Does o-Substitution Preserve Recognizability?

Andreas Maletti

Technische Universität Dresden
Department of Computer Science
01062 Dresden, Germany
maletti@tcs.inf.tu-dresden.de

Abstract. Substitution operations on tree series are at the basis of systems of equations (over tree series) and tree series transducers. Tree series transducers seem to be an interesting transformation device in syntactic pattern matching. In this contribution, it is shown that o-substitution preserves recognizable tree series provided that the target tree series is linear and the semiring is idempotent, commutative, and continuous. This result is applied to prove that the range of the o-t-ts transformation computed by a linear recognizable tree series transducer is pointwise recognizable.

1 Introduction

Tree series transducers [1] were introduced as a joint generalization of tree transducers [2, 3] and weighted tree automata [4, 5]. They thereby serve as the transducing devices corresponding to weighted tree automata. Both historical predecessors of tree series transducers have successfully been motivated from and applied in practice. Specifically, tree transducers are motivated from syntax-directed translations in compilers [6], and they are applied in, *e. g.*, computational linguistics [7] and query languages of XML databases [8]. Weighted tree automata have been applied to code selection in compilers [9] and tree pattern matching [10].

In [11] a tree-based syntactic pattern matching approach is presented and shown to be competetive. The approach is tailored to digit recognition. Using a training procedure for regular tree grammars a tree automaton is trained. To accomodate for training errors, usually a refined model using probabilities is applied. Essentially this corresponds to a weighted tree automaton. A common observation is that the recognized digit is invariant under small translations of the input image (such as, *e. g.*, small tiltings). Finitely presentable transformations (also respecting the probabilities) on the input tree can be realised by tree series transducers. Another application of tree series transducers (using the semiring of probabilities) is demonstrated in [12], where tree series transducers are trained to perform machine translation. Yet another application of tree series transducers is presented in [13], where tree series transducer are applied to code selection.

Let us illustrate one application of tree series transducers in the setting of natural language processing. Imagine a statistical channel model that is applied to a channel that translates Japanese text into English text [14]. Statistical models are built from a large corpus of hand-annotated and translated input sentences. Any such channel model gives rise to an automatically generated (statistical) translation system, which may assist translators by providing suitable candidate translations. In [14] the simple IBM model 1 [15, 16] is displayed. This model consists of several stages: reordering, insertion, and word translation. The first stage just reorders parse subtrees to accomodate for different word order (English: Subject-Verb-Object and Japanese: Subject-Object-Verb); the second stage inserts words that have no direct translation; and the final stage just perform word-to-word translation. All operations are probabilistic, so with a certain probability, the reordering TO NN $\rightarrow$ NN TO takes place. In fact, all stages are simple weighted tree to weighted tree (where the weight is a probability) transformations, which can easily be modelled by a tree series transducer. We depict the working of a tree series transducer for the reordering stage in Figure 1.

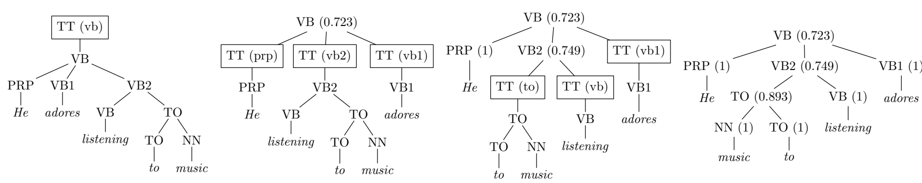

Fig. 1. Reordering performed by a tree series transducer

Tree substitution is at the core of the semantics of tree transducers, and tree series substitution fulfills this purpose for tree series transducers. In this paper we investigate o-substitution [17]. A tree series is a mapping from a set of output trees into some semiring. Let us illustrate o-substitution using the semiring of probabilities. The probability (*e. g.*, reliability), that is associated with an output tree, is taken to the nth power, if the output tree is used in n copies (is copied n times into some other tree). In this approach, an output tree stands for a composite, and the probability associated with the output tree reflects, *e. g.*, the reliability of this particular composite. When we combine composites into a new composite, then we obtain the reliability of the composite by a simple multiplication of the reliabilities of its components; each component taken as often as needed to assemble the composite (under the assumption that each component is critical for the correct functioning of the composite).

Tree series substitutions have also been studied in relation with recognizable tree series [4]. Substitution is a standard operation on tree series, and in particular, OI-substitution [18] was studied with respect to preservation of recognizability [19]. A tree series is called recognizable, if there exists a finite state automaton that computes this tree series. Recognizable tree series are of particular interest, because they are finitely representable. It is known that the result of certain substitutions is not recognizable. We study the limit of recognizability under o-substitution. Which o-substitutions will lead to recognizable tree series? Thus we aim towards classes of transformations that preserve the ability to finitely represent tree series.

Our main result states that o-substitution preserves recognizable tree series in semirings that are commutative, idempotent, and continuous [20], whenever the participating tree series are linear (*i. e.*, each variable may occur at most once in the trees in the support).

We apply this result to show that the o-t-ts transformation computed by a linear recognizable tree series transducer over a commutative, idempotent, and continuous semiring is pointwise recognizable.

2 Preliminaries

We use $\mathbb{N}$ and $\mathbb{N}_+$ to represent the nonnegative and positive integers, respectively. Further let $[k]$ be an abbreviation for $\{n \in \mathbb{N} \mid 1 \leqslant n \leqslant k\}$. A set Σ which is nonempty and finite is also called an *alphabet*. As usual, Σ^* denotes the set of all (finite) words over Σ. Given $w \in \Sigma^*$, the *length of* w is denoted by $|w|$.

A *ranked alphabet* is an alphabet Σ with a mapping $\mathrm{rk}_\Sigma \colon \Sigma \longrightarrow \mathbb{N}$. We use Σ_k to represent $\{\sigma \in \Sigma \mid \mathrm{rk}_\Sigma(\sigma) = k\}$. Moreover, we use the set $\mathrm{X} = \{\mathrm{x}_i \mid i \in \mathbb{N}_+\}$ of *variables* and $\mathrm{X}_k = \{\mathrm{x}_i \mid i \in [k]\}$. Given a ranked alphabet Σ and $V \subseteq \mathrm{X}$, the set of *Σ-trees indexed by V*, denoted by $T_\Sigma(V)$, is inductively defined to be the smallest set T such that (i) $V \subseteq T$ and (ii) for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T$ also $\sigma(t_1, \ldots, t_k) \in T$. Since we generally assume that $\Sigma \cap \mathrm{X} = \emptyset$, we write α instead of $\alpha()$ whenever $\alpha \in \Sigma_0$. Moreover, we also write T_Σ to denote $T_\Sigma(\emptyset)$.

For every $t \in T_\Sigma(\mathrm{X})$, we denote by $|t|_x$ the number of occurrences of $x \in \mathrm{X}$ in t. Let $I \subseteq \mathbb{N}_+$ be finite, $u \in T_\Sigma(\mathrm{X})$ and $u_i \in T_\Sigma(\mathrm{X})$ for every $i \in I$. By $u[u_i]_{i \in I}$ we denote the tree obtained from u by replacing every occurrence of a variable x_i with $i \in I$ by u_i. We write $u[u_1, \ldots, u_n]$ for $u[u_i]_{i \in I}$ if $I = [n]$. Let $V \subseteq \mathrm{X}$. We say that $u \in T_\Sigma(\mathrm{X})$ is *linear* and *nondeleting* in V, if every $x \in V$ occurs at most once and at least once in t, respectively. Moreover, we use $\mathrm{var}(u)$ to represent the set of variables that occur in u.

A *(commutative) semiring* is an algebraic structure $\mathcal{A} = (A, +, \cdot, 0, 1)$ consisting of two commutative monoids $(A, +, 0)$ and $(A, \cdot, 1)$ such that $\cdot$ distributes over $+$ and 0 is absorbing with respect to $\cdot$. As usual we use $\sum_{i \in I} a_i$ for sums of families $(a_i)_{i \in I}$ of $a_i \in A$ where for only finitely many $i \in I$ we have $a_i \neq 0$. A semiring $\mathcal{A} = (A, +, \cdot, 0, 1)$ is called *idempotent*, if $1 + 1 = 1$, and $\mathcal{A}$ is called *complete*, if it is possible to define an infinitary sum operation such that for arbitrary index sets I and $(a_i)_{i \in I}$ of $a_i \in A$ we have

- $\sum_{i \in \{j_1, j_2\}} a_i = a_{j_1} + a_{j_2}$ with $j_1 \neq j_2$;
- $\sum_{i \in I} a_i = \sum_{j \in J} \bigl(\sum_{i \in I_j} a_i\bigr)$ for all $(I_j)_{j \in J}$ such that $\bigcup_{j \in J} I_j = I$ and for every $j_1 \neq j_2$ we have $I_{j_1} \cap I_{j_2} = \emptyset$; and
- $a \cdot \bigl(\sum_{i \in I} a_i\bigr) = \sum_{i \in I}(a \cdot a_i)$ for all $a \in A$.

Whenever we speak of a complete semiring, we silently assume that the infinitary sum operation is given. A semiring is *naturally ordered*, whenever $\sqsubseteq \subseteq A^2$, defined by $a \sqsubseteq b$ iff there exists a $c \in A$ such that $a + c = b$, constitutes a partial order on A. Let $\mathcal{A}$ be complete and naturally ordered. We say that $\mathcal{A}$ is *continuous*, if for every index set I and $(a_i)_{i \in I}$ of $a_i \in A$ the following supremum exists and $\sum_{i \in I} a_i = \sup\{\sum_{i \in F} a_i \mid F \subseteq I \text{ with } F \text{ finite}\}$ where the supremum is taken with respect to the natural order $\sqsubseteq$. Examples of continuous semirings are

- the Boolean semiring $\mathbb{B} = (\{0,1\}, \vee, \wedge, 0, 1)$,
- the natural number semiring $\mathbb{N} = (\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$, and
- the arctic semiring $\mathbb{A} = (\mathbb{N} \cup \{\infty, -\infty\}, \max, +, -\infty, 0)$.

Let S be a set and $\mathcal{A} = (A, +, \cdot, 0, 1)$ be a semiring. A *(formal) power series* ψ is a mapping $\psi \colon S \longrightarrow A$. Given $s \in S$, we denote $\psi(s)$ also by (ψ, s) and write the series as $\sum_{s \in S} (\psi, s)\, s$. The *support* of ψ is $\mathrm{supp}(\psi) = \{s \in S \mid (\psi, s) \neq 0\}$. Power series with finite support are called *polynomials*. We denote the set of all power series by $\mathcal{A}\langle\!\langle S \rangle\!\rangle$ and the set of polynomials by $\mathcal{A}\langle S \rangle$. The polynomial with empty support is denoted by $\widetilde{0}$. Power series $(\psi_i)_{i \in I} \in \mathcal{A}\langle\!\langle S \rangle\!\rangle$ are added componentwise; *i. e.*, $(\sum_{i \in I} \psi_i, s) = \sum_{i \in I} (\psi_i, s)$ for every $s \in S$, and we multiply $\psi \in \mathcal{A}\langle\!\langle S \rangle\!\rangle$ with a coefficient $a \in A$ componentwise; *i. e.*, $(a \cdot \psi, s) = a \cdot (\psi, s)$ for every $s \in S$.

In this paper, we only consider power series in which the set S is a set of trees. Such power series are also called *tree series*. Let Δ be a ranked alphabet. A tree series $\psi \in \mathcal{A}\langle\!\langle T_\Delta(\mathrm{X}) \rangle\!\rangle$ is said to be *linear* and *nondeleting* in $V \subseteq \mathrm{X}$, if every $t \in \mathrm{supp}(\psi)$ is linear and nondeleting in V, respectively. We also use $\mathrm{var}(\psi) = \bigcup_{u \in \mathrm{supp}(\psi)} \mathrm{var}(u)$.

Now let $\mathcal{A}$ be a complete semiring and $\psi \in \mathcal{A}\langle\!\langle T_\Delta(\mathrm{X}) \rangle\!\rangle$ and let $I \subseteq \mathbb{N}_+$ be finite and $\psi_i \in \mathcal{A}\langle\!\langle T_\Delta(\mathrm{X}) \rangle\!\rangle$ for every $i \in I$. The o-*substitution of* $(\psi_i)_{i \in I}$ *into* ψ, denoted by $\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_i)_{i \in I}$, is defined by

$$\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_i)_{i \in I} = \sum_{\substack{u \in \mathrm{supp}(\psi), \\ (\forall i \in I) \colon u_i \in \mathrm{supp}(\psi_i)}} \Big((\psi, u) \cdot \prod_{i \in I} (\psi_i, u_i)^{|u|_{\mathrm{x}_i}} \Big)\, u[u_i]_{i \in I} \enspace .$$

If we suppose that $I = [n]$, then we also write $\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_1, \ldots, \psi_n)$ instead of $\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_i)_{i \in I}$. In an expression $\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_1, \ldots, \psi_n)$ the series ψ is called the *target* and every ψ_i is called a *source*.

Let us recall the notion of recognizable tree series [4, 5, 18, 21]. Let Σ be a ranked alphabet and $\mathcal{A} = (A, +, \cdot, 0, 1)$ be a semiring. A *(bottom-up) weighted tree automaton M (over Σ and $\mathcal{A}$)*, abbreviated to wta, is a tuple $(Q, \Sigma, \mathcal{A}, F, \mu)$ where Q is an alphabet of *states*, $F \colon Q \longrightarrow A$ is a *final weight distribution* and $\mu = (\mu_k)_{k \in \mathbb{N}}$ with $\mu_k \colon \Sigma_k \longrightarrow A^{Q \times Q^k}$ is a *tree representation*. The *initial algebra semantics* of M is determined by the mapping $h_\mu \colon T_\Sigma \longrightarrow A^Q$ given by

$$h_\mu(\sigma(t_1, \ldots, t_k))_q = \sum_{q_1, \ldots, q_k \in Q} \mu_k(\sigma)_{q, q_1, \ldots, q_k} \cdot h_\mu(t_1)_{q_1} \cdot \ldots \cdot h_\mu(t_k)_{q_k}$$

for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, $q \in Q$, and $t_1, \ldots, t_k \in T_\Sigma$. The tree series *recognized by* M, denoted by $\|M\|$, is defined by $(\|M\|, t) = \sum_{q \in Q} F_q \cdot h_\mu(t)_q$ for every $t \in T_\Sigma$.

We use the method of [22, 21] to graphically represent wta. Note that we write $\mu_0(\alpha)_q$ instead of $\mu_0(\alpha)_{q,()}$ for every $\alpha \in \Sigma_0$ and $q \in Q$. A tree series $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$ is termed *recognizable*, if there exists a wta M over Σ and $\mathcal{A}$ such that $\psi = \|M\|$. The class of all recognizable tree series over Σ and $\mathcal{A}$ is denoted by $\mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma \rangle\rangle$.

Let Q be an alphabet. We write $Q(V)$ for $\{q(v) \mid q \in Q, v \in V\}$. Now let $\mathcal{A} = (A, +, \cdot, 0, 1)$ be a semiring and Σ and Δ be ranked alphabets. A *tree representation* μ (over Q, Σ, Δ, and $\mathcal{A}$) [1] is a family $(\mu(\sigma))_{\sigma \in \Sigma}$ of matrices $\mu(\sigma) \in \mathcal{A}\langle\langle T_\Delta(\mathrm{X}) \rangle\rangle^{Q \times Q(\mathrm{X}_k)^*}$ where $k = \mathrm{rk}_\Sigma(\sigma)$ such that for every $q \in Q$ and $w \in Q(\mathrm{X}_k)^*$ it holds that $\mu(\sigma)_{q,w} \in \mathcal{A}\langle\langle T_\Delta(\mathrm{X}_n) \rangle\rangle$ with $n = |w|$, and $\mu(\sigma)_{q,w} \neq \widetilde{0}$ for only finitely many $(q, w) \in Q \times Q(\mathrm{X}_k)^*$. A tree representation μ is said to be *recognizable* and *linear*, if $\mu(\sigma)_{q,w}$ is recognizable and linear for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $(q, w) \in Q \times Q(\mathrm{X}_k)^*$, respectively. A *tree series transducer* [1, 20], in the sequel abbreviated to tst, is a sixtuple $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ consisting of

- an alphabet Q of *states*,
- ranked alphabets Σ and Δ, also called *input* and *output ranked alphabet*, respectively,
- a complete semiring $\mathcal{A} = (A, +, \cdot, 0, 1)$,
- a vector $F \in \mathcal{A}\langle\langle T_\Delta(\mathrm{X}_1) \rangle\rangle^Q$, called *top-most output*, such that for all $q \in Q$: F_q is nondeleting and linear in X_1, and
- a tree representation μ over Q, Σ, Δ, and $\mathcal{A}$.

Tst inherit the properties *recognizable* and *linear* from their tree representation. Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ be a tst. Then M induces a mapping $\|M\|^{\mathrm{o}} \colon T_\Sigma \longrightarrow \mathcal{A}\langle\langle T_\Delta \rangle\rangle$ as follows. For every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T_\Sigma$ we define the mapping $h^{\mathrm{o}}_\mu \colon T_\Sigma \longrightarrow \mathcal{A}\langle\langle T_\Delta \rangle\rangle^Q$ inductively for every $q \in Q$ by

$$h^{\mathrm{o}}_\mu(\sigma(t_1, \ldots, t_k))_q = \sum_{\substack{w \in Q(\mathrm{X}_k)^*,\\ w = q_1(\mathrm{x}_{i_1}) \cdots q_n(\mathrm{x}_{i_n})}} \mu_k(\sigma)_{q,w} \xleftarrow[\mathrm{o}]{} (h^{\mathrm{o}}_\mu(t_{i_1})_{q_1}, \ldots, h^{\mathrm{o}}_\mu(t_{i_n})_{q_n}) \ .$$

For every $t \in T_\Sigma$ the *o-tree-to-tree-series (for short: o-t-ts) transformation computed by* M is $\|M\|^{\mathrm{o}}(t) = \sum_{q \in Q} F_q \xleftarrow[\mathrm{o}]{} (h^{\mathrm{o}}_\mu(t)_q)$.

3 Preservation of recognizability

In this section we consider the question whether o-substitution preserves recognizability. Let Σ be a ranked alphabet. It is known that IO substitution does not, in general, preserve recognizability. However, IO substitution on linear tree languages preserves recognizability [23].

In [24] a first result on tree series is presented for OI substitution. For every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $\psi_1, \ldots, \psi_k \in \mathcal{A}\langle\langle T_\Sigma(\mathrm{X})\rangle\rangle$, we define

$$\sigma(\psi_1, \ldots, \psi_k) = \sum_{t_1,\ldots,t_k \in T_\Sigma(\mathrm{X})} \big((\psi_1, t_1) \cdot \ldots \cdot (\psi_k, t_k)\big)\, \sigma(t_1, \ldots, t_k) \enspace .$$

Note that this sum is always well-defined. Let $t \in T_\Sigma(\mathrm{X})$ be a tree, $n \in \mathbb{N}$, and $\psi_1, \ldots, \psi_n \in \mathcal{A}\langle\langle T_\Sigma(\mathrm{X})\rangle\rangle$. For every $j \in [n]$, $\ell \in \mathbb{N}_+ \setminus [n]$ let

$$\mathrm{x}_j \underset{\mathrm{OI}}{\longleftarrow} (\psi_1, \ldots, \psi_n) = \psi_j \quad \text{and} \quad \mathrm{x}_\ell \underset{\mathrm{OI}}{\longleftarrow} (\psi_1, \ldots, \psi_n) = 1\, \mathrm{x}_\ell$$

and for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T_\Sigma(\mathrm{X})$ let

$$t \underset{\mathrm{OI}}{\longleftarrow} (\psi_1, \ldots, \psi_n) = \sigma(t_1 \underset{\mathrm{OI}}{\longleftarrow} (\psi_1, \ldots, \psi_n), \ldots, t_k \underset{\mathrm{OI}}{\longleftarrow} (\psi_1, \ldots, \psi_n)) \enspace ,$$

where $t = \sigma(t_1, \ldots, t_k)$. Now let $\psi \in \mathcal{A}\langle\langle T_\Sigma(\mathrm{X})\rangle\rangle$. We define $\psi \underset{\mathrm{OI}}{\longleftarrow} (\psi_1, \ldots, \psi_n)$ by

$$\psi \underset{\mathrm{OI}}{\longleftarrow} (\psi_1, \ldots, \psi_n) = \sum_{t \in T_\Sigma(\mathrm{X})} (\psi, t) \cdot \big(t \underset{\mathrm{OI}}{\longleftarrow} (\psi_1, \ldots, \psi_n)\big) \enspace .$$

Note that also this sum is always well-defined. With the help of [24] we can easily relate o-substitution and OI substitution. Recall that our semirings are always commutative.

Proposition 1. *Let $n \in \mathbb{N}$, $\psi \in \mathcal{A}\langle\langle T_\Sigma(\mathrm{X}_n)\rangle\rangle$ be nondeleting and linear in X_n, and $\psi_1, \ldots, \psi_n \in \mathcal{A}\langle\langle T_\Sigma(\mathrm{X})\rangle\rangle$.*

$$\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_1, \ldots, \psi_n) = \psi \underset{\mathrm{OI}}{\longleftarrow} (\psi_1, \ldots, \psi_n)$$

Proof. Clearly, $t \underset{\mathrm{OI}}{\longleftarrow} (1\; t_i)_{i \in I} = 1\; t[t_i]_{i \in I}$ for every $t \in T_\Sigma(\mathrm{X}_I)$ and family $(t_i)_{i \in I} \in T_\Sigma(\mathrm{X})^I$.

$$\begin{aligned}
\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_i)_{i \in I} &= \sum_{\substack{t \in T_\Sigma(\mathrm{X}_I),\\ (\forall i \in I)\colon t_i \in T_\Sigma(\mathrm{X})}} \Big((\psi, t) \cdot \prod_{i \in I} (\psi_i, t_i)\Big) \cdot (1\; t[t_i]_{i \in I}) \\
&= \qquad (\text{by } t \underset{\mathrm{OI}}{\longleftarrow} (1\; t_i)_{i \in I} = 1\; t[t_i]_{i \in I}) \\
&\qquad \sum_{\substack{t \in T_\Sigma(\mathrm{X}_I),\\ (\forall i \in I)\colon t_i \in T_\Sigma(\mathrm{X})}} \Big((\psi, t) \cdot \prod_{i \in I} (\psi_i, t_i)\Big) \cdot \big(t \underset{\mathrm{OI}}{\longleftarrow} (1\; t_i)_{i \in I}\big) \\
&= \qquad (\text{by [24, Theorem 6] and definition of } \underset{\mathrm{OI}}{\longleftarrow}) \\
&\qquad \sum_{t \in T_\Sigma(\mathrm{X}_I)} (\psi, t) \cdot \big(t \underset{\mathrm{OI}}{\longleftarrow} (\psi_i)_{i \in I}\big) \;=\; \psi \underset{\mathrm{OI}}{\longleftarrow} (\psi_i)_{i \in I}
\end{aligned}$$

Theorem 2 (*cf.* [24]). *For every $n \in \mathbb{N}$, $\psi \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma(\mathrm{X}_n)\rangle\rangle$ such that ψ is nondeleting and linear in X_n, and every $\psi_1, \ldots, \psi_n \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma\rangle\rangle$ we have that $\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_1, \ldots, \psi_n) \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma\rangle\rangle$.*

Proof. The statement is proved for OI-substitution in [24, Corollary 14]. Since OI-substitution coincides with o-substitution on nondeleting and linear target tree series (see Proposition 1), we obtain the statement.

We would like to achieve a result which does not depend on nondeletion of ψ (see Theorem 2). Let us show the main idea in a simple setting. Let $\psi \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma(\mathrm{X}_1)\rangle\rangle$ be linear in X_1 and $\psi_1 \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma\rangle\rangle$. Our goal is to show that $\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_1)$ is recognizable, thus we need to present a wta $M' = (Q', \Sigma, \mathcal{A}, F', \mu')$ that recognizes $\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_1)$. Let $M = (Q, \Delta, \mathcal{A}, F, \mu)$ and $M_1 = (Q_1, \Sigma, \mathcal{A}, F_1, \mu_1)$ be wta that recognize ψ and ψ_1, respectively. We employ a standard idea for the construction of M'. Roughly speaking, we take the disjoint union of M and M_1 and add transitions that nondeterministically change from M_1 to M. More precisely, for every $k \in \mathbb{N}_+$, $\sigma \in \Sigma_k$, $q \in Q$, and $q_1, \ldots, q_k \in Q_1$ we set

$$\mu'_k(\sigma)_{q,q_1,\ldots,q_k} = \sum_{p \in Q_1} \mu_0(\mathrm{x}_1)_q \cdot (F_1)_p \cdot (\mu_1)_k(\sigma)_{p,q_1,\ldots,q_k} \enspace .$$

Roughly, for each state p of M_1 we take $(\mu_1)_k(\sigma)_{p,q_1\cdots q_k}$ of M_1, multiply $(F_1)_p$, and multiply $\mu_0(\mathrm{x}_1)_q$ for entering M (via x_1) in state q. Nullary symbols σ are treated similarly. We employ a proof method, which requires us to make the input alphabets Σ and Δ disjoint. This simplifies the proof because each tree then admits a unique decomposition into (at most one) part that needs to be processed by M_1 and a part that needs to be processed by M.

Proposition 3. *Let $\mathcal{A}$ be idempotent and continuous. Let $J \subseteq I \subseteq \mathbb{N}_+$ be finite, $\psi \in \mathcal{A}\langle\langle T_\Delta(\mathrm{X})\rangle\rangle$ such that $J \cap \mathrm{var}(\psi) = I \cap \mathrm{var}(\psi)$, and for every $i \in I$ let $\psi_i \in \mathcal{A}\langle\langle T_\Delta(\mathrm{X})\rangle\rangle$ such that $\psi_i \neq \widetilde{0}$ for every $i \in I \setminus J$.*

$$\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_i)_{i\in I} = \psi \underset{\mathrm{o}}{\longleftarrow} (\psi_j)_{j \in J}$$

Theorem 4. *Let $\mathcal{A}$ be a continuous and idempotent semiring. Let $n \in \mathbb{N}$, $\psi \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma(\mathrm{X}_n)\rangle\rangle$ be linear in X_n, and $\psi_1, \ldots, \psi_n \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma\rangle\rangle$.*

$$\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_1, \ldots, \psi_n) \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma\rangle\rangle$$

Proof. Let $\psi_i = \widetilde{0}$ for some $i \in [n]$. Then $\psi \underset{\mathrm{o}}{\longleftarrow} (\psi_1, \ldots, \psi_n) = \widetilde{0}$, which is recognizable. Thus let $\psi_i \neq \widetilde{0}$ for all $i \in [n]$. For every $k \in \mathbb{N}_+$ let $\Delta_k = \Sigma_k$ and $\Delta_0 = \Sigma_0 \cup \mathrm{X}_n$. Since $\psi \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma(\mathrm{X}_n)\rangle\rangle$ and $\psi_1, \ldots, \psi_n \in \mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma\rangle\rangle$, there exist wta $M = (Q, \Delta, \mathcal{A}, F, \mu)$ and $M_i = (Q_i, \Sigma, \mathcal{A}, F_i, \mu_i)$ such that $\|M\| = \psi$ and $\|M_i\| = \psi_i$ for every $i \in [n]$.

For every $i \in [n]$ and $k \in \mathbb{N}$ let $\overline{\Sigma}^i$ be $\overline{\Sigma}^i_k = \{\overline{\sigma}^i \mid \sigma \in \Sigma_k\}$. We define $\mathrm{bar}_i \colon T_\Sigma \longrightarrow T_{\overline{\Sigma}^i}$ by $\mathrm{bar}_i(\sigma(t_1, \ldots, t_k)) = \overline{\sigma}^i(\mathrm{bar}_i(t_1), \ldots, \mathrm{bar}_i(t_k))$ for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T_\Sigma$. Moreover, we define the mapping $\mathrm{bar}_i \colon \mathcal{A}\langle\langle T_\Sigma\rangle\rangle \longrightarrow \mathcal{A}\langle\langle T_{\overline{\Sigma}^i}\rangle\rangle$ for every $\varphi \in \mathcal{A}\langle\langle T_\Sigma\rangle\rangle$ by

$$\mathrm{bar}_i(\varphi) = \sum_{t \in T_\Sigma} (\varphi, t)\ \mathrm{bar}_i(t) \enspace .$$

Without loss of generality, we assume that for every $i \in [n]$ we have that (i) Σ and $\overline{\Sigma}^i$ are disjoint and (ii) Q and Q_i are disjoint. Let $\Sigma'_k = \Sigma_k \cup \bigcup_{1 \leq i \leq n} \overline{\Sigma}^i_k$ for every $k \in \mathbb{N}$, and $Q' = Q \cup \bigcup_{1 \leq i \leq n} Q_i$. We construct a wta M' recognizing $\psi \underset{\mathrm{o}}{\longleftarrow} (\mathrm{bar}_1(\psi_1), \ldots, \mathrm{bar}_n(\psi_n))$ as follows. Let $M' = (Q', \Sigma', \mathcal{A}, F', \mu')$ where for every $i \in [n]$, $k \in \mathbb{N}$, $\sigma \in \Sigma_k$:

- $F'_q = F_q$ for every $q \in Q$ and $F'_p = 0$ for every $p \in \bigcup_{1 \leq i \leq n} Q_i$;
- $\mu'_k(\overline{\sigma}^i)_{p,w} = (\mu_i)_k(\sigma)_{p,w}$ for every $p \in Q_i$ and $w \in (Q_i)^k$;
- $\mu'_k(\sigma)_{q,w} = \mu_k(\sigma)_{q,w}$ for every $q \in Q$ and $w \in Q^k$; and
- $\mu'_k(\overline{\sigma}^i)_{q,w} = \sum_{p \in Q_i} \mu_0(\mathrm{x}_i)_q \cdot (F_i)_p \cdot (\mu_i)_k(\sigma)_{p,w}$ for every $q \in Q$ and $w \in (Q_i)^k$.

All the remaining entries in μ' are set to 0.

We claim that $\psi' = \psi \underset{\mathrm{o}}{\longleftarrow} (\mathrm{bar}_1(\psi_1), \ldots, \mathrm{bar}_n(\psi_n))$ is recognizable. In fact, M' recognizes ψ'. Clearly, $h_{\mu'}(\mathrm{bar}_i(t))_p = h_{\mu_i}(t)_p$ for every $i \in [n]$, $t \in T_\Sigma$, and $p \in Q_i$. Next we prove that for every $q \in Q$ and $t \in T_\Sigma(\mathrm{X}_n)$, which is linear in X_n, and family $(u_i)_{i \in \mathrm{var}(t)} \in T_\Sigma^{\mathrm{var}(t)}$ we have

$$h_{\mu'}(t[\mathrm{bar}_i(u_i)]_{i \in \mathrm{var}(t)})_q = h_\mu(t)_q \cdot \prod_{i \in \mathrm{var}(t)} \big(\|M_i\|, u_i\big) \enspace .$$

We prove this statement inductively, so let $t = \mathrm{x}_j$ for some $j \in [n]$. Moreover, let $u_j = \sigma(t_1, \ldots, t_k)$ for some $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T_\Sigma$.

$$\begin{aligned}
& h_{\mu'}(\mathrm{x}_j[\mathrm{bar}_i(u_i)]_{i \in \mathrm{var}(\mathrm{x}_j)})_q \\
= \quad & \text{(by substitution and definition of } \mathrm{bar}_j) \\
& h_{\mu'}\big(\overline{\sigma}^j(\mathrm{bar}_j(t_1), \ldots, \mathrm{bar}_j(t_k))\big)_q \\
= & \sum_{q_1, \ldots, q_k \in Q'} \mu'_k(\overline{\sigma}^j)_{q, q_1 \cdots q_k} \cdot \prod_{i \in [k]} h_{\mu'}(\mathrm{bar}_j(t_i))_{q_i} \\
= \quad & \text{(by definition of } \mu' \text{ and } h_{\mu'}(\mathrm{bar}_j(t_i))_{q_i} = h_{\mu_j}(t_i)_{q_i}) \\
& \sum_{q_1, \ldots, q_k \in Q_j} \sum_{p \in Q_j} \mu_0(\mathrm{x}_j)_q \cdot (F_j)_p \cdot (\mu_j)_k(\sigma)_{p, q_1 \cdots q_k} \cdot \prod_{i \in [k]} h_{\mu_j}(t_i)_{q_i} \\
= & \sum_{p \in Q_j} \mu_0(\mathrm{x}_j)_q \cdot (F_j)_p \cdot h_{\mu_j}(\sigma(t_1, \ldots, t_k))_p \\
= & \; \mu_0(\mathrm{x}_j)_q \cdot \big(\|M_j\|, \sigma(t_1, \ldots, t_k)\big) \\
= & \; h_\mu(\mathrm{x}_j)_q \cdot \prod_{i \in \mathrm{var}(\mathrm{x}_j)} \big(\|M_i\|, u_i\big)
\end{aligned}$$

Let $t = \sigma(t_1, \ldots, t_k)$ for some $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T_\Sigma(\mathrm{X}_n)$.

$$\begin{aligned}
& h_{\mu'}(\sigma(t_1, \ldots, t_k)[\mathrm{bar}_i(u_i)]_{i \in \mathrm{var}(t)})_q \\
= \quad & \text{(by substitution)} \\
& h_{\mu'}(\sigma(t_1[\mathrm{bar}_i(u_i)]_{i \in \mathrm{var}(t_1)}, \ldots, t_k[\mathrm{bar}_i(u_i)]_{i \in \mathrm{var}(t_k)}))_q
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{q_1,\dots,q_k \in Q'} \mu'_k(\sigma)_{q,q_1\cdots q_k} \cdot \prod_{j\in[k]} h_{\mu'}(t_j[\mathrm{bar}_i(u_i)]_{i\in\mathrm{var}(t_j)})_{q_j} \\
&= \qquad \text{(by induction hypothesis and definition of } \mu'\text{)} \\
&\quad \sum_{q_1,\dots,q_k \in Q} \mu_k(\sigma)_{q,q_1\cdots q_k} \cdot \prod_{j\in[k]} \Big(h_\mu(t_j)_{q_j} \cdot \prod_{i\in\mathrm{var}(t_j)} \big(\|M_i\|, u_i\big)\Big) \\
&= h_\mu(\sigma(t_1,\dots,t_k))_q \cdot \prod_{j\in[k], i\in\mathrm{var}(t_j)} \big(\|M_i\|, u_i\big) \\
&= \qquad \text{(because } t \text{ is linear in } \mathrm{X}_n\text{)} \\
&\quad h_\mu(\sigma(t_1,\dots,t_k))_q \cdot \prod_{i\in\mathrm{var}(t)} \big(\|M_i\|, u_i\big)
\end{aligned}
$$

This completes the proof of the auxiliary statement. Consequently,

$$
\begin{aligned}
(\|M'\|, t[\mathrm{bar}_i(u_i)]_{i\in\mathrm{var}(t)}) &= (\|M\|, t) \cdot \prod_{i\in\mathrm{var}(t)} (\|M_i\|, u_i) \\
&= (\psi, t) \cdot \prod_{i\in\mathrm{var}(t)} (\psi_i, u_i) \ . \qquad (1)
\end{aligned}
$$

Using this result, we can show that $\psi' = \psi \underset{\mathrm{o}}{\longleftarrow} (\mathrm{bar}_i(\psi_i))_{i\in I}$ is recognizable. In fact, this is the tree series that is recognized by M'.

$$
\begin{aligned}
&\quad \psi \underset{\mathrm{o}}{\longleftarrow} (\mathrm{bar}_1(\psi_1), \dots, \mathrm{bar}_n(\psi_n)) \\
&= \qquad \text{(by distributivity)} \\
&\quad \sum_{t\in\mathrm{supp}(\psi)} (\psi, t) \cdot \Big((1\ t) \underset{\mathrm{o}}{\longleftarrow} \big(\mathrm{bar}_1(\psi_1), \dots, \mathrm{bar}_n(\psi_n)\big)\Big) \\
&= \qquad \text{(by Proposition 3)} \\
&\quad \sum_{t\in\mathrm{supp}(\psi)} (\psi, t) \cdot \Big((1\ t) \underset{\mathrm{o}}{\longleftarrow} \big(\mathrm{bar}_i(\psi_i)\big)_{i\in\mathrm{var}(t)}\Big) \\
&= \qquad \text{(by definition of } \underset{\mathrm{o}}{\longleftarrow} \text{ because } t \text{ is linear)} \\
&\quad \sum_{\substack{t\in\mathrm{supp}(\psi),\\ (\forall i\in\mathrm{var}(t)):\ u_i\in\mathrm{supp}(\mathrm{bar}_i(\psi_i))}} \Big((\psi, t) \cdot \prod_{i\in\mathrm{var}(t)} (\mathrm{bar}_i(\psi_i), u_i)\Big)\ t[u_i]_{i\in\mathrm{var}(t)} \\
&= \qquad \text{(by definition of } \mathrm{bar}_i\text{)} \\
&\quad \sum_{\substack{t\in T_\Sigma(\mathrm{X}_n),\\ (\forall i\in\mathrm{var}(t)):\ u_i\in T_\Sigma}} \Big((\psi, t) \cdot \prod_{i\in\mathrm{var}(t)} (\psi_i, u_i)\Big)\ t[\mathrm{bar}_i(u_i)]_{i\in\mathrm{var}(t)} \\
&= \qquad \text{(by (1))} \\
&\quad \sum_{\substack{t\in T_\Sigma(\mathrm{X}_n),\\ (\forall i\in\mathrm{var}(t)):\ u_i\in T_\Sigma}} \big(\|M\|, t[\mathrm{bar}_i(u_i)]_{i\in\mathrm{var}(t)}\big)\ t[\mathrm{bar}_i(u_i)]_{i\in\mathrm{var}(t)}
\end{aligned}
$$

$$= \sum_{u \in T_{\Sigma'}} \Big(\sum_{\substack{t \in T_\Sigma(\mathrm{X}_n), \\ (\forall i \in \mathrm{var}(t))\colon u_i \in T_\Sigma}} \big(\|M\|, t[\mathrm{bar}_i(u_i)]_{i \in \mathrm{var}(t)}\big)\; t[\mathrm{bar}_i(u_i)]_{i \in \mathrm{var}(t)} \; , \; u \Big)\; u$$

$$= \qquad \text{(because } t \text{ and } u_i \text{ are uniquely determined by } u\text{)}$$

$$\sum_{u \in T_{\Sigma'}} \big(\|M\|, u\big)\; u \;=\; \|M\|$$

Finally, we need to remove the annotation. To this end we define the mapping $\mathrm{unbar}\colon T_{\Sigma'}(\mathrm{X}) \longrightarrow T_\Sigma(\mathrm{X})$ for every $x \in \mathrm{X}$, $k \in \mathbb{N}$, $i \in [n]$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T_{\Sigma'}(\mathrm{X})$ by

$$\begin{aligned} \mathrm{unbar}(x) &= x \\ \mathrm{unbar}(\sigma(t_1, \ldots, t_k)) &= \sigma(\mathrm{unbar}(t_1), \ldots, \mathrm{unbar}(t_k)) \\ \mathrm{unbar}(\overline{\sigma}^i(t_1, \ldots, t_k)) &= \sigma(\mathrm{unbar}(t_1), \ldots, \mathrm{unbar}(t_k)) \enspace . \end{aligned}$$

Finally, let $\mathrm{unbar}\colon \mathcal{A}\langle\!\langle T_{\Sigma'}(\mathrm{X})\rangle\!\rangle \longrightarrow \mathcal{A}\langle\!\langle T_\Sigma(\mathrm{X})\rangle\!\rangle$ be defined by

$$\mathrm{unbar}(\varphi) = \sum_{t \in T_{\Sigma'}(\mathrm{X})} (\varphi, t)\; \mathrm{unbar}(t)$$

for every $\varphi \in \mathcal{A}\langle\!\langle T_{\Sigma'}(\mathrm{X})\rangle\!\rangle$. Clearly, $\mathrm{unbar}(\psi') = \psi \xleftarrow[\mathrm{o}]{} (\psi_1, \ldots, \psi_n)$. Moreover, unbar can be realized by a nondeleting, linear tree transducer (with one state) of [24] (which uses OI substitution). Since ψ' is a recognizable tree series and nondeleting, linear tree transducers of [24] preserve recognizability, also $\mathrm{unbar}(\psi')$ is recognizable, which proves the statement.

Let us illustrate the previous theorem on an example.

Example 5. Let $\Sigma = \{\gamma^{(1)}, \alpha^{(0)}\}$ and consider the arctic semiring. Let

$$\psi = \max_{u \in T_\Sigma(\mathrm{X}_1)} \mathrm{height}(u)\; u \quad \text{and} \quad \psi' = \max_{u \in T_\Sigma} \mathrm{height}(u)\; u \enspace .$$

Then $\psi \xleftarrow[\mathrm{o}]{} (\psi')$ is recognizable. In fact, $\psi \xleftarrow[\mathrm{o}]{} (\psi') = \psi'$. We show the wta that recognize ψ and $\psi \xleftarrow[\mathrm{o}]{} (\psi')$ [the automaton that is constructed in Theorem 4] in Fig. 2.

4 Application to tree series transducers

In Theorem 4 we showed that o-substitution preserves recognizability under certain conditions. We now apply this theorem to tst. In fact this means that theorems about wta can be applied. We demonstrate such an application after the theorem.

Theorem 6. *Let $\mathcal{A}$ be an idempotent and continuous semiring. Moreover, let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ be a linear recognizable tst. Then $\|M\|^{\mathrm{o}}(t)$ is recognizable for every $t \in T_\Sigma$.*

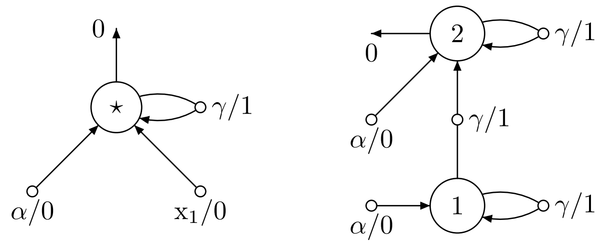

Fig. 2. Wta recognizing ψ [left] and $\psi \underset{\mathrm{o}}{\longleftarrow} (\psi')$ [right] over $\mathbb{A}$.

Proof. We first prove that $h_\mu^{\mathrm{o}}(t)_q$ is recognizable for every $t \in T_\Sigma$ and $q \in Q$ by induction on t. Let $t = \sigma(t_1, \ldots, t_k)$ for some $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T_\Sigma$.

$$h_\mu^{\mathrm{o}}(\sigma(t_1, \ldots, t_k))_q = \sum_{\substack{w \in Q(\mathrm{X}_k)^*, \\ w = q_1(\mathrm{x}_{i_1}) \cdots q_n(\mathrm{x}_{i_n})}} \mu_k(\sigma)_{q,w} \underset{\mathrm{o}}{\longleftarrow} (h_\mu^{\mathrm{o}}(t_{i_1})_{q_1}, \ldots, h_\mu^{\mathrm{o}}(t_{i_n})_{q_n}) \enspace .$$

By induction hypothesis $h_\mu^{\mathrm{o}}(t_{i_j})_{q_j}$ is recognizable for every $j \in [n]$. Since M is recognizable, $\mu_k(\sigma)_{q,w}$ is recognizable. By Theorem 4 also

$$\mu_k(\sigma)_{q,w} \underset{\mathrm{o}}{\longleftarrow} (h_\mu^{\mathrm{o}}(t_{i_1})_{q_1}, \ldots, h_\mu^{\mathrm{o}}(t_{i_n})_{q_n})$$

is recognizable because $\mu_k(\sigma)_{q,w}$ is linear in X_n. Since recognizable tree series are closed under finite sums [4] we obtain that $h_\mu^{\mathrm{o}}(t)_q$ is recognizable.

For every $t \in T_\Sigma$ we have $\|M\|^{\mathrm{o}}(t) = \sum_{q \in Q} F_q \underset{\mathrm{o}}{\longleftarrow} (h_\mu^{\mathrm{o}}(t)_q)$. We showed that $h_\mu^{\mathrm{o}}(t)_q$ is recognizable. Moreover, $F_q \underset{\mathrm{o}}{\longleftarrow} (h_\mu^{\mathrm{o}}(t)_q)$ is recognizable due to Theorem 4. Thus, also $\|M\|^{\mathrm{o}}(t)$ is recognizable.

Since idempotent semirings are zero-sum free [25], we obtain the following corollary. Other results on recognizable tree series can be applied similarly.

Corollary 7. *Let $\mathcal{A}$ be an idempotent and continuous semiring with recursive operations. Moreover, let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ be a linear recognizable tst. Then for every $t \in T_\Sigma$ it is decidable whether $\|M\|^{\mathrm{o}}(t) = \widetilde{0}$ or not.*

Proof. By Theorem 6 we have that $\psi = \|M\|^{\mathrm{o}}(t)$ is recognizable and by [26] we can decide whether $\psi = \widetilde{0}$.

References

1. Engelfriet, J., Fülöp, Z., Vogler, H.: Bottom-up and top-down tree series transformations. J. Autom. Lang. Combin. **7**(1) (2002) 11–70
2. Rounds, W.C.: Mappings and grammars on trees. Math. Systems Theory **4**(3) (1970) 257–287
3. Thatcher, J.W.: Generalized^2 sequential machine maps. J. Comput. System Sci. **4**(4) (1970) 339–367

4. Berstel, J., Reutenauer, C.: Recognizable formal power series on trees. Theoret. Comput. Sci. **18**(2) (1982) 115–148
5. Kuich, W.: Formal power series over trees. In Bozapalidis, S., ed.: Proc. 3rd Int. Conf. Develop. in Lang. Theory, Aristotle University of Thessaloniki (1998) 61–101
6. Irons, E.T.: A syntax directed compiler for ALGOL 60. Commun. ACM **4**(1) (1961) 51–55
7. Morawietz, F., Cornell, T.: The MSO logic-automation connection in linguistics. In Lecomte, A., Lamarche, F., Perrier, G., eds.: Proc. 2nd Int. Conf. Logical Aspects of Comput. Linguist.V̇olume 1582 of LNCS. (1999) 112–131
8. Bex, G.J., Maneth, S., Neven, F.: A formal model for an expressive fragment of XSLT. Inform. Systems **27**(1) (2002) 21–39
9. Ferdinand, C., Seidl, H., Wilhelm, R.: Tree automata for code selection. Acta Inform. **31**(8) (1994) 741–760
10. Seidl, H.: Finite tree automata with cost functions. Theoret. Comput. Sci. **126**(1) (1994) 113–142
11. López, D., Piñaga, I.: Syntactic pattern recognition by error correcting analysis on tree automata. In Ferri, F.J., Quereda, J.M.I., Amin, A., Pudil, P., eds.: Proc. Joint IAPR Int. Workshops Advances in Pattern Recognition. Volume 1876 of LNCS., Springer (2000) 133–142
12. Graehl, J., Knight, K.: Training tree transducers. In: HLT-NAACL. (2004) 105–112
13. Borchardt, B.: Code selection by tree series transducers. In Domaratzki, M., Okhotin, A., Salomaa, K., Yu, S., eds.: Proc. 9th Int. Conf. Implementation and Application of Automata. Volume 3317 of LNCS., Springer (2004) 57–67
14. Yamada, K., Knight, K.: A syntax-based statistical translation model. In: Proc. 39th Annual Meeting Assoc. Comput. Ling., Morgan Kaufmann (2001) 523–530
15. Brown, P.F., Cocke, J., Della Pietra, S., Della Pietra, V.J., Jelinek, F., Mercer, R.L., Roossin, P.S.: A statistical approach to language translation. In: Proc. 12th Int. Conf. Comput. Ling. (1988) 71–76
16. Brown, P.F., Della Pietra, S., Della Pietra, V.J., Mercer, R.L.: The mathematics of statistical machine translation: Parameter estimation. Comput. Linguist. **19**(2) (1993) 263–311
17. Fülöp, Z., Vogler, H.: Tree series transformations that respect copying. Theory Comput. Systems **36**(3) (2003) 247–293
18. Bozapalidis, S.: Equational elements in additive algebras. Theory Comput. Systems **32**(1) (1999) 1–33
19. Kuich, W.: Full abstract families of tree series I. In Karhumäki, J., Maurer, H.A., Paun, G., Rozenberg, G., eds.: Jewels Are Forever. Springer (1999) 145–156
20. Ésik, Z., Kuich, W.: Formal tree series. J. Autom. Lang. Combin. **8**(2) (2003) 219–285
21. Borchardt, B.: The Theory of Recognizable Tree Series. PhD thesis, Technische Universität Dresden (2005)
22. Pech, C.: Kleene-Type Results for Weighted Tree-Automata. PhD thesis, Technische Universität Dresden (2003)
23. Engelfriet, J.: Tree automata and tree grammars. Technical Report DAIMI FN-10, Aarhus University (1975)
24. Kuich, W.: Tree transducers and formal tree series. Acta Cybernet. **14**(1) (1999) 135–149
25. Golan, J.S.: Semirings and their Applications. Kluwer Academic, Dordrecht (1999)
26. Maletti, A.: Relating tree series transducers and weighted tree automata. Int. J. Found. Comput. Sci. **16**(4) (2005) 723–741