The Myhill-Nerode Theorem for Recognizable Tree Series

Björn Borchardt*

Dresden University of Technology, Faculty of Computer Science, D-01062 Dresden.
borchard@tcs.inf.tu-dresden.de

Abstract. In this paper we prove a Myhill-Nerode theorem for recognizable tree series over commutative semifields and thereby present a minimization of bottom-up finite state weighted tree automata over a commutative semifield, where minimal means with respect to the number of states among all equivalent, deterministic devices.

1 Introduction

The concept of finite state weighted tree automata is a generalization of (i) finite state (unweighted) tree automata and (ii) finite state weighted (string) automata in the sense that trees are accepted with a weight, which is taken from a semiring. Let us briefly recall the two latter underlying models.

Trees can be traversed in two directions: from the leaves towards the root, which is modeled by bottom-up (also: frontier-to-root) finite state tree automata [22, 17, 5], or from the root towards the leaves, which is modeled by top-down (also: root-to-frontier) finite state tree automata [17, 19]. The class of recognizable tree languages is the class of tree languages accepted by these devices. For more details on finite state tree automata we refer the reader to [12, 13]. An algebraic characterization of recognizable tree languages is provided by the Myhill-Nerode theorem, which, as stated in [13] (Chapter II, Theorem 7.1, also cf. [10, 14]), says that for every tree language $L \subseteq T_\Sigma$ the following statements are equivalent: (i) L is recognizable (by a bottom-up finite state tree automaton), (ii) the term algebra $\mathcal{F}_\Sigma(X)$ has a congruence of finite index saturating L (i.e., L is the union of some equivalence classes), (iii) the index of the Nerode congruence $\equiv_L$ is finite (where for every $s, t \in T_\Sigma$, $s \equiv_L t$ iff for all trees $C \in T_\Sigma(\{x\})$ with exactly one occurrence of the variable x, the equivalence $C[x \leftarrow s] \in L \Leftrightarrow C[x \leftarrow t] \in L$ holds).

As the second underlying model, finite state weighted (string) automata [20] (also: string-to-weight transducers) arise from unweighted (string) automata by associating a weight with each of its transitions; the weight is taken from a semiring. For more details on this concept the reader may consult [16, 2, 18, 6, 7]. The second mentioned paper also investigates the minimization in terms of syntactic

* Research was financially supported by the German Research Council under grant (DFG, GRK 433/2).

algebras and proves that a power series is recognizable iff its syntactic algebra has finite rank (cf. [2], Chapter II, Theorem 1.1). Besides others [18] provides a minimization of deterministic weighted automata over the tropical semiring, where minimization means with respect to the number of states among all equivalent, deterministic devices. This construction is based on a generalization of the Myhill-Nerode congruence and of the Myhill-Nerode construction for weighted automata (cf. [18], Theorem 14).

As mentioned above, a finite state weighted tree automaton $M = (Q, \Sigma, Q_d, A, \mu)$ accepts trees with a weight, which is taken from a semiring $(A, \oplus, \odot, 0, 1)$. Similar to unweighted tree automata we distinguish between bottom-up (for short: bu-w-fta) and top-down devices. In this paper we will only consider the model bu-w-fta. Every transition has a weight, which is implemented in the tree representation $\mu = (\mu_k : \Sigma^{(k)} \to A^{Q^k \times Q} \mid k \geq 0)$. M can be entered by reading an arbitrary symbol α of rank 0 in every state q with the weight $\mu_0(\alpha)_{(),q}$. M can be left in a final state $q \in Q_d$. Now we let M run on a tree $t \in T_\Sigma$. We obtain the weight of a run on t by multiplying the weights of the applied transitions. The weight of t is the sum of the weights of all runs accepting t and the tree series accepted by M is a mapping $S_M \in A\langle\langle T_\Sigma \rangle\rangle$, where $A\langle\langle T_\Sigma \rangle\rangle = \{S \mid S : T_\Sigma \to A\}$ is the class of formal tree series over Σ and A. We formalize this procedure by a Σ-algebra, where we perform the summation arising from the nondeterminism at each step (compare [11] for a proof of the equivalence of automata-theoretic semantics, in which the summation is only performed on successful runs, and the initial algebra semantics, in which the summation is done at every transition step, in the more general context of tree series transducers). We note that there exist other weighted tree automata models, e.g. the concept of recognizable tree series of [1] (which if the underlying semiring is commutative, is equivalent to the concept of bu-w-fta), R-cost automata of [21], and A'-automata of [15] (cf. [3] for a discussion on the relationship between the latter two models and our model of weighted tree automata). As a survey paper on formal tree series we refer the reader to [9].

In the present paper we consider bu-w-fta as an extension of classical bottom-up finite state tree automata and generalize the Myhill-Nerode congruence and the Myhill-Nerode theorem as stated above. More precisely, we show that for every tree series S over a commutative semifield the following statements are equivalent (cf. Theorem 2):

(i) $S \in A\langle\langle T_\Sigma \rangle\rangle$ is recognizable by a deterministic bu-w-fta.
(ii) There exists a congruence relation $\equiv$ of the term algebra (T_Σ, Σ) of finite index satisfying Properties (MN1) and (MN2).
(iii) The Myhill-Nerode congruence $\equiv_S$ is of finite index.

We note that, by Theorem 6.2 of [3] if the underlying semiring is a locally finite semifield, then S is recognizable by a nondeterministic bu-w-fta iff S is recognizable by a deterministic bu-w-fta. Hence Theorem 2 gives algebraic characterizations for tree series which are accepted by (a) deterministic bu-w-fta over a commutative semifield and (b) nondeterministic bu-w-fta over a locally finite, commutative semifield. Roughly speaking, (MN1) and (MN2) are two technical conditions which, applied in the Boolean semiring and thereby considering unweighted tree automata, collapse to the classical property that "$\equiv$ saturates the tree language $supp(S)$". Hence Theorem 2 is a canonical generalization of the Myhill-Nerode theorem for tree languages. Similar to unweighted tree automata we derive from the Myhill-Nerode theorem a minimization for (a) deterministic bu-w-fta over a commutative semifield and (b) nondeterministic bu-w-fta over a locally finite, commutative semifield, where minimization means minimal with respect to the number of states among all deterministic, equivalent devices.

[4] considers weighted tree automata as a generalization of weighted (string) automata and proves the following characterization for a tree series S over a field (cf. [4], corollary on page 353): S is recognizable (in the sense of [1]) iff the syntactic algebra $\mathcal{A}_S$ is of finite dimension. Although the above stated theorem of Bozapalidis and Theorem 2 of the present paper seem to be very similar, none of these results implies the other one: we show that the tree series $(S,t) = size(t) \in Real\langle\langle T_\Sigma \rangle\rangle$ is recognizable by a nondeterministic bu-w-fta (cf. Example 1), but not by a deterministic bu-w-fta (cf. Example 3). Hence the syntactic algebra $\mathcal{A}_S$ is finite dimensional (by the corollary on page 353 of [4]), but the Myhill-Nerode congruence $\equiv_S$ is of infinite index (by Theorem 2 of this paper). We note that the characterization of [4] holds for tree series over a field, while our result only assumes a commutative semifield.

2 Preliminaries

2.1 Notions and Trees

We denote the set of all nonnegative integers by $\mathbb{N} = \{0, 1, \ldots\}$. For a set S, the cardinality and the power set of S are denoted by $card(H)$ and $\mathfrak{P}(S)$, respectively. If Σ is a nonempty finite set and $rk : \Sigma \to \mathbb{N}$, then the tuple (Σ, rk) is called *ranked alphabet*. In order to be short in notation we will write Σ rather than (Σ, rk). For every $k \geq 0$ we define the set $\Sigma^{(k)} = \{\sigma \in \Sigma \mid rk(\sigma) = k\}$ of all symbols of Σ, which have rank k. An element $\sigma \in \Sigma^{(k)}$ is also written as $\sigma^{(k)}$.

For the rest of this paper let Σ be a ranked alphabet. Moreover, let X be a set (of variables) disjoint with Σ. The set of *t*rees over Σ indexed by X, denoted by $T_\Sigma(X)$, is the smallest set $T_\Sigma(X) \subseteq (\Sigma \cup X \cup \{(,),,\})^*$ such that (i) $X \cup \Sigma^{(0)} \subseteq T_\Sigma(X)$ and (ii) if $k \geq 1$, $\sigma \in \Sigma^{(k)}$, and $s_1, \ldots, s_k \in T_\Sigma(X)$, then $\sigma(s_1, \ldots, s_k) \in T_\Sigma(X)$. The set $T_\Sigma(\emptyset)$ is denoted by T_Σ. Clearly, $T_\Sigma(X) = T_{\Sigma \cup X}$. The *size* of a tree $t \in T_\Sigma(X)$ is given by the mapping $size : T_\Sigma(X) \to \mathbb{N}$, where $size(\sigma(s_1, \ldots, s_k)) = \sum_{i=1}^{k} size(s_i) + 1$ for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $s_1, \ldots, s_k \in T_\Sigma$. Now let $x_1, x_2, \ldots$ be variables and $X_m = \{x_1, \ldots, x_m\}$ for every $m \geq 0$. Let $t \in T_\Sigma(X_m)$ for some $m \geq 0$. The *substitution of* $x_1, \ldots, x_m$ *by* $s_1, \ldots, s_m \in T_\Sigma(X_m)$ *in* t, denoted by $t[s_1, \ldots, s_m]$, is the tree, which we obtain from t by replacing every occurrence of x_j in t by s_j for every $1 \leq j \leq m$. Moreover, a tree $t \in T_\Sigma(X_m)$ is called a *Σ-m-context* if every variable $x \in X_m$ occurs precisely once in t. We denote the class of all Σ-m-contexts by $C_{\Sigma,m}(X_m)$.

2.2 Algebraic Notions

In this section we introduce semirings, semifields, term algebras, matrices, and equivalence and congruence relations (also cf. [16]).

Let A be a nonempty set, $0, 1 \in A$, and $\oplus$ and $\odot$ binary, associative operations on A. The tuple $(A, \oplus, \odot, 0, 1)$ is called *semiring* if (SR1) $\oplus$ is commutative, (SR2) 0 is the neural element of $\oplus$, (SR3) 1 is the neural element of $\odot$, (SR4) $\odot$ is distributive over $\oplus$, and (SR5) $0 \odot a = a \odot 0 = 0$ for every $a \in A$. Whenever $\oplus$, $\odot$, 0, and 1 are clear from the context, then we denote the semiring $(A, \oplus, \odot, 0, 1)$ by A. We call a semiring *commutative* if $\odot$ is commutative. A is *zero divisor free* if for every $a, b \in A$, $a \odot b = 0$ implies $a = 0$ or $b = 0$.

Let $^{-1}$ be a unary operation on $A \setminus \{0\}$. The tuple $(A, \oplus, \odot, ^{-1}, 0, 1)$ is called *semifield* if (SF1) $(A, \oplus, \odot, 0, 1)$ is a semiring and, (SF2) for every $a \in A \setminus \{0\}$, a^{-1} is the multiplicative inverse of a. Whenever it does not lead to confusion, we will write A rather than $(A, \oplus, \odot, ^{-1}, 0, 1)$. Let $A = (A, \oplus, \odot, ^{-1}, 0, 1)$ be a semifield. A is called *commutative* (*zero divisor free*) if $(A, \oplus, \odot, 0, 1)$ is a commutative (zero divisor free, respectively) semiring. A is called *locally finite* if, for every finite $B \subseteq A$, there exists a semifield $(\tilde{B}, \oplus, \odot, ^{-1}, 0, 1)$ such that $B \subseteq \tilde{B}$ and $\tilde{B}$ is finite (in other words: the closure of B under $\{\oplus, \odot, ^{-1}\}$ is finite). Typical (commutative) semifields are

- the *Boolean semifield* $\mathbb{B} = (\{0, 1\}, \vee, \wedge, ^{-1}, 0, 1)$, where the operations are conjunction and disjunction, and $1^{-1} = 1$,
- the *tropical semifield* $Trop = (\mathbb{R} \cup \{+\infty\}, min, +, ^{-1}, +\infty, 0)$ with minimum and addition, where $min(a, +\infty) = a$ and $a + (+\infty) = (+\infty) + a = +\infty$ for every $a \in \mathbb{R} \cup \{+\infty\}$, and the inverse is defined for every $a \in \mathbb{R}$ to be $-a$,
- the *real numbers* $Real = (\mathbb{R}, +, \cdot, ^{-1}, 0, 1)$ with the usual operations,

By standard arguments one proves the following lemma.

Lemma 1. *Every semifield is zero divisor free.*

Now consider an arbitrary ranked alphabet Σ and let $k \geq 0$. The *term algebra (with respect to Σ)* is the algebra (T_Σ, Σ), where every $\sigma \in \Sigma^{(k)}$ is identified with the mapping $\sigma : {T_\Sigma}^k \to T_\Sigma$ (*top concatenation with σ*), which is defined for every $t_1, \ldots, t_k \in T_\Sigma$ to be the tree $\sigma(t_1, \ldots, t_k)$.

We now recall some definitions concerning matrices. An *$I \times J$-matrix* is a mapping $M : I \times J \to A$, where A, I, and J are arbitrary sets. As usual, $M(i, j)$ is also denoted by $M_{i,j}$. The set of all $I \times J$-matrices over A is denoted by $A^{I \times J}$. If either $I = \{i\}$ or $J = \{j\}$ are singletons, then we write A^J and M_j (A^I and M_i, respectively) rather than $A^{I \times J}$ and $M_{i,j}$.

Let us now recall some definitions concerning relations. Let S be a set. A *binary relation ρ (on S)* is a subset $\rho \subseteq S \times S$. If $(s, t) \in \rho$, then we also write $s \, \rho \, t$. We call a *reflexive* ($(a, a) \in \rho$ for every $a \in S$), *symmetric* ($(a, b) \in \rho$ implies $(b, a) \in \rho$), and *transitive* ($(a, b), (b, c) \in \rho$ implies $(a, c) \in \rho$) relation ρ an *equivalence relation*.

Let ρ be an equivalence relation and $a \in S$. We call the set $[a]_\rho = \{b \in S \mid (a, b) \in \rho\}$ the *equivalence class of a (with respect to ρ)*. The set of all equivalence classes is denoted by $S/_\rho = \{[a]_\rho \mid a \in S\}$ and the *index of* ρ is defined to be the cardinality of $S/_\rho$. A set $S' \subseteq S$ is *saturated or respected by* ρ if S is the union of some equivalence classes of ρ. Moreover, a *representative mapping (with respect to* ρ*)* is a mapping $\varphi : S/_\rho \to S$ such that $\varphi([a]_\rho) \in [a]_\rho$ for every $[a]_\rho \in S/_\rho$. We denote the *set of all representative mappings with respect to* ρ by $RepMap(\rho)$. Clearly, $RepMap(\rho) \neq \emptyset$.

Now consider a ranked alphabet Σ and an equivalence relation $\rho \subseteq T_\Sigma \times T_\Sigma$. ρ is a *congruence relation (of the term algebra* (T_Σ, Σ)*)* if for every $k \geq 0$, $s \in \Sigma^{(k)}$, and $s_1, \ldots, s_k, t_1, \ldots, t_k \in T_\Sigma$, $s_1 \, \rho \, t_1, \ldots, s_k \, \rho \, t_k$ implies $\sigma(s_1, \ldots, s_k) \, \rho \, \sigma(t_1, \ldots, t_k)$.

2.3 Formal Tree Series

Let A be a semiring and Σ be a ranked alphabet. A *(formal) tree series (over* Σ *and* A*)* is a mapping $S : T_\Sigma \to A$. The image $S(t) \in A$ of $t \in T_\Sigma$ is called *coefficient* of t and we write (S, t) rather than $S(t)$. The tree series S can be written as the sum $\sum_{t \in T}(S, t)\ t$ (cf. [8], Chapter VI, 3.1). The *support of* S is defined to be the set $supp(S) = \{t \in T_\Sigma \mid (S, t) \neq 0\}$.

3 Bottom-Up Finite State Weighted Tree Automata

We now briefly recall the model of bottom-up finite state weighted tree automata. For more details we refer the reader to [3].

Definition 1 ([3], Definition 3.1). *Let* Q *be a finite set,* Σ *a ranked alphabet, and* A *a semiring. A* bottom-up tree representation (over Q, Σ, and A) *is a family* $\mu = (\mu_k \mid k \geq 0)$ *of mappings* $\mu_k : \Sigma^{(k)} \longrightarrow A^{Q^k \times Q}$. μ *is called* deterministic *if for every* $k \geq 0$, $\sigma \in \Sigma^{(k)}$, *and* $(q_1, \ldots, q_k) \in Q^k$ *there is at most one* $q \in Q$ *such that* $\mu_k(\sigma)_{(q_1,\ldots,q_k),q} \neq 0$.

Every tree representation μ over a finite set Q induces a family of mappings $(\overline{\mu_k(\sigma)} : A^Q \times \cdots \times A^Q \to A^Q \mid k \geq 0, \sigma \in \Sigma^{(k)})$ in the following way: for every $q \in Q$ and $V_1, \ldots, V_K \in A^Q$,

$$\overline{\mu_k(\sigma)}(V_1, \ldots, V_k)_q = \sum_{(q_1,\ldots,q_k) \in Q^k} (V_1)_{q_1} \odot \cdots \odot (V_k)_{q_k} \odot \mu_k(\sigma)_{(q_1,\ldots,q_k),q}.$$

Clearly, $\left(A^Q, \left(\overline{\mu_k(\sigma)} \mid k \geq 0, \sigma \in \Sigma^{(k)}\right)\right)$ is a Σ-algebra. Its initial homomorphism $h_\mu : T_\Sigma \to A^Q$ is given for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $t_1, \ldots, t_k \in T_\Sigma$ by $h_\mu(\sigma(t_1, \ldots t_k)) = \overline{\mu_k(\sigma)}(h_\mu(t_1), \ldots, h_\mu(t_k))$. We call $h_\mu(t)$ the *characteristic vector of* t *(with respect to* μ*)*. Let us now define bottom-up finite state weighted tree automata.

Definition 2 ([3], Definition 3.3). *Let* Q *be a finite set,* Σ *a ranked alphabet,* $Q_d \subseteq Q$, A *a semiring, and* μ *a bottom-up tree representation. The tuple* $M = (Q, \Sigma, Q_d, A, \mu)$ *is called* bottom-up finite state weighted tree automaton (for short bu-w-fta). *A bu-w-fta is called* deterministic *if its tree representation is deterministic. The tree series S_M* accepted *by M is defined for every $t \in T_\Sigma$ by $(S_M, t) = \sum_{q \in Q_d} h_\mu(t)_q$. We denote by $A^{n,bu}\langle\!\langle T_\Sigma \rangle\!\rangle$ and $A^{d,bu}\langle\!\langle T_\Sigma \rangle\!\rangle$ the classes of all tree series, which are accepted by a bu-w-fta and a deterministic bu-w-fta, respectively.*

Example 1. Consider the bu-w-fta $M = (Q, \Sigma, Q_d, A, \mu)$ given by $Q = \{p, q\}$, $\Sigma = \{\alpha^{(0)}, \sigma^{(2)}\}$, $Q_d = \{p\}$, $A = Real$, and

$$\mu_0(a) = \begin{array}{c} \begin{array}{cc} p & q \end{array} \\ \begin{pmatrix} 1 & 1 \end{pmatrix} \end{array}, \qquad \mu_2(\sigma) = \begin{array}{cc} \begin{array}{cc} p & q \end{array} & \\ \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} & \begin{array}{c} (p,p) \\ (p,q) \\ (q,p) \\ (q,q) \end{array} \end{array}.$$

One can easily show that $h_\mu(t)_p = size(t)$ and $h_\mu(t)_q = 1$ for every $t \in T_\Sigma$. Thus every input tree is accepted by M with its size: $(S_M, t) = size(t)$. $\diamond$

Later on, for a given $t \in T_\Sigma$, we work with the set of all those states $q \in Q$ such that there is a "non zero derivation" of t in M ending up in q. Formally, let $\hat{\mu} : T_\Sigma \to \mathfrak{P}(Q)$ be the mapping, which is defined for every $t \in T_\Sigma$ by $\hat{\mu}(t) = \{q \in Q \mid h_\mu(t)_q \neq 0\}$. We conclude this section by presenting three properties of the mapping $\hat{\mu}$.

Lemma 2 ([3], Lemma 3.6). *Let $M = (Q, \Sigma, Q_d, A, \mu)$ be a deterministic bu-w-fta. Then $card(\hat{\mu}(t)) \leq 1$ for every tree $t \in T_\Sigma$. If $\hat{\mu}(t) = \{q\}$ for some $q \in Q$, then we write $\hat{\mu}(t) = q$.*

Lemma 3 ([3], Lemma 3.7). *Let $M = (Q, \Sigma, Q_d, A, \mu)$ be a bu-w-fta. Also, let $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $t_1, \ldots, t_k \in T_\Sigma$. If there exists an $1 \leq i \leq k$ such that $\hat{\mu}(t_i) = \emptyset$, then $\hat{\mu}(\sigma(t_1, \ldots, t_k)) = \emptyset$.*

By a repeated application of Lemma 3 we obtain the following corollary:

Corollary 1. *Let $M = (Q, \Sigma, Q_d, A, \mu)$ be a bu-w-fta. Moreover, let $m \geq 0$, $s_1, \ldots, s_m \in T_\Sigma$, and $t \in T_\Sigma(X_m)$. If $\hat{\mu}(s_j) = \emptyset$ for some $1 \leq j \leq m$, then $\hat{\mu}(t[s_1, \ldots, s_m]) = \emptyset$.*

4 Substitution and Weights

In this section we investigate the behavior of the characteristic vectors if we substitute variables by trees. In particular, we consider a finite set X_m of variables, a tree $t \in T_\Sigma(X_m)$, and trees $s_1, \ldots, s_m \in T_\Sigma$ and show a relationship between the characteristic vectors of t, $s_1, \ldots, s_m$, and $t[s_1, \ldots, s_m]$. Intuitively a run on $t[s_1, \ldots, s_m]$ can be decomposed into runs on each of the copies of s_j, which are substituted in $t[s_1, \ldots, s_m]$, and a run on t. For our purpose it suffices to prove the relationship under the assumption that, for every $1 \leq j \leq m$, the non zero runs on the copies of s_j end up in the same state q_j. This additional requirement we can ensure by assuming that M is a deterministic bu-w-fta. Then the weight of a run on $t[s_1, \ldots, s_m]$ is the product of the weight of the run on $t[q_1, \ldots, q_m] \in T_\Sigma(Q)$ and the powers (according to the number of occurrences of x_j) of the weights of the runs on $s_1, \ldots, s_m$. In order to let an automaton run on $t[q_1, \ldots, q_m]$ we extend the given bu-w-fta.

Definition 3. *Let $M = (Q, \Sigma, Q_d, A, \mu)$ be a bu-w-fta. The* extension of M (with respect to Q) *is the bu-w-fta $M_Q = (Q, \Sigma \cup \{q^{(0)} \mid q \in Q\}, Q_d, A, \mu_Q)$, where for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $p, q, q_1, \ldots, q_k \in Q$,*

$$(\mu_Q)_0(p)_q = \begin{cases} 1 & \text{if } p = q, \\ 0 & \text{otherwise,} \end{cases}$$
$$(\mu_Q)_k(\sigma)_{(q_1,\ldots,q_k),q} = \mu_k(\sigma)_{(q_1,\ldots,q_k),q}.$$

Let us now prove the relationship between the characteristic vectors of $s_1, \ldots, s_m \in T_\Sigma$, $t[s_1, \ldots, s_m] \in T_\Sigma$, and $t[q_1, \ldots, q_m] \in T_\Sigma(Q)$ with respect to a deterministic bottom-up tree representation.

Theorem 1. *Let $M = (Q, \Sigma, Q_d, A, \mu)$ be a deterministic bu-w-fta over a commutative semiring, M_Q the exention of M with respect to Q, and $t \in T_\Sigma(X_m)$ for some $m \geq 0$. Moreover, let ε_j denote the number of occurences of x_j in t for every $1 \leq j \leq m$. Then, for every $s_1, \ldots, s_m \in T_\Sigma$ and $q \in Q$,*

$$h_\mu(t[s_1, \ldots, s_m])_q = \begin{cases} h_{\mu_Q}(t[\hat{\mu}(s_1), \ldots, \hat{\mu}(s_m)])_q \odot \prod_{j=1}^{m} h_\mu(s_j)_{\hat{\mu}(s_j)}{}^{\varepsilon_j} \\ \qquad \text{if } \hat{\mu}(s_j) \neq \emptyset \text{ for every } 1 \leq j \leq m, \\ 0 \qquad \text{otherwise.} \end{cases}$$

We note that $h_\mu(t[s_1, \ldots, s_m])$ can also be expressed by the result of the o-substitution of tree series (cf. [11], Definition 3.2): the monomial $d\ t[s_1, \ldots, s_m]$ with $d = h_\mu(t[s_1, \ldots, s_m])_q$ is equal to

$$c\ t \xleftarrow{o} (c_1\ s_1, \ldots, c_m\ s_m)$$

where $c = h_{\mu_Q}(t[\hat{\mu}(s_1), \ldots, \hat{\mu}(s_m)])$ and $c_j = h_\mu(s_j)_{\hat{\mu}(s_j)}$ for every $1 \leq j \leq m$.

Corollary 2. *Let $M = (Q, \Sigma, Q_d, A, \mu)$ be a deterministic bu-w-fta over a zero divisor free, commutative semiring and $t \in T_\Sigma(X_m)$ for some $m \geq 0$. Also, let $s_1, s_1', \ldots, s_m, s_m' \in T_\Sigma$ such that $\hat{\mu}(s_1) = \hat{\mu}(s_1'), \ldots, \hat{\mu}(s_m) = \hat{\mu}(s_m')$. Then $\hat{\mu}(t[s_1, \ldots, s_m]) = \hat{\mu}(t[s_1', \ldots, s_m'])$.*

5 The Myhill-Nerode Theorem for Tree Series over Commutative Semifields

In this section we generalize the Myhill-Nerode theorem from tree languages to tree series. For every tree series $S \in A\langle\langle T_\Sigma \rangle\rangle$ over a commutative semifield A we introduce a congruence relation $\equiv_S$ of the term algebra (T_Σ, Σ), which extends the Myhill-Nerode congruence for tree languages in a natural way. $\equiv_S$

can be used to characterize tree series, which are accepted by deterministic bu-w-fta over a commutative semifield. We show that the following statements are equivalent: (i) $S \in A^{d,bu}$, (ii) there exists a congruence relation of the underlying term algebra which has finite index and satisfies Properties (MN1) and (MN2) (which we define on page 9), and (iii) $\equiv_S$ is of finite index. Similar to the proof of the classical Myhill-Nerode theorem we show the implications (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i). From the above equivalences we obtain a characterization for tree series $S \in A^{n,bu}$ over a locally finite, commutative semifield by applying the equality $A^{d,bu}\langle\langle T_\Sigma \rangle\rangle = A^{n,bu}\langle\langle T_\Sigma \rangle\rangle$ (cf. Theorem 6.2 of [3]). Finally we present a minimization of (a) deterministic bu-w-fta over a commutative semifield and (b) nondeterministic bu-w-fta over a locally finite, commutative semifield, which is a straightforward application of the Myhill-Nerode theorem. We note that minimization means minimal with respect to the number of states among all equivalent, deterministic devices.

We also note that the proof of (ii) $\Rightarrow$ (i) (Myhill-Nerode construction) requires multiplicative inverses, which in general do not exist in a semiring. Hence, for the rest of this section let A be a commutative semifield. Also, let $S \in A\langle\langle T_\Sigma \rangle\rangle$. Recall that $C_{\Sigma,1}(X_1)$ denotes the class of all Σ-1-contexts, i.e., the class of all trees over $\Sigma \cup X_1$ with precisely one occurrence of the variable x_1. We start by defining the binary relation $\equiv_S \subseteq T_\Sigma \times T_\Sigma$ by

$$\equiv_S = \{(s,t) \in T_\Sigma \times T_\Sigma \mid \exists a \in A \setminus \{0\} \forall C \in C_{\Sigma,1}(X_1) : (S, C[s]) = a \odot (S, C[t])\}.$$

Let us now investigate $\equiv_S$. In particular, we show that $\equiv_S$ is a congruence relation of the term algebra (T_Σ, Σ), which is stated in Lemma 5.

Lemma 4. *Let $k \geq 0$, $\sigma \in \Sigma^{(k)}$, $s_1, \ldots, s_k, t_1, \ldots, t_k \in T_\Sigma$, and $a_1, \ldots, a_k \in A \setminus \{0\}$ such that $s_i \equiv_S t_i$ and $(S, C[s_i]) = a_i \odot (S, C[t_i])$ for every $1 \leq i \leq k$ and $C \in C_{\Sigma,1}(X_1)$. Then $\big(S, C\big[\sigma(s_1, \ldots, s_k)\big]\big) = a_1 \odot \cdots \odot a_k \odot \big(S, C\big[\sigma(t_1, \ldots, t_k)\big]\big)$.*

Clearly, $\equiv_S$ is an equivalence relation. From Lemma 4 follows that $\equiv_S$ is also a congruence relation of the term algebra.

Lemma 5. *$\equiv_S$ is a congruence relation of the term algebra (T_Σ, Σ).*

Let us now prove that, if S is recognizable by a deterministic bu-w-fta M, then $\equiv_S$ is of finite index. This we finally show in Corollary 3. We start the proof of the aforementioned statement by showing a tight relationship between $\equiv_S$ and the state behaviour of M. Recall that the tree series, which is recognized by M, is denoted by S_M.

Lemma 6. *Let $M = (Q, \Sigma, Q_d, A, \mu)$ be a deterministic bu-w-fta such that $S_M = S$. Also, let $s, t \in T_\Sigma$. If $\hat{\mu}(s) = \hat{\mu}(t)$, then $s \equiv_S t$.*

Does the claim of Lemma 6 also hold for nondeterministic bu-w-fta? The answer is no, which follows from Example 2.

Example 2. In Example 1 we have presented a nondeterministic bu-w-fta M over the commuatative semifield *Real* with $(S_M, t) = size(t)$ and $\hat{\mu}(t) = \{p, q\}$ for every $t \in T_\Sigma$. In particular, $\hat{\mu}(\alpha) = \{p, q\} = \hat{\mu}(\sigma(\alpha, \alpha))$. On the other hand, by applying the equality $size(C[s]) = size(s) + size(C) - 1$ for every $s \in T_\Sigma$ and $C \in C_{\Sigma,1}(X_1)$ which can be shown by induction on $size(C[t])$, one can easily prove that $s \equiv_{S_M} t$ iff $size(s) = size(t)$ for every $s, t \in T_\Sigma$. From $size(\alpha) = 1 \neq 3 = size(\sigma(\alpha, \alpha))$ follows that $\alpha \not\equiv_{S_M} \sigma(\alpha, \alpha)$ even though $\hat{\mu}(\alpha) = \hat{\mu}(\sigma(\alpha, \alpha))$. $\diamond$

In order to obtain an upper bound for the index of $\equiv_S$ we introduce and investigate the set $L_S = \{s \in T_\Sigma \mid \forall C \in C_{\Sigma,1}(X_1)\ (S, C[s]) = 0)\}$.

Observation 1. *Either L_S is the empty set or it is an equivalence class of $\equiv_S$.*

Lemma 7. *Let $M = (Q, \Sigma, Q_d, A, \mu)$ be a deterministic bu-w-fta such that $S_M = S$. Then $\hat{\mu}(t) \in Q$ for every $t \in T_\Sigma \setminus L_S$.*

Corollary 3. *For every deterministic bu-w-fta M over A such that $S_M = S$, $index(\equiv_S) \leq card(Q) + 1$. Moreover, if $L_S = \emptyset$ then $index(\equiv_S) \leq card(Q)$. In particular, if $S \in A\langle\langle T_\Sigma \rangle\rangle^{d,bu}$, then $\equiv_S$ is of finite index.*

Now let us prove that if $\equiv_S$ has finite index, then there exists a congruence relation of the term algebra (T_Σ, Σ) of finite index satisfying Properties (MN1) and (MN2) (which we define on page 9). We prove this claim by showing that the particular relation $\equiv_S$ has these properties, which is stated in Lemmata 5 and 11. In order to define Property (MN1) and the extension of the Myhill-Nerode construction (cf. Definition 4) we would like to work with multiplicative inverses of (S, t). But in general there exists a tree $t \in T_\Sigma$ such that $(S, t) = 0$ and then we can not take the inverse of (S, t). Therefore we define a new tree series $S^\varphi \in A\langle\langle T_\Sigma \rangle\rangle$, where φ is an arbitrary representative mapping of $\equiv_S$. This definition requires some preliminary considerations.

Lemma 8. *Let $\varphi \in RepMap(\equiv_S)$. There exists precisely one mapping $a_\varphi : T_\Sigma \to A \setminus \{0\}$, which satisfies the following two properties.*

(i) For every $t \in T_\Sigma$ and $C \in C_{\Sigma,1}(X_1)$ the equation $(S, C[t]) = a_\varphi(t) \odot (S, C[\varphi([t]_{\equiv_S})])$ holds.
(ii) For every $t \in L_S$ we have $a_\varphi(t) = 1$.

For the rest of this section let $\varphi \in RepMap(\equiv_S)$ and $a_\varphi : T_\Sigma \to A \setminus \{0\}$ the uniquely determined mapping, which satisfies Properties (i) and (ii) of Lemma 8. Now we define the tree series $S^\varphi \in A\langle\langle T_\Sigma \rangle\rangle$, which is given for every $t \in T_\Sigma$ by

$$(S^\varphi, t) = \begin{cases} (S, t) & \text{if } (S, t) \neq 0, \\ a_\varphi(t) & \text{otherwise.} \end{cases}$$

We note that $(S^\varphi, t) \neq 0$ for every $t \in T_\Sigma$, because $a_\varphi(t) \in A \setminus \{0\}$. Recall that in the Myhill-Nerode theorem for tree languages one of the characterizations of a recognizable tree language L is that there exists a congruence relation which saturates L. Since we additionally have to handle the weights of the trees, this property has to be generalized appropriately. Therefore we introduce the following properties for equivalence relations $\equiv \subseteq T_\Sigma \times T_\Sigma$:

(MN1) For every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $s_1, \ldots, s_k, t_1, \ldots, t_k \in T_\Sigma$ such that $\sigma(t_1, \ldots, t_k) \notin L_S$ and $s_i \equiv t_i$ for every $1 \leq i \leq k$,

$$\begin{aligned}&(S^\varphi, s_k)^{-1} \odot \cdots \odot (S^\varphi, s_1)^{-1} \odot (S^\varphi, \sigma(s_1, \ldots, s_k))\\ &= (S^\varphi, t_k)^{-1} \odot \cdots \odot (S^\varphi, t_1)^{-1} \odot (S^\varphi, \sigma(t_1, \ldots, t_k)).\end{aligned}$$

(MN2) The relation $\equiv$ respects $supp(S)$, i.e., for every $s, t \in T_\Sigma$ with $s \equiv_S t$, $s \in supp(S)$ implies $t \in supp(S)$.

We note that if the underlying semifield is the Boolean semifield and thereby considering unweighted tree automata, then (MN1) and (MN2) are equivalent to the classical property "$\equiv$ respects $supp(S)$". Let us now prove that $\equiv_S$ satisfies (MN1) and (MN2). This requires the following two lemmata.

Lemma 9. *Let $s, t \in T_\Sigma$. If $s \equiv_S t$, then (i) $(S, s) = a_\varphi(s) \odot a_\varphi(t)^{-1} \odot (S, t)$ and (ii) $(S^\varphi, s) = a_\varphi(s) \odot a_\varphi(t)^{-1} \odot (S^\varphi, t)$.*

Lemma 10. *Let $s, t \in T_\Sigma$ such that $s = \sigma(s_1, \ldots, s_k)$ and $t = \sigma(t_1, \ldots, t_k) \notin L_S$ for some $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $s_1, \ldots, s_k, t_1, \ldots, t_k \in T_\Sigma$. If $s_i \equiv_S t_i$ for every $1 \leq i \leq k$, then $a_\varphi(s) \odot a_\varphi(t)^{-1} = a_\varphi(s_1) \odot a_\varphi(t_1)^{-1} \odot \cdots \odot a_\varphi(s_k) \odot a_\varphi(t_k)^{-1}$.*

Lemma 11. *$\equiv_S$ satisfies (MN1) and (MN2).*

Now assume that there is a congruence relation of the term algebra (T_Σ, Σ) of finite index satisfying (MN1) and (MN2). In the following we present a deterministic bu-w-fta, which accepts S (cf. Definition 4 and Lemma 13).

Definition 4. *Let $\equiv \subseteq T_\Sigma \times T_\Sigma$ be a congruence relation of the term algebra (T_Σ, Σ) of finite index, which satisfies (MN1) and (MN2). The* bu-w-fta associated with $\equiv$ *is the bu-w-fta $M = (Q, \Sigma, Q_d, A, \mu)$, where $Q = (T_\Sigma/_\equiv) \setminus \{L_S\}$, $Q_d = \{[s]_\equiv \in Q \mid \forall t \in [s]_\equiv\ t \in supp(S)\}$, and for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $[s_1]_\equiv, \ldots, [s_k]_\equiv, [s]_\equiv \in Q$,*

$$\mu_k(\sigma)_{([s_1]_\equiv, \ldots, [s_k]_\equiv), [s]_\equiv} = \begin{cases} (S^\varphi, s_k)^{-1} \odot \cdots \odot (S^\varphi, s_1)^{-1} \odot (S^\varphi, s) \\ \qquad \text{if } [s]_\equiv = [\sigma(s_1, \ldots, s_k)]_\equiv, \\ 0 \quad \text{otherwise.} \end{cases}$$

We note that the tree representation μ of the bu-w-fta associated with $\equiv$ is independent from the choice of the representatives of the equivalence classes, because $\equiv$ satisfies (MN1). Hence M is well defined.

Lemma 12. *Let $\equiv \subseteq T_\Sigma \times T_\Sigma$ as in Definition 4 and μ be the tree representation of the bu-w-fta associated with $\equiv$. Then, for every $t \in T_\Sigma$,*

$$\text{(i)}\ \ \hat{\mu}(t) = \begin{cases} \emptyset & \text{if } t \in L_S, \\ [t]_\equiv & \text{otherwise,} \end{cases}$$

(ii) if $t \notin L_S$, then $h_\mu(t)_{[t]_\equiv} = (S^\varphi, t)$.

Lemma 13. *Let $\equiv \subseteq T_\Sigma \times T_\Sigma$ be as in Definition 4. The bu-w-fta associated with $\equiv$ is a deterministic bu-w-fta recognizing S.*

Let us now sum up the results of this section. We thereby obtain a Myhill-Nerode theorem for recognizable tree series.

Theorem 2. *Let A be a commutative semifield and $S \in A\langle\langle T_\Sigma \rangle\rangle$. The following statements are equivalent:*

(i) $S \in A^{d,bu}\langle\langle T_\Sigma \rangle\rangle$.
(ii) There exists a congruence relation of the term algebra (T_Σ, Σ) of finite index satisfying Properties (MN1) and (MN2).
(iii) The relation $\equiv_S$ is of finite index.

If A is a locally finite, commutative semifield, then the following statement and (i), (ii), and (iii) are equivalent:

(i)' $S \in A^{n,bu}\langle\langle T_\Sigma \rangle\rangle$.

We note that the equivalence of Statements (i) and (i)' of Theorem 2 was shown in [3] (Theorem 6.2). Similar to classical tree automata we obtain a minimization for (a) deterministic bu-w-fta over a commutative semifield or (b) nondeterministic bu-w-fta over a locally finite, commutative semifield by applying the Myhill-Nerode construction with the Myhill-Nerode congruence. Recall that minimal means with respect to the number of states among all equivalent, deterministic devices.

Theorem 3. *If (a) A is a commutative semifield and $S \in A^{d,bu}\langle\langle T_\Sigma \rangle\rangle$ or if (b) A is a locally finite, commutative semifield and $S \in A^{n,bu}\langle\langle T_\Sigma \rangle\rangle$, then $M_{\equiv_S}$ recognizes S. Moreover, $M_{\equiv_S}$ is minimal with respect to the number of states among all equivalent, deterministic bu-w-fta.*

Example 3. In Example 1 we have presented a nondeterministic bu-w-fta M with semantics $(S, t) = size(t)$ for every $t \in T_\Sigma$, where S is a tree series over the commutative semifield $Real$. By applying Theorem 2 we show that there does not exist a deterministic bu-w-fta over $Real$ accepting S_M. As mentioned in Example 2, $s \equiv_S t$ iff $size(s) = size(t)$ for every $s, t \in T_\Sigma$. Hence $\equiv_S$ is of infinite index and thus S is not accepted by a deterministic bu-w-fta over $Real$.

On the other hand the size of a tree can be generated by a deterministic bu-w-fta over the tropical semifield $Trop$. Let $S' \in Trop\langle\langle T_\Sigma \rangle\rangle$ such that $(S', t) = size(t)$ for every $t \in T_\Sigma$. By applying the equality $size(C[s]) = size(s) + size(C) - 1$ for every $s \in T_\Sigma$ and $C \in C_{\Sigma,1}(X_1)$ one can easily show that $s \equiv_{S'} t$ for every $s, t \in T_\Sigma$. Hence $index(\equiv_{S'}) = 1$ and thus, by Theorem 2, $S' \in Trop^{d,bu}\langle\langle T_\Sigma \rangle\rangle$. Let us now construct a minimal bu-w-fta over $Trop$, which recognizes S'. Theorem 3 states that the bu-w-fta $M_{\equiv_{S'}} = (Q', \Sigma, Q'_d, A, \mu)$ is minimal with respect to the number of states, where $Q' = Q'_d = \{[\alpha]_{\equiv_{S'}}\}$, $\mu'_0(\alpha) = (1)$, and $\mu'_2(\sigma) = (1)$. ◇

References

[1] J. Berstel and C. Reutenauer. Recognizable formal power series on trees. *Theoretical Computer Science*, 18:115–148, 1982.

[2] J. Berstel and Ch. Reutenauer. *Rational Series and Their Languages*, volume 12 of *EATCS-Monographs*. Springer Verlag, 1988.

[3] B. Borchardt and H. Vogler. Determinization of Finite State Weighted Tree Automata. *J. Automata, Languages and Combinatorics, accepted*, 2003.

[4] S. Bozapalidis. Effective constuction of the syntactic algebra of a recognizable series on trees. *Acta Informatica*, 28:351–363, 1991.

[5] J. Doner. Tree acceptors and some of their applications. *J. Comput. System Sci.*, 4:406–451, 1970.

[6] M. Droste and P. Gastin. The Kleene-Schützenberger theorem for formal power series in partially commuting variables. *Information and Computation*, 153:47–80, 1999. extended abstract in: *24th ICALP*, LNCS vol. 1256, Springer, 1997, pp. 682-692.

[7] M. Droste and D. Kuske. Skew and infinitary formal power series. Technical Report 2002-38, Department of Mathematics and Computer Science, University of Leicester, 2002.

[8] S. Eilenberg. *Automata, Languages, and Machines, Vol.A*. Academic Press, 1974.

[9] Z. Esik and W. Kuich. Formal Tree Series. *J. Automata, Languages and Combinatorics, accepted*, 2003.

[10] Z. Fülöp and S. Vágvölgyi. Congruential tree languages are the same as recognizable tree languages. *Bulletin of EATCS*, 30:175–185, 1989.

[11] Z. Fülöp and H. Vogler. Tree series transformations that respect copying. *Theory of Computing Systems*, to appear, 2003.

[12] F. Gécseg and M. Steinby. *Tree Automata*. Akadémiai Kiadó, Budapest, 1984.

[13] F. Gécseg and M. Steinby. Tree languages. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3, chapter 1, pages 1–68. Springer-Verlag, 1997.

[14] D. Kozen. On the Myhill-Nerode theorem for trees. *EATCS Bulletin*, 47:170–173, June 1992.

[15] W. Kuich. Formal power series over trees. In S. Bozapalidis, editor, *Proc. of the 3rd International Conference Developments in Language Theory*, pages 61–101. Aristotle University of Thessaloniki, 1998.

[16] W. Kuich and A. Salomaa. *Semirings, Automata, Languages*. EATCS Monographs on Theoretical Computer Science, Springer Verlag, 1986.

[17] M. Magidor and G. Moran. Finite automata over finite trees. Technical Report 30, Hebrew University, Jerusalem, 1969.

[18] M. Mohri. Finite-state transducers in language and speech processing. *Computational Linguistics*, 23:269–311 (1–42), 1997.

[19] M. O. Rabin. Decidability of second-order theories and automata in infinite trees. *Transactions of the Amer. Math. Soc.*, 141:1–35, 1969.

[20] M.P. Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.

[21] H. Seidl. Finite tree automata with cost functions. *Theoret. Comput. Sci.*, 126:113–142, 1994.

[22] J.W. Thatcher and J.B. Wright. Generalized finite automata theory with application to a decision problem of second-order logic. *Math. Systems Theory*, 2(1):57–81, 1968.