Acta Cybernetica 14 (1999) 135-149.

Tree Transducers and Formal Tree Series*

Werner Kuich †

Gécseg Ferenc barátomnak 60. születésnapja alkalmából

Abstract

We introduce tree transducers over formal tree series as a generalization of a restricted type of root-to-frontier tree transducers and show that linear nondeleting recognizable tree transducers do preserve recognizability of tree series.

1 Introduction and preliminaries

In this paper we give a uniform treatment of tree transducers and tree automata in terms of tree series and matrices.

In Section 2 we define tree transducers that map tree series into tree series. These tree transducers are a generalization of a restricted type of root-to-frontier tree transducers as described in Gécseg, Steinby [4, 5].

In Section 3 we consider linear and nondeleting tree representations and show certain algebraic properties of these tree representations.

In the last section we consider linear nondeleting recognizable tree transducers. Intuitively, these are generalizations of linear nondeleting root-to-frontier tree transducers with infinitely many productions whose right sides form recognizable tree languages. The main result of Section 4 is that linear nondeleting recognizable tree transducers do preserve recognizability of tree series. Our main result is a generalization of the following theorem of Thatcher [10]: *Linear root-to-frontier tree transducers preserve recognizability* (see also Gécseg, Steinby [4], Corollary IV.6.6).

It is assumed that the reader is familiar with the basics of semiring theory (see Kuich, Salomaa [9] and Kuich [6], Section 2). Throughout the paper, $\langle A, +, \cdot, 0, 1\rangle$ denotes a *commutative continuous* semiring. This means:

(o) the multiplication $\cdot$ is commutative;

*Partially supported by the "Stiftung Aktion Österreich-Ungarn".

†Abteilung für Theoretische Informatik Institut für Algebra und Diskrete Mathematik Technische Universität Wien, Wiedner Hauptstraße 8–10, A–1040 Wien e-mail: kuich@tuwien.ac.at

(i) A is partially ordered by the relation $\sqsubseteq$: $a \sqsubseteq b$ iff there exists a c such that $a + c = b$,

(ii) $\langle A, +, \cdot, 0, 1\rangle$ is a complete semiring,

(iii) $\sum_{i \in I} a_i = \sup(\sum_{i \in E} a_i \mid E \subseteq I,\ E \text{ finite})$, $a_i \in A$, $i \in I$, for an arbitrary index set I, where sup denotes the least upper bound with respect to $\sqsubseteq$.

In the sequel, we denote $\langle A, +, \cdot, 0, 1\rangle$ briefly by A.

Furthermore, $\Sigma = \Sigma_0 \cup \Sigma_1 \cup \ldots \cup \Sigma_k \cup \ldots$ will always denote a *ranked alphabet*, where Σ_k, $k \geq 0$, contains the symbols of rank k and X will denote an *alphabet of leaf symbols*. By $T_\Sigma(X)$ we denote the set of *trees* formed by $\Sigma \cup X$. This set $T_\Sigma(X)$ is the smallest set formed according to the following conventions:

(i) if $\omega \in \Sigma_0 \cup X$ then $\omega \in T_\Sigma(X)$,

(ii) if $\omega \in \Sigma_k$, $k \geq 1$, and $t_1, \ldots, t_k \in T_\Sigma(X)$ then $\omega(t_1, \ldots, t_k) \in T_\Sigma(X)$.

If $\Sigma_0 \neq \emptyset$ then X may be the empty set.

By $A\langle\langle T_\Sigma(X)\rangle\rangle$ we denote the set of *formal tree series* over $T_\Sigma(X)$, i. e., the set of mappings $s : T_\Sigma(X) \to A$ written in the form $\sum_{t \in T_\Sigma(X)} (s, t)t$, where the coefficient (s, t) is the value of s for the tree $t \in T_\Sigma(X)$. For a formal tree series $s \in A\langle\langle T_\Sigma(X)\rangle\rangle$, we define the *support* of s, $\text{supp}(s) = \{t \in T_\Sigma(X) \mid (s, t) \neq 0\}$. By $A\langle T_\Sigma(X)\rangle$ we denote the set of tree series in $A\langle\langle T_\Sigma(X)\rangle\rangle$ that have finite support. A power series with finite support is termed *polynomial*. (For more definitions see Kuich [7].)

Formal tree series induce continuous mappings called *substitutions* as follows. Let Y denote a set of variables, where $Y \cap (\Sigma \cup X) = \emptyset$ ($\emptyset$ denotes the empty set), and consider a mapping $h : Y \to A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$. This mapping can be extended to a mapping $h : T_\Sigma(X \cup Y) \to A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$ by $h(x) = x$, $x \in X$, and

$$h(\omega(t_1, \ldots, t_k)) = \bar{\omega}(h(t_1), \ldots, h(t_k)) = \\ \sum_{t'_1, \ldots, t'_k \in T_\Sigma(X \cup Y)} (h(t_1), t'_1) \ldots (h(t_k), t'_k)\omega(t'_1, \ldots, t'_k)\,,$$

for $\omega \in \Sigma_k$ and $t_1, \ldots, t_k \in T_\Sigma(X \cup Y)$, $k \geq 0$. One more extension of h yields a mapping $h : A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle \to A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$ by defining $h(s) = \sum_{t \in T_\Sigma(X \cup Y)} (s, t)h(t)$. This last extension of h is a complete semiring morphism from $A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$ into $A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$. It is a continuous mapping (see Corollary 2.15 of Kuich [7]).

Let now $s \in A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$. Then, by definition, the formal tree series s *induces a mapping* $s : (A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle)^Y \to A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$ as follows: given $h : Y \to A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$, the value of s with argument h is simply $h(s)$, where h is the extended mapping. If $Y = \{y_1, \ldots, y_n\}$ is finite, we use the following notation: $h : Y \to A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$, where $h(y_i) = s_i$, $1 \leq i \leq n$, is denoted by $(s_i,\ 1 \leq i \leq n)$ or $(s_1, \ldots, s_n)$ and the value of s with argument h is denoted by $s(s_i,\ 1 \leq i \leq n)$ or $s(s_1, \ldots, s_n)$. Intuitively, this is simply the *substitution* of the formal tree series $s_i \in A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$ into the variables y_i,

$1 \le i \le n$, of $s \in A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$. The mapping $s : (A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle)^Y \to A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$, i. e., the substitution of formal tree series into the variables of Y, is a continuous mapping (see Theorem 2.18 of Kuich [7]). Observe that $s(s_1, \ldots, s_n) = \sum_{t \in T_\Sigma(X \cup Y)} (s,t) t(s_1, \ldots, s_n)$.

In certain situations, formulae are easier readable if we use the notation $s[s_i/y_i,\ 1 \le i \le n]$ for the substitution of the formal tree series s_i into the variables y_i, $1 \le i \le n$, of s instead of the notation $s(s_i,\ 1 \le i \le n)$. So we will use sometimes this notation $s[s_i/y_i,\ 1 \le i \le n]$.

In the same way, $s \in A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle$ also induces a mapping $s : (A\langle\langle T_\Sigma(X)\rangle\rangle)^Y \to A\langle\langle T_\Sigma(X)\rangle\rangle$.

Our tree automata and tree transducers will be defined by transition matrices. Let $Y_k = \{y_1, \ldots, y_k\}$, $k \ge 1$, and Y be sets of variables. A *matrix* $M \in (A\langle\langle T_\Sigma(X \cup Y_k)\rangle\rangle)^{I' \times I^k}$, $k \ge 1$, I' and I arbitrary index sets, *induces a mapping*

$$M : (A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle)^{I \times 1} \times \ldots \times (A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle)^{I \times 1} \to (A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle)^{I' \times 1}$$

(there are k argument vectors), defined by the entries of the resulting vector as follows: For $P_1, \ldots, P_k \in (A\langle\langle T_\Sigma(X \cup Y)\rangle\rangle)^{I \times 1}$ we define, for all $i \in I'$,

$$M(P_1, \ldots, P_k)_i = \sum_{i_1, \ldots, i_k \in I} M_{i,(i_1,\ldots,i_k)}((P_1)_{i_1}, \ldots, (P_k)_{i_k}) = \sum_{i_1, \ldots, i_k \in I} \sum_{t \in T_\Sigma(X \cup Y_k)} (M_{i,(i_1,\ldots,i_k)}, t) t((P_1)_{i_1}, \ldots, (P_k)_{i_k}).$$

Throughout the whole paper, I (resp. Q) will denote an arbitrary (resp. a finite) index set.

2 Tree transducers

In this section we introduce tree transducers based on formal tree series and matrices. We show that these tree transducers are a generalization of a restricted type of root-to-frontier tree transducers as described in Gécseg, Steinby [4, 5].

In the sequel, Σ and Σ' denote finite ranked alphabets, X and X' denote leaf alphabets and $Z = \{z_i \mid i \ge 1\}$ denotes an alphabet of variables. We denote $Z_k = \{z_i \mid 1 \le i \le k\}$, $k \ge 1$, and $Z_0 = \emptyset$.

A *tree representation* (with state set Q) is a mapping μ from $\Sigma \cup X$ into matrices with entries in $A\langle\langle T_{\Sigma'}(X' \cup Z)\rangle\rangle$ such that

$$\mu : \Sigma_k \to (A\langle\langle T_{\Sigma'}(X' \cup Z_k)\rangle\rangle)^{Q \times Q^k},\ k \ge 1,$$
$$\mu : \Sigma_0 \cup X \to (A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q \times 1}.$$

For $f \in \Sigma_k$, $k \ge 1$, $\mu(f)$ *induces a mapping*

$$\mu(f) : (A\langle\langle T_{\Sigma'}(X' \cup Z)\rangle\rangle)^{Q \times 1} \times \cdots \times (A\langle\langle T_{\Sigma'}(X' \cup Z)\rangle\rangle)^{Q \times 1} \to (A\langle\langle T_{\Sigma'}(X' \cup Z)\rangle\rangle)^{Q \times 1}$$

(there are k argument vectors), defined by the entries of the resulting vector as follows: For $P_1, \ldots, P_k \in (A\langle\langle T_{\Sigma'}(X' \cup Z)\rangle\rangle)^{Q \times 1}$ and $q \in Q$, the mapping is given by

$$\mu(f)(P_1, \ldots, P_k)_q = \sum_{q_1,\ldots,q_k \in Q} \mu(f)_{q,(q_1,\ldots,q_k)}((P_1)_{q_1}, \ldots, (P_k)_{q_k}) .$$

Observe that for $P_1, \ldots, P_k \in (A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1}$, the vector $\mu(f)(P_1, \ldots, P_k)$ is again in $(A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1}$. This means that

$$\langle (A\langle\langle T_{\Sigma'}(X' \cup Z)\rangle\rangle)^{Q\times 1}, (\mu(f) \mid f \in \Sigma)\rangle \text{ and } \langle (A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1}, (\mu(f) \mid f \in \Sigma)\rangle$$

are Σ-algebras. Hence, the mapping $\mu : X \to (A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1}$ can be uniquely extended to a morphism

$$\mu : T_\Sigma(X) \to (A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1} .$$

This morphic extension is defined inductively as follows:

$$\mu(f(t_1, \ldots, t_k)) = \mu(f)(\mu(t_1), \ldots, \mu(t_k))$$

for $f \in \Sigma_k$, $t_1, \ldots, t_k \in T_\Sigma(X)$.

A tree representation μ is called *polynomial* iff $\mu(f) \in (A\langle T_{\Sigma'}(X' \cup Z_k)\rangle)^{Q\times Q^k}$ for $f \in \Sigma_k$, $k \geq 1$, and $\mu(f) \in (A\langle T_{\Sigma'}(X')\rangle)^{Q\times 1}$ for $f \in \Sigma_0 \cup X$. Observe that, for $|Q| = 1$ and $A = \mathbb{B}$, our polynomial tree representations are nothing else than tree homomorphisms (see Gécseg, Steinby [5], page 18).

For $s \in A\langle\langle T_\Sigma(X)\rangle\rangle$ we define $\mu(s) = \sum_{t\in T_\Sigma(X)} (s,t) \otimes \mu(t)$, where $\otimes$ denotes the Kronecker product (see Kuich, Salomaa [9], Section 4). We are now in the position to define the notion of a tree transducer.

A *tree transducer* (with input alphabet Σ, input leaf alphabet X, output alphabet Σ', output leaf alphabet X')

$$\mathfrak{T} = (Q, \mu, S)$$

is given by

(i) a nonempty finite set Q of *states*,

(ii) a *tree representation* μ with state set Q,

(iii) $S \in (A\langle\langle T_{\Sigma'}(X' \cup Z_1)\rangle\rangle)^{1\times Q}$, called the *initial state vector*.

The *mapping* $||\mathfrak{T}|| : A\langle\langle T_\Sigma(X)\rangle\rangle \to A\langle\langle T_{\Sigma'}(X')\rangle\rangle$ *realized by a tree transducer* $\mathfrak{T} = (Q, \mu, S)$ is defined by

$$||\mathfrak{T}||(s) = S(\mu(s)) = S(\sum_{t\in T_\Sigma(X)} (s,t) \otimes \mu(t)) .$$

A tree transducer $\mathfrak{T} = (Q, \mu, S)$ is called *polynomial* iff μ is a polynomial tree representation, and the entries of S are of the form $S_q = a_q z_1$, $a_q \in A$, $q \in Q$.

We now connect our notion of tree transducer with the root-to-frontier tree transducers. By Gécseg, Steinby [5], a *root-to-frontier tree transducer*

$$\mathfrak{A} = (\Sigma, X, Q, \Sigma', X', P, Q')$$

is a system, where

(1) $\Sigma, \Sigma', X, X', Q$ are specified in the same way as in the definition of our tree transducer;

(2) P is a finite set of *productions* of the following types:
(i) $qx \to t$, where $q \in Q$, $x \in X$, $t \in T_{\Sigma'}(X')$;
(ii) $qf(z_1, \ldots, z_k) \to t$, where $q \in Q$, $f \in \Sigma_k$, $k \geq 0$, $t \in T_{\Sigma'}(X' \cup QZ_k)$;

(3) $Q' \subseteq Q$ is the set of *initial states.*

A root-to-frontier tree transducer $\mathfrak{A}$ is called *nondeterministically simple* iff for each production of type (ii) $qf(z_1, \ldots, z_k) \to t$ there exists a set $C_{qf\to t} = \{q_{i_1}z_1, \ldots, q_{i_k}z_k\}$ such that $t \in T_{\Sigma'}(X' \cup C_{qf\to t})$. (Compare with Gécseg, Steinby) [4], Exercise 4 on page 213.) Observe that not all elements $q_{i_1}z_1, \ldots, q_{i_k}z_k$ of $C_{qf\to t}$ have to appear in t, i. e., $\mathfrak{A}$ needs not to be nondeleting.

For the forthcoming considerations in this section, our basic semiring is the Boolean semiring $\mathbb{B}$ and we use without mentioning the isomorphism between $\mathbb{B}\langle\langle T_\Sigma(X)\rangle\rangle$ and $\mathfrak{P}(T_\Sigma(X))$. Given a nondeterministically simple root-to-frontier tree transducer $\mathfrak{A} = (\Sigma, X, Q, \Sigma', X', P, Q')$, we construct a polynomial tree transducer $\mathfrak{T} = (Q, \mu, S)$ that behaves analogous to $\mathfrak{A}$. The polynomial tree representation μ is defined as follows:

(i) For $x \in X$, $q \in Q$, $t \in T_{\Sigma'}(X')$, $(\mu(x)_q, t) = 1$ if $qx \to t \in P$.

(ii) For $f \in \Sigma_k$, $q, q_{i_1}, \ldots, q_{i_k} \in Q$, $t(z_1, \ldots, z_k) \in T_{\Sigma'}(X' \cup Z_k)$, $k \geq 0$, $(\mu(f)_{q,(q_{i_1},\ldots,q_{i_k})}, t(z_1, \ldots, z_k)) = 1$ if $qf(z_1, \ldots, z_k) \to t(q_{i_1}z_1, \ldots, q_{i_k}z_k) \in P$ and $C_{qf\to t} = \{q_{i_1}z_1, \ldots, q_{i_k}z_k\}$.

The initial state vector S is defined by $S_q = z_1$ if $q \in Q'$, $S_q = 0$ if $q \in Q - Q'$.

We claim that, for $s \in T_\Sigma(X)$, $t \in T_{\Sigma'}(X')$ and $q \in Q$

$$(\mu(s)_q, t) = 1 \text{ iff } qs \Rightarrow^* t$$

and prove it by induction on the form of trees in $T_\Sigma(X)$. Clearly, the claim holds true for trees in $X \cup \Sigma_0$. Let now $f \in \Sigma_k$ for some $k \geq 1$, $s_1, \ldots, s_k \in T_\Sigma(X)$, and $t = f(s_1, \ldots, s_k)$. By induction hypothesis, we have $(\mu(s_j)_{q_j}, t_j) = 1$ iff $q_js_j \Rightarrow^* t_j$, $q_j \in Q$, $t_j \in T_{\Sigma'}(X')$, $1 \leq j \leq k$. Let now $(\mu(f(s_1, \ldots, s_k))_q, t) = 1$, i. e., for some $q_1, \ldots, q_k \in Q$

$$(\mu(f)_{q,(q_1,\ldots,q_k)}(\mu(s_1)_{q_1}, \ldots, \mu(s_k)_{q_k}), t) = 1\,.$$

Then there exist $v \in T_{\Sigma'}(X' \cup Z_k)$ and $t_1, \dots, t_k \in T_{\Sigma'}(X')$ such that $(\mu(f)_{q,(q_1,\dots,q_k)}, v) = 1$, $(\mu(s_j)_{q_j}, t_j) = 1$, $1 \le j \le k$, and $t = v(t_1, \dots, t_k)$. This implies

$$qf(s_1, \dots, s_k) \Rightarrow^* v(q_1 s_1, \dots, q_k s_k) \Rightarrow^* v(t_1, \dots, t_k) = t\,.$$

Similarly, we can show that $qf(s_1, \dots, s_k) \Rightarrow^* t$ implies $(\mu(f(s_1, \dots, s_k))_q, t) = 1$. This yields our first theorem.

Theorem 1 *Let $\mathfrak{A}$ be a nondeterministically simple root-to-frontier tree transducer and $\mathfrak{T}$ be the polynomial tree transducer constructed from $\mathfrak{A}$. Let $L \subseteq T_\Sigma(X)$ be a tree language. Then $\mathfrak{A}$ maps L to the tree language $supp(||\mathfrak{T}||(char(L))) \subseteq T_{\Sigma'}(X')$.*

3 Linear and nondeleting tree transducers

In this section we introduce linear and nondeleting tree representations and tree transducers. The main result of this section is that a linear nondeleting representation can be extended to a morphism over matrices of formal tree series.

A tree $t \in T_\Sigma(X \cup Z_k)$, $k \ge 1$, is called *linear* iff the variable z_j appears at most once in t, $1 \le j \le k$. A tree $t \in T_\Sigma(X \cup Z_k)$, $k \ge 1$, is called *nondeleting* iff the variable z_j appears at least once in t, $1 \le j \le k$. A tree series $s \in A\langle\langle T_\Sigma(X \cup Z_k)\rangle\rangle$, $k \ge 1$, is called *linear* or *nondeleting* iff all $t \in \text{supp}(s)$ are linear or nondeleting, respectively. A tree representation μ is called *linear* or *nondeleting* iff all entries of $\mu(f)$, $f \in \Sigma_k$, $k \ge 1$, are linear or nondeleting tree series, respectively. A tree transducer $\mathfrak{T} = (Q, \mu, S)$ is called *linear* or *nondeleting* iff μ is linear or nondeleting, respectively, and the entries of S are of the form $S_q = a_q z_1$, $a_q \in A$, $q \in Q$.

Before we can state and prove our main result of this section we need a series of technical lemmas.

Lemma 2 *Let, for some $k \ge 1$, $t \in T_\Sigma(X \cup Z_k)$, and $s_j \in A\langle\langle T_\Sigma(X)\rangle\rangle$, $a_j \in A$, $1 \le j \le k$. Assume that the variable z_j appears $m_j \ge 0$ times in t, $1 \le j \le k$. Then*

$$t(a_1 s_1, \dots, a_k s_k) = a_1^{m_1} \dots a_k^{m_k} t(s_1, \dots, s_k)\,.$$

Proof. The proof is by induction on the form of trees in $T_\Sigma(X \cup Z_k)$. The lemma is trivial for $t = x$ or $t = z_j$, $1 \le j \le k$. Let now t be of the form $f(t_1, \dots, t_m)$, where $f \in \Sigma_m$ for some $m \ge 1$, and $t_1, \dots, t_m \in T_\Sigma(X \cup Z_k)$. Let z_j appear u_{ij} times in t_i, $1 \le j \le k$, $1 \le i \le m$. Then we have by induction hypothesis $t_i(a_1 s_1, \dots, a_k s_k) = a_1^{u_{i1}} \dots a_k^{u_{ik}} t_i(s_1, \dots, s_k)$, $1 \le i \le m$. Hence, $t(a_1 s_1, \dots, a_k s_k) = f(t_1(a_1 s_1, \dots, a_k s_k), \dots, t_m(a_1 s_1, \dots, a_k s_k)) = f(a_1^{u_{11}} \dots a_k^{u_{1k}} t_1(s_1, \dots, s_k), \dots, a_1^{u_{m1}} \dots a_k^{u_{mk}} t_m(s_1, \dots, s_k)) = a_1^{u_{11}+\dots+u_{m1}} \dots a_k^{u_{1k}+\dots+u_{mk}} f(t_1(s_1, \dots, s_k), \dots, t_m(s_1, \dots, s_k)) = a_1^{u_{11}+\dots+u_{m1}} \dots a_k^{u_{1k}+\dots+u_{mk}} t(s_1, \dots, s_k)$. □

Lemma 3 *Let, for some $k \geq 1$, $s \in A\langle\langle T_\Sigma(X \cup Z_k)\rangle\rangle$ be linear and nondeleting, and $s_j \in A\langle\langle T_\Sigma(X)\rangle\rangle$, $a_j \in A$ for $1 \leq j \leq k$. Then*

$$s(a_1 s_1, \ldots, a_k s_k) = a_1 \ldots a_k s(s_1, \ldots, s_k)\,.$$

Proof. We have $s = \sum_{t \in T_\Sigma(X \cup Z_k)} (s,t)t$. Since s is linear and nondeleting, we obtain, by Lemma 2, $t(a_1 s_1, \ldots, a_k s_k) = a_1 \ldots a_k t(s_1, \ldots, s_k)$ for all $t \in T_\Sigma(X \cup Z_k)$ such that $(s,t) \neq 0$. Hence $s(a_1 s_1, \ldots, a_k s_k) = \sum_{t \in T_\Sigma(X \cup Z_k)} a_1 \ldots a_k (s,t)$ $t(s_1, \ldots, s_k) = a_1 \ldots a_k s(s_1, \ldots, s_k)$. □

Lemma 4 *Assume that the variable z_1 appears exactly once in $t \in T_\Sigma(X \cup Z_k)$, $k \geq 1$. Let $s_i \in A\langle\langle T_\Sigma(X)\rangle\rangle$ for $i \in I$ and $r_2, \ldots, r_k \in A\langle\langle T_\Sigma(X)\rangle\rangle$. Then*

$$t(\sum_{i \in I} s_i, r_2, \ldots, r_k) = \sum_{i \in I} t(s_i, r_2, \ldots, r_k)\,.$$

Proof. The proof is by induction on the form of trees in $T_\Sigma(X \cup Z_k)$. The lemma is trivial for $t = z_1$. Let now t be of the form $f(t_1, \ldots, t_m)$, where $f \in \Sigma_m$, and $t_1, \ldots, t_m \in T_\Sigma(X \cup Z_k)$. The variable z_1 appears in exactly one of the subtrees t_j, $1 \leq j \leq k$, say in t_u. By induction hypothesis we have $t_u(\sum_{i \in I} s_i, r_2, \ldots, r_k) = \sum_{i \in I} t_u(s_i, r_2, \ldots, r_k)$. Hence,

$$\begin{aligned} &f(t_1(\textstyle\sum_{i \in I} s_i, r_2, \ldots, r_k), \ldots, t_u(\sum_{i \in I} s_i, r_2, \ldots, r_k), \ldots, \\ &\qquad\qquad t_m(\textstyle\sum_{i \in I} s_i, r_2, \ldots, r_k)) = \\ &f(t_1(r_1, r_2, \ldots, r_k), \ldots, \textstyle\sum_{i \in I} t_u(s_i, r_2, \ldots, r_k), \ldots, t_m(r_1, r_2, \ldots, r_k)) = \\ &\textstyle\sum_{i \in I} f(t_1(r_1, r_2, \ldots, r_k), \ldots, t_u(s_i, r_2, \ldots, r_k), \ldots, t_m(r_1, r_2, \ldots, r_k)) \end{aligned}$$

for all $r_1 \in A\langle\langle T_\Sigma(X)\rangle\rangle$. Hence, the last sum is equal to

$$\begin{aligned} &\textstyle\sum_{i \in I} f(t_1(s_i, r_2, \ldots, r_k), \ldots, t_u(s_i, r_2, \ldots, r_k), \ldots, t_m(s_i, r_2, \ldots, r_k)) = \\ &\textstyle\sum_{i \in I} (f(t_1, \ldots, t_u, \ldots, t_m))(s_i, r_2, \ldots, r_k) = \sum_{i \in I} t(s_i, r_2, \ldots, r_k)\,. \end{aligned}$$

□

Lemma 5 *Let, for some $k \geq 1$, $s \in A\langle\langle T_\Sigma(X \cup Z_k)\rangle\rangle$ be linear and nondeleting, and $s_i \in A\langle\langle T_\Sigma(X)\rangle\rangle$, for $i \in I$. Moreover, let $r_2, \ldots, r_k \in A\langle\langle T_\Sigma(X)\rangle\rangle$. Then*

$$s(\sum_{i \in I} s_i, r_2, \ldots, r_k) = \sum_{i \in I} s(s_i, r_2, \ldots, r_k)\,.$$

Proof. We have, by Lemma 4, $s(\sum_{i \in I} s_i, r_2, \ldots, r_k) = \sum_{t \in T_\Sigma(X \cup Z_k)} (s,t) t(\sum_{i \in I} s_i, r_2, \ldots, r_k) = \sum_{i \in I} \sum_{t \in T_\Sigma(X \cup Z_k)} (s,t) t(s_i, r_2, \ldots, r_k) = \sum_{i \in I} s(s_i, r_2, \ldots, r_k)$. □

Clearly, Lemma 5 also holds for argument places different from one.

Theorem 6 *Let, for some $k \geq 1$, $s \in A\langle\langle T_\Sigma(X \cup Z_k)\rangle\rangle$ be linear and nondeleting, and $s_{i_j} \in A\langle\langle T_\Sigma(X)\rangle\rangle$, $a_{i_j} \in A$ for $i_j \in I_j$, $1 \leq j \leq k$. Then*

$$s(\sum_{i_1 \in I_1} a_{i_1} s_{i_1}, \ldots, \sum_{i_k \in I_k} a_{i_k} s_{i_k}) = \sum_{i_1 \in I_1} \cdots \sum_{i_k \in I_k} a_{i_1} \ldots a_{i_k} s(s_{i_1}, \ldots, s_{i_k})\,.$$

Proof. By Lemmas 3 and 5. □

Given a tree representation μ, we now extend μ to mappings

$$\mu : (A\langle\langle\Sigma_k\rangle\rangle)^{I\times I^k} \to ((A\langle\langle T_{\Sigma'}(X' \cup Z_k)\rangle\rangle)^{Q\times Q^k})^{I\times I^k} \text{ for } k \geq 0\,,$$

and

$$\mu : (A\langle\langle T_\Sigma(X)\rangle\rangle)^{I\times 1} \to ((A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1})^{I\times 1}$$

by

$$\mu(M) = \sum_{f\in\Sigma_k} (M,f) \otimes \mu(f), \quad M \in (A\langle\langle\Sigma_k\rangle\rangle)^{I\times I^k},\ k \geq 0\,,$$

and

$$\mu(P) = \sum_{t\in T_\Sigma(X)} (P,t) \otimes \mu(t), \quad P \in (A\langle\langle T_\Sigma(X)\rangle\rangle)^{I\times 1}\,.$$

Observe that $\mu(M)$, $M \in (A\langle\langle\Sigma_k\rangle\rangle)^{I\times I^k}$, $k \geq 0$, induces a mapping $\mu(M) : ((A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1})^{I\times 1} \times \ldots \times ((A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1})^{I\times 1} \to ((A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1})^{I\times 1}$ (there are k argument vectors). The next theorem implies that a linear nondeleting tree representation μ is a morphism from the Σ-algebra

$$\langle (A\langle\langle T_\Sigma(X)\rangle\rangle)^{I\times 1}, (M_f \mid f \in \Sigma)\rangle, \text{ for some } M_f \in (A\langle\langle\Sigma_k\rangle\rangle)^{I\times I^k},\ f\in\Sigma_k,\ k\geq 0\,,$$

into the Σ-algebra

$$\langle ((A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1})^{I\times 1}, (\mu(M_f) \mid f \in \Sigma)\rangle\,.$$

Theorem 7 *Let $M \in (A\langle\langle\Sigma_k\rangle\rangle)^{I\times I^k}$, $P_1,\ldots,P_k \in (A\langle\langle T_\Sigma(X)\rangle\rangle)^{I\times 1}$ for some $k \geq 1$, and μ be a linear nondeleting tree representation with state set Q. Then*

$$\mu(M)(\mu(P_1),\ldots,\mu(P_k)) = \mu(M(P_1,\ldots,P_k))\,.$$

Proof. We first compute the left side of the equality of the theorem for indices $i \in I$ and $q \in Q$:

$$\begin{aligned}
&(\mu(M)(\mu(P_1),\ldots,\mu(P_k))_i)_q = \\
&\textstyle\sum_{i_1,\ldots,i_k\in I}\sum_{q_1,\ldots,q_k\in Q}(\mu(M)_{i,(i_1,\ldots,i_k)})_{q,(q_1,\ldots,q_k)} \\
&\qquad\qquad ((\mu(P_1)_{i_1})_{q_1},\ldots,(\mu(P_k)_{i_k})_{q_k}) = \\
&\textstyle\sum_{i_1,\ldots,i_k\in I}\sum_{q_1,\ldots,q_k\in Q}((\sum_{f\in\Sigma_k}(M,f)\otimes\mu(f))_{i,(i_1,\ldots,i_k)})_{q,(q_1,\ldots,q_k)} \\
&\textstyle(((\sum_{t_1\in T_\Sigma(X)}(P_1,t_1)\otimes\mu(t_1))_{i_1})_{q_1},\ldots,((\sum_{t_k\in T_\Sigma(X)}(P_k,t_k)\otimes\mu(t_k))_{i_k})_{q_k}) = \\
&\textstyle\sum_{i_1,\ldots,i_k\in I}\sum_{q_1,\ldots,q_k\in Q}\sum_{f\in\Sigma_k}(M,f)_{i,(i_1,\ldots,i_k)}\mu(f)_{q,(q_1,\ldots,q_k)} \\
&\qquad\textstyle(\sum_{t_1\in T_\Sigma(X)}(P_1,t_1)_{i_1}\mu(t_1)_{q_1},\ldots,\sum_{t_k\in T_\Sigma(X)}(P_k,t_k)_{i_k}\mu(t_k)_{q_k}) = \\
&\textstyle\sum_{i_1,\ldots,i_k\in I}\sum_{q_1,\ldots,q_k\in Q}\sum_{f\in\Sigma_k}\sum_{t_1,\ldots,t_k\in T_\Sigma(X)}(M,f)_{i,(i_1,\ldots,i_k)} \\
&\qquad\qquad (P_1,t_1)_{i_1}\ldots(P_k,t_k)_{i_k}\mu(f)_{q,(q_1,\ldots,q_k)}(\mu(t_1)_{q_1},\ldots,\mu(t_k)_{q_k})\,.
\end{aligned}$$

Here the last equality follows by Theorem 6. We now compute the right side of the equality of the theorem for indices $i \in I$ and $q \in Q$:

$$
\begin{aligned}
&(\mu(M(P_1, \ldots, P_k))_i)_q = \\
&\sum_{t \in T_\Sigma(X)} (M(P_1, \ldots, P_k), t)_i \mu(t)_q = \\
&\sum_{t \in T_\Sigma(X)} (\sum_{i_1, \ldots, i_k \in I} (M_{i,(i_1,\ldots,i_k)}((P_1)_{i_1}, \ldots, (P_k)_{i_k}), t)) \mu(t)_q = \\
&\sum_{t \in T_\Sigma(X)} \sum_{i_1, \ldots, i_k \in I} \sum_{f \in \Sigma_k} ((M_{i,(i_1,\ldots,i_k)}, f) f((P_1)_{i_1}, \ldots, (P_k)_{i_k}), t) \mu(t)_q = \\
&\sum_{t_1, \ldots, t_k \in T_\Sigma(X)} \sum_{i_1, \ldots, i_k \in I} \sum_{f \in \Sigma_k} (M_{i,(i_1,\ldots,i_k)}, f) \\
&\qquad\qquad ((P_1)_{i_1}, t_1) \ldots ((P_k)_{i_k}, t_k) \mu(f(t_1, \ldots, t_k))_q = \\
&\sum_{t_1, \ldots, t_k \in T_\Sigma(X)} \sum_{i_1, \ldots, i_k \in I} \sum_{f \in \Sigma_k} \sum_{q_1, \ldots, q_k \in Q} (M_{i,(i_1,\ldots,i_k)}, f) \\
&\qquad ((P_1)_{i_1}, t_1) \ldots ((P_k)_{i_k}, t_k) \mu(f)_{q,(q_1,\ldots,q_k)} (\mu(t_1)_{q_1}, \ldots, \mu(t_k)_{q_k}) .
\end{aligned}
$$

Here the fourth equality follows by the fact that $(f((P_1)_{i_1}, \ldots, (P_k)_{i_k}), t)$ is unequal to 0 only if t is of the form $f(t_1, \ldots, t_k)$.

Since both sides of the equation of our theorem coincide, the theorem is proven. □

4 Recognizable tree transducers and recognizable tree series

It is easy to see that our tree transducers do not preserve the recognizability of tree series. (See the example in the last paragraph of page 18 of Gecseg, Steinby [5].) On the other hand, linear root-to-frontier tree transducers do preserve recognizability of tree languages. (See Thatcher [10]; and Gécseg, Steinby [4], Theorem 2.7, Lemma 6.5 and Corollary 6.6.) In this section we show that linear *nondeleting* recognizable tree transducers do preserve recognizability of tree series. We show this by two different constructions: one is based on finite linear systems, the other is based on finite tree automata.

We start with the construction based on finite linear systems.

A *finite linear system* (see Berstel, Reutenauer [1], Bozapalidis [2, 3], Kuich [7, 8]) is a system of formal equations $z_i = p_i$, $1 \leq i \leq n$, for some $n \geq 1$, where each p_i is in $A\langle\langle T_\Sigma(X \cup Z_n)\rangle\rangle$. A *solution* to the finite linear system $z_i = p_i$, $1 \leq i \leq n$, is given by $\sigma \in (A\langle\langle T_\Sigma(X)\rangle\rangle)^{n \times 1}$ such that $\sigma_i = p_i(\sigma_1, \ldots, \sigma_n)$, $1 \leq i \leq n$. A solution σ of $z_i = p_i$, $1 \leq i \leq n$, is termed *least solution* iff $\sigma \sqsubseteq \tau$ for all solutions τ of $z_i = p_i$, $1 \leq i \leq n$. The *approximation sequence* $(\sigma^j \mid j \in \mathbb{N})$, $\sigma^j \in (A\langle\langle T_\Sigma(X)\rangle\rangle)^{n \times 1}$, $j \geq 0$, *associated to the finite linear system* $z_i = p_i$, $1 \leq i \leq n$, is defined as follows:

$$\sigma_i^0 = 0, \qquad \sigma_i^{j+1} = p_i(\sigma_1^j, \ldots, \sigma_n^j),\ 1 \leq i \leq n,\ j \geq 0 .$$

The least upper bound $\sigma = \sup(\sigma^j \mid j \in \mathbb{N})$ of the approximation sequence exists and is the least solution of the finite linear system.

A finite linear system $z_i = p_i$, $1 \leq i \leq n$ is called *proper* iff $(p_i, z_j) = 0$ for all $1 \leq j \leq n$, i. e., iff there do not appear linear terms in p_i.

A finite linear system $z_i = p_i$, $1 \le i \le n$ is called *polynomial* iff each p_i is in $A\langle T_\Sigma(X \cup Z_n)\rangle$. The collection of all the components of least solutions of finite polynomial linear systems is denoted by $A^{\text{rec}}\langle\langle T_\Sigma(X)\rangle\rangle$. The tree series in $A^{\text{rec}}\langle\langle T_\Sigma(X)\rangle\rangle$ are called *recognizable tree series.*

A finite linear system $z_i = p_i$, $1 \le i \le n$ is called *recognizable* iff each p_i is in $A^{\text{rec}}\langle\langle T_\Sigma(X \cup Z_n)\rangle\rangle$.

An adaption of the proof of Proposition 6.1 of Berstel, Reutenauer [1] yields the following result.

Theorem 8 *For each finite (resp. recognizable finite or polynomial finite) linear system there exists a proper finite (resp. proper recognizable finite or proper polynomial finite) linear system with the same least solution. A proper finite linear system has a unique solution.*

We now show that the least solution of a recognizable finite linear system has recognizable components.

Theorem 9 *Let $z_i = p_i$, $1 \le i \le n$, be a recognizable finite linear system with least solution σ. Then $\sigma_i \in A^{\text{rec}}\langle\langle T_\Sigma(X)\rangle\rangle$ for all $1 \le i \le n$.*

Proof. Without loss of generality let $z_i = p_i$, $1 \le i \le n$, be a proper recognizable finite linear system. Since $p_i \in A^{\text{rec}}\langle\langle T_\Sigma(X \cup Z_n)\rangle\rangle$, $1 \le i \le n$, there exist proper polynomial finite linear systems $y_{ij} = q_{ij}$, $1 \le j \le m_i$, $m_i \ge 1$, where the y_{ij} are new variables and $q_{ij} \in A\langle T_\Sigma(X \cup Z_n \cup \{y_{i1}, \dots, y_{im_i}\})\rangle$, $1 \le i \le n$, such that the q_{i1}-components of their least solutions τ_i are equal to p_i. Consider now the polynomial finite linear system $z_i = q_{i1}(z_1, \dots, z_n, y_{i1}, \dots, y_{im_i})$, $y_{ij} = q_{ij}(z_1, \dots, z_n, y_{i1}, \dots, y_{im_i})$, $1 \le j \le m_i$, $1 \le i \le n$, and observe that this polynomial finite linear system has a unique solution. We claim that this unique solution is given by $\sigma \cup ((\tau_i)_j(\sigma_1, \dots, \sigma_n) \mid 1 \le j \le m_i,\ 1 \le i \le n)$. Substitution of this vector yields, for $1 \le j \le m_i$, $1 \le i \le n$,

$$\begin{aligned} q_{i1}(\sigma_1, \dots, \sigma_n, (\tau_i)_1(\sigma_1, \dots, \sigma_n), \dots, (\tau_i)_{m_i}(\sigma_1, \dots, \sigma_n)) &= \\ (\tau_i)_1(\sigma_1, \dots, \sigma_n) = p_i(\sigma_1, \dots, \sigma_n) &= \sigma_i\,, \\ q_{ij}(\sigma_1, \dots, \sigma_n, (\tau_i)_1(\sigma_1, \dots, \sigma_n), \dots, (\tau_i)_{m_i}(\sigma_1, \dots, \sigma_n)) &= (\tau_i)_j(\sigma_1, \dots, \sigma_n)\,. \end{aligned}$$

Hence $\sigma \cup ((\tau_i)_j(\sigma_1, \dots, \sigma_n) \mid 1 \le j \le m_i,\ 1 \le i \le n)$ is the unique solution of the polynomial finite linear system and $\sigma \in (A^{\text{rec}}\langle\langle T_\Sigma(X)\rangle\rangle)^{n \times 1}$. □

Let $Y = \{y_i \mid i \ge 1\}$ be an alphabet of new variables and denote $Y_k = \{y_1, \dots, y_k\}$, $k \ge 1$, $Y_0 = \emptyset$. Let $s(y_1, \dots, y_k) \in A^{\text{rec}}\langle\langle T_\Sigma(X \cup Y_k)\rangle\rangle$ and $\tau_j \in A^{\text{rec}}\langle\langle T_\Sigma(X \cup Y_k)\rangle\rangle$, $1 \le j \le k$. Then, by Bozapalidis [2], $s(\tau_1, \dots, \tau_n)$ is again in $A^{\text{rec}}\langle\langle T_\Sigma(X \cup Y_k)\rangle\rangle$, i. e., $A^{\text{rec}}\langle\langle T_\Sigma(X \cup Y_k)\rangle\rangle$ is closed under substitution.

Theorem 10 $A^{\text{rec}}\langle\langle T_\Sigma(X \cup Y_k)\rangle\rangle$, $k \ge 1$, *is closed under substitution.*

Consider a finite linear system $y_i = p_i(y_1, \dots, y_n)$, $1 \le i \le n$, where $p_i \in A\langle\langle T_\Sigma(X \cup Y_n)\rangle\rangle$, and a tree representation μ with state set Q, where $\mu : \Sigma_k \to (A\langle\langle T_{\Sigma'}(X' \cup Z_k)\rangle\rangle)^{Q \times Q^k}$, $k \ge 1$, and $\mu : \Sigma_0 \cup X \to (A\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q \times 1}$. Let $(y_i)_q$,

$1 \leq i \leq n$, $q \in Q$, be new variables and denote $Y_Q^k = \{(y_i)_q \mid 1 \leq i \leq k,\ q \in Q\}$. Extend the definition of μ to the domain $\Sigma \cup X \cup Y_n$, by

$$\mu : Y_n \to (A\langle\langle T_{\Sigma'}(Y_Q^n)\rangle\rangle)^{Q\times 1},$$

where $\mu(y_j)_q = (y_j)_q$, $1 \leq j \leq n$, $q \in Q$. By this extension, we obtain that

$$\mu : T_\Sigma(X \cup Y_n) \to (A\langle\langle T_{\Sigma'}(X' \cup Y_Q^n)\rangle\rangle)^{Q\times 1}.$$

Lemma 11 *Consider $s(y_1, \ldots, y_n) \in A\langle\langle T_\Sigma(X \cup Y_n)\rangle\rangle$ and a linear nondeleting tree representation μ with domain $\Sigma \cup X \cup Y_n$. Let $s_1, \ldots, s_n \in A\langle\langle T_\Sigma(X)\rangle\rangle$. Then*

$$\mu(s)[\mu(s_j)_q/(y_j)_q,\ 1 \leq j \leq n,\ q \in Q] = \mu(s(s_1, \ldots, s_n)).$$

Proof. We first consider a tree $t \in T_\Sigma(X \cup Y_n)$ and show by induction on the form of t that $\mu(t)[\mu(s_j)_q/(y_j)_q,\ 1 \leq j \leq n,\ q \in Q] = \mu(t(s_1, \ldots, s_n))$.
(i) For $t = y_i$, $1 \leq i \leq n$, we obtain $\mu(y_i)[\mu(s_j)/\mu(y_j),\ 1 \leq j \leq n] = \mu(s_i) = \mu(y_i(s_1, \ldots, s_n))$.
(ii) For $t = x$, $x \in \Sigma_0 \cup X$, we obtain $\mu(x)[\mu(s_j)/\mu(y_j),\ 1 \leq j \leq n] = \mu(x) = \mu(x(s_1, \ldots, s_n))$.
(iii) For $t = f(t_1, \ldots, t_k)$, $f \in \Sigma_k$, $t_1, \ldots, t_k \in T_\Sigma(X \cup Y_n)$, $k \geq 1$, we obtain

$$\begin{aligned}
&\mu(f(t_1, \ldots, t_k))[\mu(s_j)/\mu(y_j),\ 1 \leq j \leq n] = \\
&\mu(f)(\mu(t_1)[\mu(s_j)/\mu(y_j),\ 1 \leq j \leq n], \ldots, \mu(t_k)[\mu(s_j)/\mu(y_j),\ 1 \leq j \leq n]) = \\
&\mu(f)(\mu(t_1(s_1, \ldots, s_n)), \ldots, \mu(t_k(s_1, \ldots, s_n))) = \\
&\mu(f(t_1(s_1, \ldots, s_n), \ldots, t_k(s_1, \ldots, s_n))) = \\
&\mu((f(t_1, \ldots, t_k))(s_1, \ldots, s_n)).
\end{aligned}$$

Here we have applied the induction hypothesis in the second equality and Theorem 7 in the third equality.

Finally, we obtain

$$\begin{aligned}
&\mu(s)[\mu(s_j)/\mu(y_j),\ 1 \leq j \leq n] = \\
&\textstyle\sum_{t\in T_\Sigma(X\cup Y_n)}(s,t) \otimes \mu(t)[\mu(s_j)/\mu(y_j),\ 1 \leq j \leq n] = \\
&\textstyle\sum_{t\in T_\Sigma(X\cup Y_n)}(s,t) \otimes \mu(t(s_1, \ldots, s_n)) = \\
&\mu(\textstyle\sum_{t\in T_\Sigma(X\cup Y_n)}(s,t)t(s_1, \ldots, s_n)) = \mu(s(s_1, \ldots, s_n)).
\end{aligned}$$

□

Theorem 12 *Consider a linear nondeleting tree representation μ with domain $\Sigma \cup X \cup Y_n$. Let $y_i = p_i(y_1, \ldots, y_n)$, $1 \leq i \leq n$, where $p_i \in A\langle\langle T_\Sigma(X \cup Y_n)\rangle\rangle$, be a finite linear system with least solution σ. Then $\mu(\sigma)$ is the least solution of the finite linear system $\mu(y_i) = \mu(p_i(y_1, \ldots, y_n))$, $1 \leq i \leq n$.*

Proof. Let $(\sigma^j \mid j \in \mathbb{N})$ and $(\tau^j \mid j \in \mathbb{N})$ be the approximation sequences of $y_i = p_i(y_1, \ldots, y_n)$, $1 \leq i \leq n$, and $\mu(y_i) = \mu(p_i(y_1, \ldots, y_n))$, $1 \leq i \leq n$, respectively. We claim that $\tau_i^j = \mu(\sigma_i^j)$, $1 \le i \le n$, $j \ge 0$, and show it by induction on j. The case $j = 0$ is clear. Let $j \ge 0$. Then, for $1 \le i \le n$,

$$\tau_i^{j+1} = \mu(p_i(y_1, \dots, y_n))[\tau_k^j / \mu(y_k),\ 1 \le k \le n] =$$
$$\mu(p_i(y_1, \dots, y_n))[\mu(\sigma_k^j)/\mu(y_k),\ 1 \le k \le n] =$$
$$\mu(p_i(\sigma_1^j, \dots, \sigma_n^j)) = \mu(\sigma_i^{j+1}).$$

Here we have applied the induction hypothesis in the second equality and Lemma 11 in the third equality. The claim now implies our theorem. □

A tree representation μ is called *recognizable* iff $\mu(f) \in (A^{\mathrm{rec}}\langle\langle T_{\Sigma'}(X' \cup Z_k)\rangle\rangle)^{Q \times Q^k}$ for $f \in \Sigma_k$, $k \ge 1$, and $\mu(f) \in (A^{\mathrm{rec}}\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q \times 1}$ for $f \in \Sigma_0 \cup X$. A tree transducer $\mathfrak{T} = (Q, \mu, S)$ is called *recognizable* iff μ is a recognizable tree representation and the entries of S are of the form $S_q = a_q z_1$, $a_q \in A$, $q \in Q$.

Corollary 13 *Consider a linear nondeleting recognizable tree representation μ. Let s be in $A^{\mathrm{rec}}\langle\langle T_\Sigma(X)\rangle\rangle$. Then $\mu(s)$ is in $(A^{\mathrm{rec}}\langle\langle T_{\Sigma'}(X')\rangle\rangle)^{Q\times 1}$.*

Corollary 14 *Consider a linear nondeleting recognizable tree transducer $\mathfrak{T}$ and a recognizable tree series s. Then $||\mathfrak{T}||(s)$ is again recognizable.*

We now turn to the automata-based construction.

Our tree automata are a generalization of the nondeterministic root-to-frontier tree recognizers (see Gécseg, Steinby [4, 5]) and are defined in Kuich [7, 8]. A *tree automaton* (with input alphabet Σ and leaf alphabet X)

$$\mathfrak{A} = (I, M, S, P)$$

is given by

(i) a nonempty set I of *states*,

(ii) a sequence $M = (M_k \mid k \ge 1)$ of *transition matrices* $M_k \in (A\langle\langle T_\Sigma(X \cup Y_k)\rangle\rangle)^{I \times I^k}$, $k \ge 1$,

(iii) $S \in (A\langle\langle T_\Sigma(X \cup Y_1)\rangle\rangle)^{1 \times I}$, called the *initial state vector*,

(iv) $P \in (A\langle\langle T_\Sigma(X)\rangle\rangle)^{I \times 1}$, called the *final state vector*.

The *approximation sequence* $(\sigma^j \mid j \in \mathbb{N})$, $\sigma^j \in (A\langle\langle T_\Sigma(X)\rangle\rangle)^{I\times 1}$, $j \ge 0$, *associated to* $\mathfrak{A}$ is defined as follows:

$$\sigma^0 = 0, \qquad \sigma^{j+1} = \sum_{k \ge 1} M_k(\sigma^j, \dots, \sigma^j) + P, \quad j \ge 0.$$

The *behavior* $||\mathfrak{A}|| \in A\langle\langle T_\Sigma(X)\rangle\rangle$ of the tree automaton $\mathfrak{A}$ is defined by

$$||\mathfrak{A}|| = \sum_{i \in I} S_i(\sigma_i) = S(\sigma),$$

where $\sigma \in (A\langle\langle T_\Sigma(X)\rangle\rangle)^{I\times 1}$ is the least upper bound of the approximation sequence associated to $\mathfrak{A}$.

A tree automaton $\mathfrak{A} = (I, (M_k \mid k \geq 1), S, P)$ is termed *simple* iff the entries of the transition matrices M_k, $k \geq 1$, of the initial state vector S and of the final state vector P have the following specific form:

(i) the entries of M_k, $k \geq 2$, are of the form $\sum_{f\in\Sigma_k} a_f f(y_1, \ldots, y_k)$, $a_f \in A$;

(ii) the entries of M_1 are of the form $\sum_{f\in\Sigma_1} a_f f(y_1) + ay_1$, $a_f, a \in A$;

(iii) the entries of P are of the form $\sum_{\omega\in\Sigma_0\cup X} a_\omega \omega$, $a_\omega \in A$;

(iv) the entries of S are of the form dy_1, $d \in A$.

A tree automaton $\mathfrak{A} = (I, (M_k \mid k \geq 1), S, P)$ is termed *proper* iff the entries of M_1 do not contain a linear term ay_1, $a \in A$.

A tree automaton $\mathfrak{A} = (I, M, S, P)$ is called *polynomial* (resp. *recognizable*) iff the following conditions are satisfied:

(i) $M = (M_k \mid 1 \leq k \leq \bar{k})$ is a *finite sequence* of transition matrices M_k whose entries are *polynomials* in $A\langle T_\Sigma(X \cup Y_k)\rangle$ (resp. *tree series* in $A^{\text{rec}}\langle\langle T_\Sigma(X \cup Y_k)\rangle\rangle$), $1 \leq k \leq \bar{k}$. (Technically speaking, this means that all transition matrices $M_{\bar{k}+j}$, $j \geq 1$, are equal to the zero matrix.) Moreover, the matrices M_k, $1 \leq k \leq \bar{k}$, are *row finite*.

(ii) The entries of the initial state vector S are of the form $S_i = d_i y_1$, $i \in I$. Moreover, S is *row finite*.

(iii) The entries of the final state vector P are *polynomials* in $A\langle T_\Sigma(X)\rangle$ (resp. *recognizable tree series* in $A^{\text{rec}}\langle\langle T_\Sigma(X)\rangle\rangle$).

By Bozapalidis [2], by Kuich [7], and by Theorem 9 we obtain the following result.

Theorem 15 *The following statements on a formal tree series in $A\langle\langle T_\Sigma(X)\rangle\rangle$ are equivalent:*

(i) $s \in A^{\text{rec}}\langle\langle T_\Sigma(X)\rangle\rangle$,

(ii) *there exists a polynomial tree automaton $\mathfrak{A}$ with finite state set such that $s = ||\mathfrak{A}||$,*

(iii) *there exists a simple proper polynomial tree automaton $\mathfrak{A}$ with finite state set such that $s = ||\mathfrak{A}||$,*

(iv) *there exists a recognizable tree automaton $\mathfrak{A}$ with finite state set such that $s = ||\mathfrak{A}||$,*

(v) *there exists a proper recognizable tree automaton $\mathfrak{A}$ with finite state set such that $s = ||\mathfrak{A}||$,* Let $\mathfrak{A} = (I, (M_k \mid k \geq 1), S, P)$ be a simple tree automaton and $\mathfrak{T} = (Q, \mu, R)$ be a tree transducer such that $R_q = a_q y_1$, $a_q \in A$, $q \in Q$. Then $\mathfrak{T}(\mathfrak{A})$ is defined to be the tree automaton

$$\mathfrak{T}(\mathfrak{A}) = (I \times Q, (\mu(M_k) \mid k \geq 1), S \otimes R, \mu(P)) .$$

Theorem 16 *Let $\mathfrak{A} = (I, (M_k \mid k \geq 1), S, P)$ be a simple tree automaton and $\mathfrak{T} = (Q, \mu, R)$ be a linear nondeleting tree transducer. Then*

$$||\mathfrak{T}(\mathfrak{A})|| = ||\mathfrak{T}||(||\mathfrak{A}||) .$$

Proof. Consider the approximation sequences $(\sigma^j \mid j \in \mathbb{N})$ and $(\tau^j \mid j \in \mathbb{N})$ of $\mathfrak{A}$ and $\mathfrak{T}(\mathfrak{A})$ with upper bounds σ and τ, respectively. Then we prove by inducion on j that $\tau^j = \mu(\sigma^j)$, $j \geq 0$. The induction basis being clear, we proceed with the induction step. Let $j \geq 0$. Then

$$\begin{aligned}
&\tau^{j+1} = \textstyle\sum_{k\geq 1} \mu(M_k)(\mu(\sigma^j), \ldots, \mu(\sigma^j)) + \mu(P) = \\
&\textstyle\sum_{k\geq 1} \mu(M_k(\sigma^j, \ldots, \sigma^j)) + \mu(P) = \\
&\mu(\textstyle\sum_{k\geq 1} M_k(\sigma^j, \ldots, \sigma^j) + P) = \mu(\sigma^{j+1}) .
\end{aligned}$$

Here the second equality follows by Theorem 7. Hence, we obtain $\tau = \mu(\sigma)$.

We now compute the behavior of $\mathfrak{T}(\mathfrak{A})$:

$$\begin{aligned}
&||\mathfrak{T}(\mathfrak{A})|| = (S \otimes R)(\tau) = \textstyle\sum_{i\in I} \sum_{q\in Q}((S \otimes R)_i)_q (\mu(\sigma)_i)_q = \\
&\textstyle\sum_{q\in Q} \sum_{i\in I} R_q S_i \sum_{t\in T_\Sigma(X)} (\sigma_i, t)\mu(t)_q = \\
&\textstyle\sum_{q\in Q} R_q \sum_{t\in T_\Sigma(X)} \sum_{i\in I} (S_i \sigma_i, t)\mu(t)_q = \\
&\textstyle\sum_{q\in Q} R_q \sum_{t\in T_\Sigma(X)} (||\mathfrak{A}||, t)\mu(t)_q = \\
&\textstyle\sum_{q\in Q} R_q \mu(||\mathfrak{A}||)_q = ||\mathfrak{T}||(||\mathfrak{A}||).
\end{aligned}$$

□

Corollary 17 *Let $\mathfrak{A}$ be a simple polynomial tree automaton with finite state set and $\mathfrak{T}$ be a linear nondeleting recognizable tree transducer. Then $||\mathfrak{T}||(||\mathfrak{A}||)$ is in $A^{\mathrm{rec}}\langle\langle T_\Sigma(X)\rangle\rangle$.*

Acknowledgement. Many thanks are due to Ferenc Gécseg for discussions on root-to-frontier tree transducers.

References

[1] Berstel, J., Reutenauer, C.: Recognizable formal power series on trees. Theor. Comput. Sci. 18(1982) 115–148.

[2] Bozapalidis, S.: Equational elements in additive algebras. Technical Report, Aristotle University of Thessaloniki, 1997.

[3] Bozapalidis, S.: Context-free series on trees. Preprint, Aristotle University of Thessaloniki, February 1998.

[4] Gécseg, F., Steinby, M.: Tree Automata. Akademiai Kiado, 1984.

[5] Gécseg, F., Steinby, M.: Tree Languages. In: Handbook of Formal Languages (Eds.: G. Rozenberg and A. Salomaa), Springer, 1997, Vol. 3, Chapter 1, 1–68.

[6] Kuich, W.: Semirings and formal power series: Their relevance to formal languages and automata theory. In: Handbook of Formal Languages (Eds.: G. Rozenberg and A. Salomaa), Springer, 1997, Vol. 1, Chapter 9, 609–677.

[7] Kuich, W.: Formal power series over trees. In: Proceedings of the 3rd International Conference Developments in Language Theory (S. Bozapalidis, ed.), Aristotle University of Thessaloniki, 1998, 61–101.

[8] Kuich, W.: Pushdown tree automata, algebraic tree systems, and algebraic tree series. Preprint, Technische Universität Wien, 1998.

[9] Kuich, W., Salomaa, A.: Semirings, Automata, Languages. EATCS Monographs on Theoretical Computer Science, Vol. 5. Springer, 1986.

[10] Thatcher, J. W.: Generalized^2 sequential machine maps. IBM Research Report RC 2466, 1969.