p.

Let $\mathfrak{A} = (A, \Sigma)$ be an algebra and $p \in T_\Sigma(\Xi_n)$ a tree for a positive integer n. The mapping $p^{\mathfrak{A}}\colon A^n \to A$ is defined in the following way. For an arbitrary $\mathbf{a} = (a_1, \ldots, a_n) \in A^n$,

(i) if $p = x_i$ $(1 \leq i \leq n)$, then $p^{\mathfrak{A}}(\mathbf{a}) = a_i$,

(ii) if $p = \sigma(p_1, \ldots, p_m)$ $(\sigma \in \Sigma_m,\ p_1, \ldots, p_m \in T_\Sigma(\Xi_n))$, then $p^{\mathfrak{A}}(\mathbf{a}) = \sigma^{\mathfrak{A}}(p_1^{\mathfrak{A}}(\mathbf{a}), \ldots, p_m^{\mathfrak{A}}(\mathbf{a}))$.

A *tree recognizer* is a system $\mathbf{A} = (\mathfrak{A}, \mathbf{a}, A')$, where

(i) $\mathfrak{A} = (A, \Sigma)$ is an algebra,

(ii) $\mathbf{a} \in A^n$ is the *initial vector* for some positive integer n,

(iii) $A' \subseteq A$ is the set of *final states*.

The forest $T(\mathbf{A})$ *recognized* by $\mathbf{A}$ consists of all trees $p \in T_\Sigma(\Xi_n)$ for which $p(\mathbf{a}) \in A'$.

Classical recognizers are special tree recognizers. Namely, if in the above definition $\Sigma = \Sigma_1$ and $n = 1$, then $\mathfrak{A}$ is a finite state recognizer and every finite state recognizer can be obtained this way.

Let R be a rank type and take the algebras $\mathfrak{A}_i = (A_i, \Sigma^{(i)})$ $(i = 1, \ldots, k > 0)$ with rank type R, and let

$$\varphi = \{\varphi_m : (A_1 \times \ldots \times A_k)^m \times \Sigma_m \to \Sigma_m^{(1)} \times \cdots \times \Sigma_m^{(k)} \mid \sigma \in \Sigma_m,\ m \in R\}$$

be a family of mappings, where Σ is an arbitrary ranked alphabet of rank type R. Then by the *product* of $\mathfrak{A}_1, \ldots, \mathfrak{A}_k$ with respect to Σ and φ we mean the algebra

$\mathfrak{A}=(A,\Sigma)$ with $A=A_1\times\cdots\times A_k$ such that for arbitrary $\sigma\in\Sigma_m$ $(m\in R)$ and $(a_{11},\ldots,a_{1k}),\ldots,(a_{m1},\ldots,a_{mk})\in A$,

$$\sigma^{\mathfrak{A}}\big((a_{11},\ldots,a_{1k}),\ldots,(a_{m1},\ldots,a_{mk})\big)\\ =\big(\sigma_1^{\mathfrak{A}_1}(a_{11},\ldots,a_{m1}),\ldots,\sigma_k^{\mathfrak{A}_k}(a_{1k},\ldots,a_{mk})\big),$$

where $(\sigma_1,\ldots,\sigma_k)=\varphi_m\big((a_{11},\ldots,a_{1k}),\ldots,(a_{m1},\ldots,a_{mk}),\sigma\big)$ (see [13]).

For this product we use the notation $\mathfrak{A}=(\mathfrak{A}_1\times\cdots\times\mathfrak{A}_k)[\Sigma,\varphi]$. If $\mathfrak{A}_1=\cdots=\mathfrak{A}_k=\mathfrak{B}$ then we speak of a *power* of $\mathfrak{B}$ and write $\mathfrak{A}=(\mathfrak{B}^n)[\Sigma,\varphi]$.

Obviously, in the above definition of the product, and in its generalizations to be considered in this paper, φ_m can be given in the form

$$\varphi_m\big((a_{11},\ldots,a_{1k}),\ldots,(a_{m1},\ldots,a_{mk}),\sigma\big)\\ =\big(\varphi_m^{(1)}\big((a_{11},\ldots,a_{1k}),\ldots,(a_{m1},\ldots,a_{mk}),\sigma\big),\ldots,\\ \varphi_m^{(k)}\big((a_{11},\ldots,a_{1k}),\ldots,(a_{m1},\ldots,a_{mk}),\sigma\big)\big),$$

where $\varphi_m^{(i)}\colon A_1\times\cdots\times A_k\times\Sigma_m\to\Sigma_m^{(i)}$ $(i=1,\ldots,k)$ are suitable functions. If for each i $(1\le i\le k)$, every $\varphi_m^{(i)}$ may depend only on elements from algebras $\mathfrak{A}_1,\ldots,\mathfrak{A}_{i-1}$, then we speak of a *cascade product.* In this case sometimes we shall indicate only those arguments of $\varphi_m^{(i)}$ on which it may depend, i.e. we write

$$\varphi_m^{(i)}(a_{11},\ldots,a_{1i-1},\ldots,a_{m1},\ldots,a_{mi-1},\sigma)$$

for

$$\varphi_m^{(i)}\big((a_{11},\ldots,a_{1k}),\ldots,(a_{m1},\ldots,a_{mk}),\sigma\big).$$

The concept of the product can be generalized in such a way that the inputs of the component algebras are trees. For this we use the name generalized product.

Let $\mathfrak{A}_i=(A_i,\Sigma_i)$ $(i=1,\ldots,k>0)$ be arbitrary algebras, and

$$\varphi=\{\varphi_m\colon(A_1\times\cdots\times A_k)^m\times\Sigma_m\to\\ T_{\Sigma_1}(\Xi_m)\times\cdots\times T_{\Sigma_k}(\Xi_m)\mid\sigma\in\Sigma_m,\ m\in R\}$$

a family of mappings, where Σ is an arbitrary ranked alphabet. Then by the *generalized product* of $\mathfrak{A}_1,\ldots,\mathfrak{A}_k$ with respect to Σ and φ we mean the algebra $\mathfrak{A}=(A,\Sigma)$ with $A=A_1\times\cdots\times A_k$ such that for arbitrary $\sigma\in\Sigma_m$ $(m\in R)$ and $(a_{11},\ldots,a_{1k}),\ldots,(a_{m1},\ldots,a_{mk})\in A$,

$$\sigma^{\mathfrak{A}}\big((a_{11},\ldots,a_{1k}),\ldots,(a_{m1},\ldots,a_{mk})\big)\\ =\big(p_1^{\mathfrak{A}_1}(a_{11},\ldots,a_{m1}),\ldots,p_k^{\mathfrak{A}_k}(a_{1k},\ldots,a_{mk})\big),$$

where $(p_1, \dots, p_k) = \varphi_m\big((a_{11}, \dots, a_{1k}), \dots, (a_{m1}, \dots, a_{mk}), \sigma\big)$.

For the generalized product we shall use the same notation as for the product. Moreover, the *generalized power*, the *generalized cascade product* and the *generalized cascade power* together with their notations are given in a natural way.

It has been shown that a special type of the generalized product, the so called quasi-product is powerful enough to represent algebras. A generalized product $\mathfrak{A} = (\mathfrak{A}_1 \times \cdots \times \mathfrak{A}_k)[\Sigma, \varphi]$ is a *quasi-product* if in

$$(p_1, \dots, p_k) = \varphi_m\big((a_{11}, \dots, a_{1k}), \dots, (a_{m1}, \dots, a_{mk}), \sigma\big)$$

every p_j $(1 \le j \le k)$ is of the form $\sigma_j(\xi_{j_1}, \dots, \xi_{j_{t_j}})$ $(1 \le j_1, \dots, j_{t_j} \le m)$. Therefore, the quasi-product is that slight generalization of the product when in the inputs of the component algebras permutations and identifications of variables are allowed.

Again, the *quasi-power*, the *quasi-cascade-product*, the *quasi-cascade-power*, and the corresponding notations are given in an obvious way.

The aim of the structure theory of tree automata is to represent tree recognizers by means of compositions of algebras. Among the representations homomorphism and isomorphism are most frequently used. It has been shown (see [3]) that we can confine ourselves to the homomorphic and isomorphic representations of the underlying algebras of recognizers.

Let composition mean any of the products introduced in this paper. Moreover, let $\mathscr{K}_1$ and $\mathscr{K}_2$ be two classes of algebras. We say that $\mathscr{K}_1$ is *homomorphically* (*isomorphically*) *complete* for $\mathscr{K}_2$ with respect to the composition if every algebra from $\mathscr{K}_2$ can be given as a homomorphic (isomorphic) image of a subalgebra of a composition of algebras from $\mathscr{K}_1$.

Finally, we introduce some operators on classes of algebras. If $\mathscr{K}$ is a class of algebras, then

- $\mathbf{S}(\mathscr{K})$ is the class of all subalgebras of algebras from $\mathscr{K}$,
- $\mathbf{H}(\mathscr{K})$ is the class of all homomorphic images of algebras from $\mathscr{K}$,
- $\mathbf{Q}(\mathscr{K})$ is the class of all quasi-cascade-products of algebras from $\mathscr{K}$.

For notions and notations not defined here, see [5].

2. Nilpotent automata

A unary algebra $\mathfrak{A} = (A, \Sigma)$ is said to be *nilpotent* if there are an element $a \in A$, the *absorbent element* of $\mathfrak{A}$, and a nonnegative integer k such that $p(b) = a$ for all $b \in A$ and $p \in T_\Sigma(\Xi_1)$ with $h(p) \ge k$. Moreover, an automaton $\mathbf{A} = (\mathfrak{A}, a_0, A')$ is *nilpotent* if $\mathfrak{A}$ is nilpotent. Finally, a language $T \subseteq T_\Sigma(\Xi_1)$ is *nilpotent* if $T = T(\mathbf{A})$ for a nilpotent automaton $\mathbf{A}$.

The following result gives a characterization of nilpotent automata by means of Boolean operations.

Theorem 1. *A language is nilpotent if and only if it can be obtained from the empty language and one-element languages by forming union and complementation.* □

There is also a set theoretic characterization of nilpotent languages.

Theorem 2. *A language T is nilpotent if and only if T or the complement of T is finite.* □

Finally, we recall a semigroup theoretic characterization of nilpotent languages.

Theorem 3. *A language is nilpotent if and only if its syntactic semigroup is nilpotent.* □

The concept of nilpotent automata has been generalized to tree automata in a natural way (see [4]).

A Σ-algebra $\mathfrak{A} = (A, \Sigma)$ is *nilpotent* if there are an $a \in A$ and a nonnegative integer k such that $p(\mathbf{a}) = a$ for all $p \in T_\Sigma(\Xi_n)$, $n > 0$, and $\mathbf{a} \in A^n$, whenever $h(p) \geq k$. Moreover, a tree automaton $\mathbf{A} = (\mathfrak{A}, \mathbf{a}, A')$ is *nilpotent* if $\mathfrak{A}$ is nilpotent. Finally, a tree language $T \subseteq T_\Sigma(\Xi_n)$ is *nilpotent* if $T = T(\mathbf{A})$ for a nilpotent tree automaton $\mathbf{A}$.

A result analogous to Theorem 2 can be easily obtained for tree automata. Here we give a structural characterization of the class of nilpotent algebras by means of quasi-cascade-products of unary algebras. For this, let $\mathscr{N}_1$ denote the class of all nilpotent unary algebras.

Theorem 4. *A Σ-algebra $\mathfrak{A} = (A, \Sigma)$ is nilpotent if and only if*

$$\mathfrak{A} \in \mathbf{H}(\mathbf{S}(\mathbf{Q}(\mathscr{N}_1))).$$

Proof. It is easy to see that any quasi-product of nilpotent algebras is nilpotent. Moreover, the formation of subalgebras and homomorphic images preserves nilpotency. Therefore, the conditions of Theorem 4 are sufficient.

To prove the necessity, take a nilpotent Σ-algebra $\mathfrak{A} = (A, \Sigma)$ with n elements. The case $n = 1$ being trivial, let $n > 1$. We may assume that $A = \{1, \ldots, n\}$. Moreover, since A is finite, we may suppose that $\sigma^{\mathfrak{A}}(i_1, \ldots, i_m) \geq i_1, \ldots, i_m$ for all $\sigma \in \Sigma_m$, $m > 0$, and $i_1, \ldots, i_m \in A$. (Observe, that $i \leq p(i_1, \ldots, i_{j-1}, i, i_{j+1}, \ldots, i_m)$ $(i_1, \ldots, i_{j-1}, i, i_{j+1}, \ldots, i_m \in A$, $m > 0$, $1 \leq i \leq m$, $p \in \widehat{T}_\Sigma(\Xi_m))$ is a partial ordering on A.) Thus, n is the absorbent element of $\mathfrak{A}$. Take a (unary) algebra $\mathfrak{B} = (B, \overline{\Sigma})$ from $\mathscr{N}_1$ with $B = \{1, \ldots, n\}$ such that for every pair (i, j) satisfying $1 \leq i < j \leq n$ there is a $\sigma^{(i,j)} \in \overline{\Sigma}$ such that $\sigma^{(i,j)}(i) = j$. Furthermore, suppose that there is a symbol $\sigma^{(n)} \in \overline{\Sigma}$ for which $\sigma^{(n)}(i) = n$ for all i $(= 1, \ldots, n)$. Let us define the quasi-power $\mathfrak{C} = (C, \Sigma) = (\mathfrak{B}^n)[\Sigma, \varphi]$ in the following way. First of all, for each i $(1 \le i \le n)$ let $\mathbf{c}_i = (c_{i1}, \dots, c_{in})$ be the element of C for which

$$c_{ij} = \begin{cases} n, & \text{if } 1 \le j < i, \\ i, & \text{if } j = i, \dots, n. \end{cases}$$

Denote by C' the set of all such elements $\mathbf{c}_i$ $(i = 1, \dots, n)$. Now for all $m > 0$, $\sigma \in \Sigma_m$, and $\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m} \in C'$, let $\varphi_m^{(1)}(\sigma) = \sigma^{(n)}(\xi_1)$. Moreover, if $1 < j \le n$, then $\varphi_m^{(j)}(\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m}, \sigma)$ is given as follows:

$$\varphi_m^{(j)}(\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m}, \sigma) = \begin{cases} \sigma^{(n)}(\xi_1), & \text{if } \sigma^{\mathfrak{A}}(c_{i_1 j-1}, \dots, c_{i_m j-1}) = u \text{ and } j < u, \\ \sigma^{c_{i_1 j-1}, u}(\xi_1), & \text{if } \sigma^{\mathfrak{A}}(c_{i_1 j-1}, \dots, c_{i_m j-1}) = u \text{ and } j \ge u. \end{cases}$$

In all other cases φ is given arbitrarily in accordance with the definition of the quasi-cascade-product. It is obvious that $\mathfrak{C}$ is a quasi-cascade-power of $\mathfrak{B}$.

We show that if $\sigma^{\mathfrak{A}}(i_1, \dots, i_m) = u$, then we have for $\sigma^{\mathfrak{C}}(\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m}) = \mathbf{c} = (c_1, \dots, c_n)$

$$c_j = \begin{cases} n, & \text{if } 1 \le j < u, \\ u, & \text{if } u \le j \le n. \end{cases}$$

Clearly, $u > 1$. If $u = n$ and $\max(i_1, \dots, i_m) = n$ then $c_1 = \dots = c_n = n$ obviously holds. If $u = n$ and $\max(i_1, \dots, i_m) = i_t < n$ $(1 \le t \le m)$, then

(I) $c_1 = \dots = c_{i_t} = n$ since $c_{i_t 1} = \dots = c_{i_t i_t - 1} = n$ whenever $i_t > 1$; thus, $\varphi_m^{(j)}(\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m}, \sigma) = \sigma^{(n)}(\xi_1)$ for all $j = 1, \dots, i_t$.

(II) $c_{i_t+1} = \dots = c_{n-1} = n$ since for every t $(i_t \le t < n-1)$, $c_{i_1 t} = i_1, \dots, c_{i_m t} = i_m$, $\sigma^{\mathfrak{A}}(i_1, \dots, i_m) = n$; thus, $\varphi_m^{(j)}(\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m}, \sigma) = \sigma^{(n)}(\xi_1)$ for all j $(i_t < j < n)$.

(III) $c_n = n$ since $c_{i_1 n-1} = i_1, \dots, c_{i_m n-1} = i_m$, $\sigma^{\mathfrak{A}}(i_1, \dots, i_m) = n$; thus, $\varphi_n(\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m}, \sigma) = \sigma^{(i_1, n)}(\xi_1)$.

Next assume that $u < n$. Again, let $\max(i_1, \dots, i_m) = i_t$. Obviously, $i_t < u$. Then

(i) $c_1 = \dots = c_{i_t} = n$ since if $i_t > 1$ then $c_{i_t 1} = \dots = c_{i_t i_t - 1} = n$; thus, $\varphi_m^{(j)}(\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m}, \sigma) = \sigma^{(n)}(\xi_1)$ for all $j = 1, \dots, n$.

(ii) $c_{i_t+1} = \dots = c_{u-1} = n$ since for every r $(i_t \le r < u-1)$, $c_{i_1 r} = i_1, \dots, c_{i_m r} = i_m$, $\sigma^{\mathfrak{A}}(i_1, \dots, i_m) = u$; thus, $\varphi_m^{(j)}(\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m}, \sigma) = \sigma^{(n)}(\xi_1)$ for all j $(v < j < u)$.

(iii) $c_u = \dots = c_n = u$ since for every r $(u - 1 \le r \le n)$, $c_{i_1 r} = i_1, \dots, c_{i_m r} = i_m$, $\sigma^{\mathfrak{A}}(i_1, \dots, i_m) = u$; thus, $\varphi_m^{(j)}(\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_m}, \sigma) = \sigma^{(i_1, u)}(\xi_1)$ for all $j = u, \dots, n$.

Thus, we have proved that C' forms a subalgebra of $\mathfrak{C}$ which is isomorphic to $\mathfrak{A}$ under the mapping $\tau \colon C' \to A$ given by $\tau(c_i) = i$ $(i = 1, \dots, n)$. $\square$

From the proof of the necessity of Theorem 4 we obtain

Corollary 1. *Every nilpotent algebra can be represented isomorphically by a quasi-cascade-product of unary nilpotent algebras.* □

3. Definite automata

A unary algebra $\mathfrak{A} = (A, \Sigma)$ is *definite* if there is a nonnegative integer k such that for arbitrary $p \in T_\Sigma(\Xi_1)$ with $h(p) \geq k$ and $a, b \in A$ we have $p(a) = p(b)$. Moreover, an automaton $\mathbf{A} = (\mathfrak{A}, a_0, A')$ is *definite* if $\mathfrak{A}$ is definite. Finally, a language $T \subseteq T_\Sigma(\Xi_1)$ is *definite* if there is a definite automaton $\mathbf{A}$ with $T = T(\mathbf{A})$.

The following result holds.

Theorem 5. *A language $T \subseteq T_\Sigma(\Xi_1)$ is definite if and only if there is an integer $k \geq 0$ such that for arbitrary two words $p, q \in T_\Sigma(X_1)$ with $h(p), h(q) \geq k$ and the same suffix of length k we have $p \in T \Leftrightarrow q \in T$.* □

Definite automata and definite languages have been generalized to trees (see [7]). Before giving the definition of definite tree automata and definite tree languages, we recall the concept of the cut-off of a tree at a given height.

For every nonnegative integer k, take the binary relations $\sim_k$ on $T_\Sigma(\Xi_n)$ given in the following way:

(i) If $k = 0$, then $p \sim_0 q$ for all $p, q \in T_\Sigma(\Xi_n)$.

(ii) Suppose that $\sim_k$ has been defined. Then for all $p, q \in T_\Sigma(\Xi_n)$, $p \sim_{k+1} q$ if and only if there are a $\sigma \in \Sigma_m$, $p_1, \ldots, p_m, q_1, \ldots, q_m \in T_\Sigma(\Xi_n)$ such that $p = \sigma(p_1, \ldots, p_m)$, $q = \sigma(q_1, \ldots, q_m)$ and $p_i \sim_k q_i$ for all $i = 1, \ldots, k$.

It is said that two trees $p, q \in T_\Sigma(\Xi_n)$ have the same *cut-off* at height k if $p \sim_k q$.

A Σ-algebra $\mathfrak{A} = (A, \Sigma)$ is *definite* if there is a nonnegative integer k such that for arbitrary $n > 0$, $\mathbf{a} \in A^n$ and $p, q \in T_\Sigma(\Xi_n)$ with $h(p), h(q) \geq k$ we have $p(\mathbf{a}) = q(\mathbf{a})$, whenever p and q have the same cut-off at height k. Moreover, a tree automaton $\mathbf{A} = (\mathfrak{A}, \mathbf{a}, A')$ is *definite* if $\mathfrak{A}$ is definite. Finally, a tree language is *definite* if it can be recognized by a definite tree automaton.

The following result, which is a generalization of Theorem 5 to trees, is from [7].

Theorem 6. *A tree language $T \subseteq T_\Sigma(\Xi_n)$ is definite if and only if there is an integer $k \geq 0$ such that for arbitrary two trees $p, q \in T_\Sigma(\Xi_n)$ with $h(p), h(q) \geq k$ and the same cut-off at height k we have $p \in T \Leftrightarrow q \in T$.* □

Take a rank type R. Let Σ be a ranked alphabet having two symbols $\sigma_m^1, \sigma_m^2 \in \Sigma_m$ for every $m \in R$. Define the Σ-algebra $\mathfrak{A}_R = (\{1,2\}, \Sigma)$ in the following way: $\sigma_m^i(i_1, \ldots, i_m) = i$ $(i = 1, 2)$ for all $m \in R$ and $i_1, \ldots, i_m \in \{1, 2\}$. It is easy to see that $\mathfrak{A}_R$ is definite under $k = 1$.

In [2] Ésik gives the following structural characterization of definite algebras.

Theorem 7. *For an algebra $\mathfrak{A}$ of rank type R the next statements are equivalent.*

(i) $\mathfrak{A}$ *is definite.*
(ii) $\mathfrak{A}$ *can be embedded isomorphically into a cascade power of* $\mathfrak{A}_R$.
(iii) $\mathfrak{A}$ *is a homomorphic image of a subalgebra of a cascade power of* $\mathfrak{A}_R$. □

It is not hard to see that $\mathfrak{A}_R$ is isomorphic to a quasi-cascade-product of $\mathfrak{A}_{\{1\}}$ with a single factor. Moreover, since the quasi-cascade-product is a special case of the generalized cascade composition, by Corollary 5 in [2], any quasi-cascade-product of definite algebras is definite. Furthermore, subalgebras and homomorphic images preserve definiteness. Finally, the formation of quasi-cascade-products is transitive. Therefore, from Theorem 7 we obtain

Theorem 8. *For an algebra* $\mathfrak{A}$ *the next statements are equivalent.*

(i) $\mathfrak{A}$ *is definite.*
(ii) $\mathfrak{A}$ *can be embedded isomorphically into a quasi-cascade-power of* $\mathfrak{A}_{\{1\}}$.
(iii) $\mathfrak{A}$ *is a homomorphic image of a subalgebra of a quasi-cascade-power of* $\mathfrak{A}_{\{1\}}$. □

4. Monotone automata

Monotone automata play a distinguished role in the structure theory of finite state automata. At the same time no characterization of languages recognized by monotone automata is known. Here we introduce the concept of monotone algebras and give some structural characterization for them.

An algebra $\mathfrak{A} = (A, \Sigma)$ is *monotone* if a partial ordering $\leq$ can be given on A such that for any $m > 0$, $\sigma \in \Sigma_m$, and $a_1, \ldots, a_m \in A$ we have $\sigma(a_1, \ldots, a_m) \geq a_1, \ldots, a_m$. Observe, that every nilpotent algebra is monotone. Theorem 9, the proof of which is trivial, shows that monotonicity can be reduced to unary algebras.

Let $\mathfrak{A} = (A, \Sigma)$ be an algebra. Moreover, let $\bar{\mathfrak{A}} = (A, \Sigma^{(\mathfrak{A})})$ be the algebra where $\Sigma^{(\mathfrak{A})} = \Sigma_1^{(\mathfrak{A})}$ consists of all $\sigma(a_1, \ldots, a_{i-1}, \xi_1, a_{i+1}, \ldots, a_m)$ $(m > 0,\ 1 \leq i \leq m,\ \sigma \in \Sigma_m,\ a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_m \in A)$, and for each $a \in A$,

$$\bar{\sigma}^{\bar{\mathfrak{A}}}(a) = \sigma^{\mathfrak{A}}(a_1, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_m)$$

if $\bar{\sigma} = \sigma(a_1, \ldots, a_{i-1}, \xi_i, a_{i+1}, \ldots, a_m)$. In other words, the operations of $\bar{\mathfrak{A}}$ are all the elementary translations of $\mathfrak{A}$.

Theorem 9. *An algebra* $\mathfrak{A} = (A, \Sigma)$ *is monotone if and only if the unary algebra* $\bar{\mathfrak{A}} = (A, \Sigma^{(\mathfrak{A})})$ *is monotone.* □

We shall need the following result, too.

Lemma 1. *Homomorphism preserves monotonicity of algebras.*

Proof. Let $\mathfrak{A} = (A, \Sigma)$ and $\mathfrak{B} = (B, \Sigma)$ be algebras and $\tau\colon A \to B$ a homomorphism of $\mathfrak{A}$ onto $\mathfrak{B}$. Assume that $\mathfrak{A}$ is monotone under the partial ordering $\leq$. Let us define the relation $\varrho_{\mathfrak{A}}$ on A by $a \varrho_{\mathfrak{A}} p^{\overline{\mathfrak{A}}}(a)$ $(a \in A,\ p \in T_{\Sigma(\mathfrak{A})}(\Xi_1))$. It is clear that $\varrho_{\mathfrak{A}}$ is reflexive and transitive. Moreover, $\varrho_{\mathfrak{A}} \subseteq \leq$. Therefore, $\varrho_{\mathfrak{A}}$ is a partial ordering and $\overline{\mathfrak{A}}$ is monotone under $\varrho_{\mathfrak{A}}$. Assume that the corresponding binary relation $\varrho_{\mathfrak{B}}$ is not a partial ordering on B. Then there are an element $b \in B$ and two words $\bar{p}, \bar{q} \in T_{\Sigma(\mathfrak{A})}(\Xi_1)$ such that $\bar{q}^{\overline{\mathfrak{B}}}(b) \neq b$ and $(\bar{p}(\bar{q}))^{\overline{\mathfrak{B}}}(b) = b$. Let $\bar{p}(\xi_1) = p(b_{11}, \ldots, b_{1m-1}, \xi_1)$ and $\bar{q}(\xi_1) = q(b_{21}, \ldots, b_{2n-1}, \xi_1)$ $(p \in \widehat{T}_{\Sigma}(\Xi_m)$, $q \in \widehat{T}_{\Sigma}(\Xi_n)$, $b_{11}, \ldots, b_{1m-1}, b_{21}, \ldots, b_{2n-1} \in B)$. Let $a_{11}, \ldots, a_{1m-1}, a_{21}, \ldots, a_{2n-1}$ be counter images of $b_{11}, \ldots, b_{1m-1}, b_{21}, \ldots, b_{2n-1}$, respectively. Moreover, let $\overline{\overline{p}}(\xi_1) = p(a_{11}, \ldots, a_{1m-1}, \xi_1)$ and $\overline{\overline{q}}(\xi_1) = q(a_{21}, \ldots, a_{2n-1}, \xi_1)$. For an arbitrary tree $t(\xi_1) \in T_{\Sigma(\mathfrak{A})}(\Xi_1)$ let $t^0 = \xi_1$ and $t^i = t\big(t^{i-1}(\xi_1)\big)$ if $i > 0$. Let a be a counter image of b under τ. Since τ is a homomorphism, $\big(\overline{\overline{p}}(\overline{\overline{q}})\big)^i(a) = b$ for every $i (= 0, 1, \ldots)$. Moreover, due to the finiteness of A, there are integers i and j such that $i < j$ and $\big(\overline{\overline{p}}(\overline{\overline{q}})\big)^i(a) = \big(\overline{\overline{p}}(\overline{\overline{q}})\big)^j(a)$. Furthermore, since $\bar{q}(b) \neq b$, $\overline{\overline{q}}\big((\overline{\overline{p}}(\overline{\overline{q}}))^i\big)(a) \neq \big(\overline{\overline{p}}(\overline{\overline{q}})\big)^i(a)$. Therefore, $\varrho_{\mathfrak{A}}$ is not a partial ordering on A, which is a contradiction. Thus, we have obtained that $\varrho_{\mathfrak{B}}$ is a partial ordering on B and $\overline{\mathfrak{B}}$ is a monotone unary algebra under $\varrho_{\mathfrak{B}}$ which, by Theorem 9, implies that $\mathfrak{B}$ is a monotone algebra. □

Let R be a rank type and let Σ be a ranked alphabet of rank type R. An algebra $\mathfrak{A} = (A, \Sigma)$ with $A = \{1, 2\}$ is a two-state *full* monotone algebra of rank type R if for all integers $m \in R$ and mappings $\tau\colon A^m \to A$ with $\tau(a_1, \ldots, a_m) \geq a_1, \ldots, a_m$ $(a_1, \ldots, a_m \in A)$, there is a $\sigma \in \Sigma_m$ such that $\sigma^{\mathfrak{A}} = \tau$.

We have

Theorem 10. *An algebra $\mathfrak{A} = (A, \Sigma)$ of rank type R is monotone if and only if it can be given as a homomorphic image of a subalgebra of a cascade power of a two-state full monotone algebra of rank type R.*

Proof. It is obvious that products and subalgebras of monotone algebras are monotone. Moreover, by Lemma 1, homomorphism preserves monotonicity.

Conversely, let $\mathfrak{A} = (A, \Sigma)$ be a monotone algebra of rank type R under a partial ordering $\leq$. Since A is finite, we may assume that $\leq$ is a linear ordering. For the sake of simplicity, suppose that $A = \{1, \ldots, n\}$ and $\leq$ is the natural ordering on A. Let $\mathfrak{B} = (\{1, 2\}, \Sigma')$ be a two-state full monotone algebra of rank type R. If $n = 2$ then $\mathfrak{A}$ is obviously isomorphic to a cascade power of $\mathfrak{B}$ with a single factor. Assume that $n > 2$. We shall show that there is an algebra $\mathfrak{A}' = (A', \Sigma)$ with $n - 1$ states $1, \ldots, n-1$ such that $\mathfrak{A}'$ is monotone under the natural ordering on A' and $\mathfrak{A}$ is a homomorphic image of a subalgebra of a cascade product $\mathfrak{C} = (C, \Sigma) = (\mathfrak{A}' \times \mathfrak{B})[\Sigma, \varphi]$. For arbitrary $m \in R$, $\sigma \in \Sigma_m$, and $i_1, \ldots, i_m \in A'$, let

$$\sigma^{\mathfrak{A}'}(i_1, \ldots, i_m) = \begin{cases} \sigma^{\mathfrak{A}}(i_1, \ldots, i_m), & \text{if } \sigma^{\mathfrak{A}}(i_1, \ldots, i_m) < n, \\ n-1 & \text{otherwise.} \end{cases}$$

For every $m \in R$, let $\sigma'_{m1} \in \Sigma'_m$ be the symbol satisfying ${\sigma'}^{\mathfrak{B}}_{m1}(1, \ldots, 1) = 1$. Moreover, $\sigma'_{m2} \in \Sigma'_m$ will denote a symbol such that ${\sigma'}^{\mathfrak{B}}_{m2}(i_1, \ldots, i_m) = 2$ for arbitrary $i_1, \ldots, i_m \in \{1, 2\}$. Now we are ready to define the feed-back function φ of $\mathfrak{C}$. For every $\sigma \in \Sigma_m$, let $\varphi_m^{(1)}(\sigma) = \sigma$. Furthermore, for all $\sigma \in \Sigma_m$ and $i_1, \ldots, i_m \in \{1, \ldots, n-1\}$, let

$$\varphi_m^{(2)}(i_1, \ldots, i_m, \sigma) = \begin{cases} \sigma_{m1}, & \text{if } \sigma(i_1, \ldots, i_m) < n, \\ \sigma_{m2} & \text{otherwise.} \end{cases}$$

Denote by C' the subset $\{(1,1), \ldots, (n-1,1), (n-1,2)\}$ of C. It is not hard to show that C' forms a subalgebra of $\mathfrak{C}$ and the mapping $\tau\colon C' \to A$ given by

$$\tau\big((i,j)\big) = \begin{cases} i, & \text{if } j = 1, \\ n, & \text{if } j = 2 \end{cases}$$

$\big((i,j) \in C'\big)$ is an isomorphism. Since the formation of the cascade product is transitive, this ends the proof of Theorem 10. □

From the proof of Theorem 10 we obtain

Corollary 2. *An algebra of rank type R is monotone if and only if it is isomorphic to a subalgebra of a cascade power of a two-state full monotone algebra of rank type R.* □

Let $\mathfrak{A} = (A, \Sigma')$ be the algebra with $A = \{1, 2\}$ and $\Sigma = \Sigma_1 = \{\sigma_1, \sigma_2\}$, where $\sigma_1(1) = 1$, $\sigma_1(2) = 2$ and $\sigma_2(1) = \sigma_2(2) = 2$. The algebra $\mathfrak{A}$ is a two-state full monotone algebra of rank type $R = \{1\}$. This, by Theorem 10, means that every unary monotone algebra can be given as a homomorphic image of a subalgebra of a cascade power of $\mathfrak{A}$. We shall give a two-state monotone algebra which cannot be represented homomorphically by a quasi-cascade-power of $\mathfrak{A}$. Therefore, since the formation of the quasi-cascade-product is transitive, this will imply that not every monotone algebra can be given as a homomorphic image of a subalgebra of a quasi-cascade-product of unary monotone algebras.

Consider the algebra $\mathfrak{B} = (B, \Sigma)$, where $B = \{1, 2\}$, $\Sigma = \Sigma_2 = \{\sigma\}$, $\sigma(1,1) = 1$ and $\sigma(1,2) = \sigma(2,1) = \sigma(2,2) = 2$. This $\mathfrak{B}$ is a monotone algebra.

Take a quasi-cascade-power $\mathfrak{C} = (C, \Sigma) = (\mathfrak{A}^n)[\Sigma, \varphi]$. Assume that a subalgebra $\mathfrak{C}' = (C', \Sigma)$ of $\mathfrak{C}$ can be mapped homomorphically onto $\mathfrak{B}$ under a mapping $\tau\colon C' \to B$. Let N denote the set of all elements $\mathbf{c}$ from C' for which $\tau(\mathbf{c}) = 1$

and $\sigma(\mathbf{c},\mathbf{c}) = \mathbf{c}$. Then N is nonvoid since $\mathfrak{C}'$ is monotone and $\sigma^{\mathfrak{B}}(1,1) = 1$. Furthermore, let N_2 be the set of all counter images of 2 under τ. Finally, let i $(1 \leq i \leq n)$ denote the maximal index for which there is a pair $(\mathbf{a},\mathbf{b})$,

$$\mathbf{a} = (j_{11},\ldots,j_{1i-1},j_{1i},j_{1i+1},\ldots,j_{1n}) \in N$$

and

$$\mathbf{b} = (j_{21},\ldots,j_{2i-1},j_{2i},j_{2i+1},\ldots,j_{2n}) \in N_2\,,$$

such that $j_{11} = j_{21},\ldots,j_{1i-1} = j_{2i-1}$ and $j_{1i} \neq j_{2i}$. Let us distinguish the following two cases.

(1) $j_{1i} = 1$ and $j_{2i} = 2$. Then $\varphi_2^{(i)}(j_{11},\ldots,j_{1i-1},j_{11},\ldots,j_{1i-1},\sigma) = \sigma_1(\xi_1)$ or $\sigma_1(\xi_2)$ since $\sigma(\mathbf{a},\mathbf{a}) = \mathbf{a}$. In the first case the ith component of $\sigma(\mathbf{a},\mathbf{b})$ is 1 and $\sigma(\mathbf{a},\mathbf{b}) \in N_2$. This contradicts either the maximality of i, or the assumption that τ is a homomorphism if $i = n$. In the second case, taking $\sigma(\mathbf{b},\mathbf{a})$, we arrive at the same conclusion.

(2) $j_{1i} = 2$ and $j_{2i} = 1$. Now for $\varphi_2^{(i)}(j_{11},\ldots,j_{1i-1},j_{11},\ldots,j_{1i-1},\sigma)$ we may have all four cases: $\varphi_2^{(i)}(j_{11},\ldots,j_{1i-1},j_{11},\ldots,j_{1i-1},\sigma) = \sigma_1(\xi_1)$, or $\sigma_1(\xi_2)$, or $\sigma_2(\xi_1)$, or $\sigma_2(\xi_2)$. Taking $\sigma(\mathbf{a},\mathbf{b})$ in the first case and $\sigma(\mathbf{b},\mathbf{a})$ in the remaining three cases, we have the same contradiction as previously. □

Let $\mathscr{M}_1$ denote the class of all monotone unary algebras and $\mathscr{M}$ the class of all monotone algebras. We can summarize the above result in

Theorem 11. *$\mathscr{M}_1$ is not homomorphically complete for $\mathscr{M}$ with respect to the quasi-cascade-product.* □

Finally, we remark that $\mathscr{M}$ and $\mathbf{H}(\mathbf{S}(\mathbf{Q}(\mathscr{M}_1)))$ are incomparable.

References

[1] Brzozowski, J.A.: Canonical regular expressions and minimal state graphs for definite events. - In: Proc. Symp. Math. Theory of Automata, Brooklyn, 1962, 529–562.

[2] Ésik, Z.: Definite tree automata and their cascade compositions. - Publ. Math. 48, 1996, 243–261.

[3] Gécseg, F.: On homomorphic representations by products of tree automata. - In: Results and Trends in Theoretical Computer Science, Proceedings of a symposium held in Graz in 1994, Lecture Notes in Comput. Sci. 812, Springer-Verlag, 131–139.

[4] Gécseg, F., and B. Imreh: On a special class of tree automata. - In: 2nd Conf. on Automata, Languages and Programming Systems, Salgótarján, 1988, 141–152.

[5] Gécseg, F., and M. Steinby: Tree Automata. - Akadémiai Kiadó, Budapest, 1984.

[6] Ginzburg, A.: About some properties of definite, reverse-definite events, and related automata. - IEEE Trans. Electronic. Comput. 15, 1966, 806–810.

[7] Heuter, U.: Definite tree languages. - EATCS Bulletin 35, 1988, 137–142.

[8] Kaphengst, H.: Axiomatizable classes of automata. - Voprosy teoretičeskoĭ kibernetiki, Kiev, 1965, 6–30 (Russian).

[9] Paz, A., and B. Peleg: Ultimate-definite and symmetric definite events and automata. - J. Assoc. Comput. Mach. 12, 1965, 399–410.

[10] PERLES, M., M.O. RABIN, and E. SHAMIR: The theory of definite automata. - IEEE Trans. Electronic Comput. 12, 1963, 233–243.

[11] SALOMAA, A.: Two complete axiom systems for the algebra of regular events. - J. Assoc. Comput. Mach. 13, 1966, 158–169.

[12] ŠEVRIN, L.N.: On some classes of abstract automata. - Uspekhi Mat. Nauk 17:6, 1962, 219 (Russian).

[13] STEINBY, M.: On the structure and realization of tree automata. - In: 2^{e} Colloque sur les arbres en algèbre et en programmation, Lille, 1977, 235–248.

[14] STEINBY, M.: A theory of tree language varieties. - In: Tree Automata and Languages, Elsevier Science Publishers, 1992, 57–81.

Received 11 August 1998