Myhill-Nerode Theorem for Recognizable Tree Series Revisited

Andreas Maletti

Institute for Theoretical Computer Science
Technische Universität Dresden
01062 Dresden, Germany

Abstract. In this contribution the MYHILL-NERODE congruence relation on tree series is reviewed and a more detailed analysis of its properties is presented. It is shown that, if a tree series is deterministically recognizable over a zero-divisor free and commutative semiring, then the MYHILL-NERODE congruence relation has finite index. By [Borchardt: *Myhill-Nerode Theorem for Recognizable Tree Series.* LNCS 2710. Springer 2003] the converse holds for commutative semifields, but not in general. In the second part, a slightly adapted version of the MYHILL-NERODE congruence relation is defined and a characterization is obtained for all-accepting weighted tree automata over multiplicatively cancellative and commutative semirings.

1 Introduction

By the MYHILL-NERODE theorem, we know that for every regular string language L, there exists a unique (up to isomorphism) minimal deterministic finite string automaton that recognizes L. This result was extended to several devices including finite tree automata (see the discussion in [1]), to weighted string automata [2] over (multiplicatively) cancellative semirings, and to weighted tree automata [3] over semifields (see [4, 5] for an introduction to semirings). For the weighted devices, the minimal deterministic automaton is no longer unique up to isomorphism. The structure of it is still unique but the distribution of the weights on the transitions may vary. In [2] this is called *unique up to pushing.* Weighted tree automata and transducers recently found promising applications (see [6] for a survey) in natural language processing, where the size of the automata is crucial and thus minimization essential.

Let us recall the MYHILL-NERODE congruence of [7]. Two trees t and u are equal in the MYHILL-NERODE congruence $\equiv_\psi$ for a given tree series ψ over the semifield $(A, +, \cdot, 0, 1)$, if there exist nonzero coefficients $a, b \in A$ such that for all contexts C we observe the equality $a^{-1} \cdot (\psi, C[t]) = b^{-1} \cdot (\psi, C[u])$. In this expression, the coefficients a and b can be understood as the weights of t and u, respectively. Both sides of the previous equation can be understood as futures; the futures ψ_t and ψ_u are given for every context C by $(\psi_t, C) = a^{-1} \cdot (\psi, C[t])$ and $(\psi_u, C) = b^{-1} \cdot (\psi, C[u])$. Roughly speaking, in ψ_t a context is assigned the weight of $C[t]$ in ψ with the weight of t cancelled out. In other words, trees t and u are equal if and only if their futures ψ_t and ψ_u coincide.

The MYHILL-NERODE congruence $\equiv_\psi$ has two major applications: (i) it exactly characterizes whether ψ is deterministically recognizable; i.e., ψ is deterministically recognizable if and only if $\equiv_\psi$ has finite index; and (ii) it presents a minimal deterministic wta that recognizes the tree series ψ. In this contribution, we consider the MYHILL-NERODE relation for semirings which are not necessarily semifields. We will show that, for all commutative and zero-divisor free (i.e., $a \cdot b = 0$ implies that $0 \in \{a, b\}$) semirings, a deterministically recognizable tree series ψ yields a MYHILL-NERODE congruence $\equiv_\psi$ with finite index. Thus whenever $\equiv_\psi$ has infinite index, then ψ is not deterministically recognizable. This extends a result of [7] from commutative semifields to commutative and zero-divisor free semirings. Secondly, we also consider the opposite direction with a particular focus on the minimal deterministic wta. We show how *all-accepting* [8] wta over cancellative semirings are related to unweighted tree automata. This connection can be used to minimize deterministic all-accepting wta over commutative and cancellative semirings. A Myhill-Nerode theorem for all-accepting wta over semifields is already presented in [8]. We contribute an explicit minimization and an extension of the result to cancellative semirings.

Finally, we also investigate the construction of a minimal wta in the general case (again over a cancellative semiring). To this end, we present a slightly adapted Myhill-Nerode relation. However, one main point remains open: In cancellative semirings (as opposed to semifields) the Myhill-Nerode congruence relation is not always implementable. It remains an open problem to define suitable properties on ψ and the underlying semiring $\mathcal{A}$ such that the refined Myhill-Nerode congruence is implementable. We demonstrate the applicability of the general approach by deriving such properties and thus a Myhill-Nerode theorem for deterministic all-accepting weighted tree automata.

2 Preliminaries

We use $\mathbb{N}$ to represent the nonnegative integers. Further we denote $\{n \in \mathbb{N} \mid 1 \leqslant n \leqslant k\}$ by $[1, k]$. A set Σ that is nonempty and finite is also called an *alphabet*. A *ranked alphabet* is an alphabet Σ with a mapping $\mathrm{rk}_\Sigma\colon \Sigma \to \mathbb{N}$. We write Σ_k for $\{\sigma \in \Sigma \mid \mathrm{rk}_\Sigma(\sigma) = k\}$. Given a ranked alphabet Σ, the set of *Σ-trees*, denoted by T_Σ, is inductively defined to be the smallest set T such that for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \dots, t_k \in T$ also $\sigma(t_1, \dots, t_k) \in T$. We generally write α instead of $\alpha()$ whenever $\alpha \in \Sigma_0$. Let $\Box$ be a distinguished nullary symbol. A *context* C is a tree from $T_{\Sigma \cup \{\Box\}}$ such that the nullary symbol $\Box$ occurs exactly once in C. The set of all contexts over Σ is denoted by C_Σ. Finally, we write $C[t]$ for the tree of T_Σ that is obtained by replacing in the context $C \in C_\Sigma$ the unique occurrence of $\Box$ with the tree $t \in T_\Sigma$.

Let $\equiv$ and $\cong$ be equivalence relations on a set S. We write $[s]_\equiv$ for the equivalence class of $s \in S$ and $(S/{\equiv}) = \{[s]_\equiv \mid s \in S\}$ for the set of equivalence classes. We drop the subscript from $[s]_\equiv$ whenever it is clear from the context. We say that $\equiv$ is *coarser* than $\cong$ if $\cong\, \subseteq\, \equiv$. Now suppose that $S = T_\Sigma$. We say that $\cong$ is a *congruence* (on the term algebra T_Σ) if for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $s_1, \dots, s_k, t_1, \dots, t_k \in T_\Sigma$ with $s_i \cong t_i$ also $\sigma(s_1, \dots, s_k) \cong \sigma(t_1, \dots, t_k)$.

A *(commutative) semiring* is an algebraic structure $(A, +, \cdot, 0, 1)$ 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 $a_i \neq 0$. The semiring $(A, +, \cdot, 0, 1)$ is called *zero-sum free* if for every $a, b \in A$ the condition $a + b = 0$ implies that $a = 0 = b$. We call a semiring $(A, +, \cdot, 0, 1)$ *zero-divisor free* if $a \cdot b = 0$ implies that $a = 0$ or $b = 0$. Moreover, $\mathcal{A}$ is called *cancellative* if $a \cdot b = a \cdot c$ implies $b = c$ for every $a, b, c \in A$ with $a \neq 0$. Generally, we write $a|b$ whenever there exists an element $c \in A$ such that $a \cdot c = b$. Note that in a cancellative semiring such an element c, if any exists, is uniquely determined unless $a = 0 = b$. In cancellative semirings, we thus write b/a for that uniquely determined element c provided that (i) $a|b$ and (ii) $a \neq 0$ or $b \neq 0$. Finally, a *semifield* $\mathcal{A} = (A, +, \cdot, 0, 1)$ is a semiring such that for every $a \in A \setminus \{0\}$ there exists an element $a^{-1} \in A$ such that $a \cdot a^{-1} = 1$.

Let S be a set and $(A, +, \cdot, 0, 1)$ be a semiring. A *(formal) power series* ψ is a mapping $\psi\colon S \to A$; the set of all such mappings is denoted by $\mathcal{A}\langle\!\langle S \rangle\!\rangle$. 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 $\mathrm{supp}(\psi) = \{s \in S \mid (\psi, s) \neq 0\}$. The series with empty support is denoted by $\widetilde{0}$. Power series $\psi, \psi' \in \mathcal{A}\langle\!\langle S \rangle\!\rangle$ are added componentwise and multiplied componentwise with a semiring element; i.e., $(\psi + \psi', s) = (\psi, s) + (\psi', s)$ and $(a \cdot \psi, s) = a \cdot (\psi, s)$ for every $s \in S$ and $a \in A$. 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*.

There exists an abundance of (conceptionally) equivalent definitions of weighted tree automata [9–11] for various restricted semirings. Here we will only consider the general notion of [11, 7]. A *weighted tree automaton* [7] (for short: wta) is a tuple $(Q, \Sigma, \mathcal{A}, F, \mu)$ where Q is a nonempty, finite set of *states*; Σ is a ranked alphabet of *input symbols*; $\mathcal{A} = (A, +, \cdot, 0, 1)$ is a semiring; $F\colon Q \to A$ is a *final weight assignment*; and $\mu = (\mu_k)_{k\in\mathbb{N}}$ with $\mu_k\colon \Sigma_k \to A^{Q^k \times Q}$ is a *tree representation*. The wta M is called *(bottom-up) deterministic* (respectively, *(bottom-up) complete*), if for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $q_1, \ldots, q_k \in Q$ there exists at most (respectively, at least) one $q \in Q$ such that $\mu_k(\sigma)_{q_1\cdots q_k,q} \neq 0$. The wta induces a mapping $h_\mu\colon T_\Sigma \to A^Q$ that is defined for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, $q \in Q$, and $t_1, \ldots, t_k \in T_\Sigma$ by

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

The wta M recognizes the tree series $S(M) \in \mathcal{A}\langle\langle T_\Sigma\rangle\rangle$ given by

$$(S(M), t) = \sum_{q\in Q} F(q) \cdot h_\mu(t)_q$$

for every tree $t \in T_\Sigma$. A tree series $\psi \in \mathcal{A}\langle\langle T_\Sigma\rangle\rangle$ is called *recognizable* (respectively, *deterministically recognizable*), if there exists a wta M (respectively, deterministic wta M) such that $S(M) = \psi$. The sets of all recognizable and deterministically recognizable tree series are denoted by $\mathcal{A}^{\mathrm{rec}}\langle\langle T_\Sigma\rangle\rangle$ and $\mathcal{A}^{\mathrm{rec}}_{\mathrm{det}}\langle\langle T_\Sigma\rangle\rangle$, respectively.

3 Recognizable yields finite index

In this section, we show that the Myhill-Nerode congruence given by [3, Section 5] is necessarily of finite index for every deterministically recognizable series over a zero-divisor free semiring. Thus we derive a necessary criterion for a series $\psi \in \mathcal{A}\langle\langle T_\Sigma\rangle\rangle$ to be recognizable by some deterministic wta. Moreover, we also obtain a lower bound on the number of states of any deterministic wta that recognizes ψ. The development in this section closely follows [7, Chapter 7] where the same statements are proved for semifields.

Let us start with the definition of the Myhill-Nerode relation for a tree series $\psi \in \mathcal{A}\langle\langle T_\Sigma\rangle\rangle$. Intuitively, two trees $t, u \in T_\Sigma$ are related if they behave equal in all contexts $C \in C_\Sigma$ (up to fixed factors). The factors can be imagined to be the weights of the trees t and u.

Definition 1 (see [7, Chapter 7]). *Let $\psi \in \mathcal{A}\langle\langle T_\Sigma\rangle\rangle$. The relation $\equiv_\psi \subseteq T_\Sigma \times T_\Sigma$ is defined for every $t, u \in T_\Sigma$ by $t \equiv_\psi u$ if and only if there exist $a, b \in A \setminus \{0\}$ such that for every $C \in C_\Sigma$ we observe*

$$a \cdot (\psi, C[t]) = b \cdot (\psi, C[u]) \enspace .$$

The relation $\equiv_\psi$ is equivalent to the Myhill-Nerode relation presented in [3, Section 5] provided that the semiring is a semifield. Our first lemma states that $\equiv_\psi$ is indeed an equivalence relation, and moreover, a congruence whenever the underlying semiring is zero-divisor free.

Lemma 2 (cf. [7, Lemma 7.1.2(ii)]). *Let $\mathcal{A}$ be a zero-divisor free semiring and $\psi \in \mathcal{A}\langle\langle T_\Sigma\rangle\rangle$. Then $\equiv_\psi$ is a congruence.*

Proof. The proof follows the proof of [7, Lemma 7.1.2(ii)], where the statement is proved for semifields. □

We presented a congruence which is uniquely determined by ψ. First we show that every deterministic and complete wta $M = (Q, \Sigma, \mathcal{A}, F, \mu)$ also induces a congruence relation. For the development of this we need some additional notions. The *fta underlying M* (see [12, 13] for a detailed introduction to *finite tree automata*; for short: fta) is defined as $\mathcal{B}(M) = (Q, \Sigma, \delta, F')$ where

– $q \in \delta_\sigma(q_1, \ldots, q_k)$ iff $\mu_k(\sigma)_{q_1 \cdots q_k, q} \neq 0$ for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $q, q_1, \ldots, q_k \in Q$; and
– $q \in F'$ iff $F_q \neq 0$ for every $q \in Q$.

We note that the fta underlying a deterministic and complete wta is deterministic and complete. Now let M be a deterministic and complete wta and $\mathcal{B}(M) = (Q, \Sigma, \delta, F')$ be the fta underlying M. We define the mapping $R_M : T_\Sigma \to Q$ for every $t \in T_\Sigma$ by $R_M(t) = q$ where $q \in Q$ is the unique state such that $q \in \delta(t)$. Existence and uniqueness are guaranteed by completeness and determinism of $\mathcal{B}(M)$, respectively. We denote $\ker(R_M)$ by $\equiv_M$. The following lemma follows the traditional unweighted approach.

Lemma 3. *Let M be a deterministic and complete wta over Σ. Then $\equiv_M$ is a congruence with finite index.*

Having two congruences, namely $\equiv_M$ and $\equiv_{S(M)}$, let us try to relate them. In fact, it turns out that $\equiv_{S(M)}$ is coarser than $\equiv_M$ for every deterministic and complete wta M over a zero-divisor free semiring. This shows that we need at least as many states as there are equivalence classes in $\equiv_\psi$ to recognize ψ with some deterministic and complete wta.

Theorem 4. *Let $\mathcal{A}$ be a zero-divisor free semiring, and let M be deterministic and complete wta over $\mathcal{A}$. Then $\equiv_{S(M)}$ is coarser than $\equiv_M$.*

Proof. Let $M = (Q, \Sigma, \mathcal{A}, F, \mu)$, and let $t, u \in T_\Sigma$ be such that $t \equiv_M u$; that is $R_M(t) = R_M(u)$. Let $p = R_M(t)$. Thus, also $R_M(C[t]) = R_M(C[u])$. Let $a = h_\mu(u)_{R_M(u)}$ and $b = h_\mu(t)_{R_M(t)}$. We claim that for every context $C \in C_\Sigma$

$$a \cdot (S(M), C[t]) = b \cdot (S(M), C[u]) \enspace .$$

Let us distinguish two cases for $q = R_M(C[t])$. First, let us suppose that $F_q = 0$. Then the displayed equation holds because $(S(M), C[t]) = 0 = (S(M), C[u])$. In the remainder suppose that $F_q \neq 0$. Clearly, since $t \equiv_M u$ also $C[t] \equiv_M C[u]$ because $\equiv_M$ is a congruence by Lemma 3. Thus

$$\begin{aligned} a \cdot (S(M), C[t]) &= h_\mu(u)_p \cdot F_q \cdot h_\mu(C[t])_q = h_\mu(u)_p \cdot F_q \cdot h_\mu(C)^p_q \cdot h_\mu(t)_p \\ &= h_\mu(t)_p \cdot F_q \cdot h_\mu(C[u])_q = b \cdot (S(M), C[u]) \end{aligned}$$

where $h_\mu(C[t])_q = h_\mu(C)^p_q \cdot h_\mu(t)_p$ can be proved in a straightforward manner. Consequently, $\equiv_{S(M)}$ is coarser than $\equiv_M$. ☐

As already argued this theorem admits an important corollary, which shows a lower bound on the number of states of any deterministic and complete wta that recognizes a certain series.

Corollary 5. *Let $\mathcal{A}$ be a zero-divisor free semiring and $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$. Every deterministic and complete wta M over $\mathcal{A}$ with $S(M) = \psi$ has at least* $\mathrm{index}(\equiv_\psi)$ *states.*

Proof. Let $M = (Q, \Sigma, \mathcal{A}, F, \mu)$. By Theorem 4, we have that $\equiv_\psi$ is coarser than $\equiv_M$. Thus $\mathrm{index}(\equiv_M) \geq \mathrm{index}(\equiv_\psi)$. Finally, we observe that $\mathrm{index}(\equiv_M) = \mathrm{card}(Q)$. ☐

Let us show that the statement does not hold, if we consider arbitrary semirings. Essentially, if the semiring admits zero-divisors, then it can store information in the weight.

Example 6. Let $\mathbb{Z}_4 = (\{0, 1, 2, 3\}, +, \cdot, 0, 1)$ where $+$ and $\cdot$ are the usual addition and multiplication, respectively, modulo 4. Clearly, $2 \cdot 2 = 0$ and thus $\mathbb{Z}_4$ is not zero-divisor free. Let $\Sigma = \{\sigma^{(2)}, \alpha^{(0)}, \beta^{(0)}\}$. We consider the series $\psi \in \mathbb{Z}_4\langle\langle T_\Sigma \rangle\rangle$ which is defined for every $t \in T_\Sigma$ by

$$(\psi, t) = \begin{cases} 1 & \text{if } |t|_\alpha = 0 \\ 2 & \text{if } |t|_\alpha = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let $M = (\{\star\}, \Sigma, \mathbb{Z}_4, F, \mu)$ with $F_\star = 1$ and

$$\mu_0(\alpha)_\star = 2 \qquad \mu_0(\beta)_\star = 1 \qquad \mu_2(\sigma)_{\star\star,\star} = 1 \ .$$

It can easily be checked that $S(M) = \psi$. Let us suppose that still $\equiv_\psi$ is coarser than $\equiv_M$. Since $\mathrm{card}(Q) = 1$ this means that $t \equiv_\psi u$ for all $t, u \in T_\Sigma$. We consider the trees $\sigma(\alpha, \alpha)$ and β. By definition, we should have that there exist $a, b \in [1, 3]$ such that for every $C \in C_\Sigma$

$$a \cdot (\psi, C[\sigma(\alpha, \alpha)]) = b \cdot (\psi, C[\beta]) \ .$$

Now consider the context $C = \Box$. Thus $a \cdot 0 = b \cdot 1$ and thus $b = 0$. However, $b \in [1, 3]$ which is the desired contradiction. Thus $\sigma(\alpha, \alpha) \not\equiv_\psi \beta$ and $\equiv_\psi$ is not coarser than $\equiv_M$.

Finally, let us conclude this section with an application of Corollary 5. We can envision at least two uses of Corollary 5. It can be used to show that some wta is minimal (or almost so), and it can be used to show that some tree series ψ is not recognizable. The standard examples for the latter use concerns the tree series size and height over the natural numbers and the arctic semiring $(\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$, respectively (see discussion at [7, Examples 7.3.2 and 8.1.8]). We demonstrate the application of Corollary 5 on another example.

Example 7. Let Σ be a ranked alphabet such that $\Sigma = \Sigma_2 \cup \Sigma_0$. We use the tropical semiring $\mathcal{T} = (\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$. It is easily checked that $\mathcal{T}$ is not cancellative but zero-divisor free. We define the mapping $\mathrm{zigzag}\colon T_\Sigma \to \mathbb{N}$ for every $\alpha \in \Sigma_0$ and $\sigma, \delta \in \Sigma_2$ and $t_1, t_2, t_3 \in T_\Sigma$ by

$$\begin{aligned} \mathrm{zigzag}(\alpha) &= 1 \\ \mathrm{zigzag}(\sigma(\alpha, t_1)) &= 2 \\ \mathrm{zigzag}(\sigma(\delta(t_1, t_2), t_3)) &= 2 + \mathrm{zigzag}(t_2) \ . \end{aligned}$$

It is straightforward to show that $\mathrm{zigzag} \in \mathcal{T}^{\mathrm{rec}}\langle\langle T_\Sigma\rangle\rangle$. In fact, zigzag can be recognized by a top-down deterministic wta [7, Section 4.2] with only 2 states. But can zigzag be recognized by a (bottom-up) deterministic wta over Σ and $\mathcal{T}$? We use Corollary 5 to show that no deterministic wta recognizes zigzag. Clearly, this is achieved by proving that $\equiv_{\mathrm{zigzag}}$ has infinite index. Let $t, u \in T_\Sigma$. Suppose that $t \equiv_{\mathrm{zigzag}} u$. Then there exist $a, b \in \mathbb{N}$ such that for every context $C \in C_\Sigma$

$$a + (\mathrm{zigzag}, C[t]) = b + (\mathrm{zigzag}, C[u]) \ .$$

Now consider the contexts $C_1 = \Box$ and $C_2 = \sigma(\alpha, \Box)$. We obtain the equations

$$\begin{aligned} a + \mathrm{zigzag}(t) &= b + \mathrm{zigzag}(u) \\ a + 2 &= b + 2 \end{aligned}$$

From the second equality we can conclude that $a = b$ and $\mathrm{zigzag}(t) = \mathrm{zigzag}(u)$. Hence $\ker(\mathrm{zigzag})$ is coarser than $\equiv_{\mathrm{zigzag}}$. However, $\ker(\mathrm{zigzag})$ has infinite index, which shows that also $\equiv_{\mathrm{zigzag}}$ has infinite index, and thus, by Corollary 5, no deterministic wta can recognize zigzag.

4 Finite index yields recognizable

In this section we investigate whether the lower bound established in the previous section can be achieved. Certainly, [7, Theorem 7.4.1] shows that for every $\psi \in \mathcal{A}^{\mathrm{rec}}_{\mathrm{det}}\langle\langle T_\Sigma\rangle\rangle$ (with $\mathcal{A}$ a semifield) there exists a deterministic and complete wta over $\mathcal{A}$ with exactly $\mathrm{index}(\equiv_\psi)$ states so that $S(M) = \psi$. In this section we investigate this issue for deterministic all-accepting wta over cancellative semirings. The principal approach can also be extended to deterministic wta over certain cancellative semirings. Let us illustrate the problem in the semiring $(\mathbb{N}, +, \cdot, 0, 1)$ that is not a semifield but cancellative.

Consider the series $\psi\colon T_\Sigma \to \mathbb{N}$ with $\Sigma = \{\gamma^{(1)}, \alpha^{(0)}\}$ and

$$(\psi, \gamma^n(\alpha)) = \begin{cases} 2 & \text{if } n = 0 \\ 3 & \text{if } n = 1 \\ 4 & \text{otherwise.} \end{cases}$$

It is easily checked that $\alpha \not\equiv_\psi \gamma(\alpha) \not\equiv_\psi \gamma^n(\alpha) \not\equiv_\psi \alpha$ for every $n > 1$ as well as $\gamma^m(\alpha) \equiv_\psi \gamma^n(\alpha)$ for every $m > 1$ and $n > 1$. Thus, $\mathrm{index}(\equiv_\psi) = 3$. However, it can be shown that there exists no deterministic all-accepting [8] wta M such that $S(M) = \psi$. On the other hand, it is surprisingly easy to construct a deterministic wta M such that $S(M) = \psi$. In fact, M can be constructed such that it has 3 states.

4.1 Minimization of deterministic all-accepting wta

Let us discuss the problem for deterministic *all-accepting* wta [8, Section 3.2]. It is known that for every deterministic all-accepting wta M over a semifield there exists a unique (up to isomorphism) minimal deterministic and complete all-accepting wta that recognises $S(M)$ [8, Lemma 3.8]. We plan to extend this result to cancellative semirings. Now let us formally introduce the all-accepting property. We say that the wta $M = (Q, \Sigma, \mathcal{A}, F, \mu)$ is *all-accepting* [8] if $F(q) = 1$ for all $q \in Q$. We abbreviate *all-accepting wta* simply to *aa-wta*.

Let M be a deterministic aa-wta M. The tree series $S(M)$ recognised by M is *subtree-closed* [8, Section 3.1]; that is, for every tree t with $(S(M), t) \neq 0$ also $(S(M), u) \neq 0$ for every subtree u of t. We repeat [8, Observation 3.1] for ease of reference.

Proposition 8 (see [8, Observation 3.1]). *Let M be a deterministic aa-wta. Then $S(M)$ is subtree-closed.*

In order to avoid several cases, we assume that $0/0 = 0$ (i.e., we allow to cancel 0 from 0) for the rest of the paper. First we begin with two conditions which are necessary for a series $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$ to be recognizable by a deterministic aa-wta. Namely, we say that ψ is *implementable* if

- $((\psi, t_1) \cdot \ldots \cdot (\psi, t_k)) | (\psi, \sigma(t_1, \ldots, t_k))$ for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T_\Sigma$; and
- for every $k \in \mathbb{N}$ the set

$$C_k(\psi) = \left\{ \frac{(\psi, \sigma(t_1, \ldots, t_k))}{(\psi, t_1) \cdot \ldots \cdot (\psi, t_k)} \mid \sigma \in \Sigma_k, t_1, \ldots, t_k \in T_\Sigma \right\}$$

  is finite.

The following proposition is a straightforward observation.

Proposition 9. *Let M be a deterministic aa-wta over a cancellative semiring. Then $S(M)$ is implementable.*

Proof. The proof is standard and hence omitted. □

This shows that a series that is not implementable cannot be recognized by any deterministic aa-wta. In fact, this is the reason why the series ψ given at the beginning of Section 4 cannot be recognized by any deterministic aa-wta. Now we will show that the notion of recognizability by deterministic aa-wta over cancellative semirings is closely related to classical unweighted recognizability (as induced by fta). In fact, the weights of the deterministic aa-wta are uniquely determined so that they can also be included in the input ranked alphabet.

Definition 10. *Let $\psi \in \mathcal{A}\langle\!\langle T_\Sigma \rangle\!\rangle$ be implementable over the cancellative semiring $\mathcal{A}$. We define the ranked alphabet Δ by $\Delta_k = \Sigma_k \times C_k(\psi)$ for every $k \in \mathbb{N}$. Moreover, let $\cdot|_1 \colon T_\Delta \to T_\Sigma$ be the mapping that replaces every node label of the form $\langle \sigma, c \rangle$ in the input tree simply by a node with label σ. Finally, we define the tree language $L(\psi) \subseteq T_\Delta$ as the smallest language L such that for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $u_1, \ldots, u_k \in L$ with $t_i = u_i|_1$ for every $i \in [1, k]$*

$$\left\langle \sigma, \frac{(\psi, \sigma(t_1, \ldots, t_k))}{(\psi, t_1) \cdot \ldots \cdot (\psi, t_k)} \right\rangle (u_1, \ldots, u_k) \in L \qquad \Longleftrightarrow \qquad \sigma(t_1, \ldots, t_k) \in \operatorname{supp}(\psi) \ .$$

Theorem 11. *Let $\psi \in \mathcal{A}\langle\!\langle T_\Sigma \rangle\!\rangle$ with $\mathcal{A}$ a cancellative semiring. Then ψ is recognizable by some deterministic aa-wta if and only if ψ is implementable and $L(\psi)$ is recognizable.*

Proof. Let $M = (Q, \Sigma, \mathcal{A}, F, \mu)$ be a deterministic aa-wta and $M' = (Q, \Delta, \delta, Q)$ be a deterministic fta. We call M and M' *related* if

$$\mu_k(\sigma)_{q_1 \cdots q_k, q} = c \qquad \Longleftrightarrow \qquad \Big(\langle \sigma, c \rangle (q_1, \ldots, q_k) \to q\Big) \in \delta$$
$$\mu_k(\sigma)_{q_1 \cdots q_k, q} = 0 \qquad \Longleftrightarrow \qquad \forall c \in A \colon \Big(\langle \sigma, c \rangle (q_1, \ldots, q_k) \to q\Big) \notin \delta$$

for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, $c \in C_k(\psi) \setminus \{0\}$, and $q, q_1, \ldots, q_k \in Q$.

Now suppose that M and M' are related. We claim that $L(M') = L(S(M))$; the proof of this statement is omitted. Finally, let us now turn to the main statement. First let us suppose that there exists a deterministic aa-wta $M = (Q, \Sigma, \mathcal{A}, F, \mu)$ such that $S(M) = \psi$. By Proposition 9 it follows that ψ is implementable. Clearly, we can construct a deterministic fta M' such that M and M' are related. By the claimed property, we then have $L(M') = L(S(M)) = L(\psi)$, which proves that $L(\psi)$ is recognizable.

For the remaining direction, let ψ be implementable and, without loss of generality, let $M' = (Q, \Delta, \delta, F')$ be a deterministic fta such that $L(M') = L(\psi)$ and every state is reachable and co-reachable. It follows from the implementability condition that ψ is subtree-closed. With this in mind, we necessarily have $F' = Q$ because any reachable state in $Q \setminus F'$ would not be co-reachable. Moreover, for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $q, q_1, \ldots, q_k \in Q$ there exists at most one $c \in C$ such that $\langle \sigma, c \rangle (q_1, \ldots, q_k) \to q \in \delta$ because every state is final and $L(M') = L(\psi)$. Finally, no tree in $L(\psi)$ can contain a node $\langle \sigma, 0 \rangle$ for some $\sigma \in \Sigma$. This is due to the fact that $0 = (\psi, \sigma(t_1, \ldots, t_k))/((\psi, t_1) \cdot \ldots \cdot (\psi, t_k))$ only if $(\psi, \sigma(t_1, \ldots, t_k)) = 0$ by zero-divisor freeness of $\mathcal{A}$ and subtree-closedness of ψ. Thus, any reachable state that can recognize a tree of which one node is $\langle \sigma, 0 \rangle$ is not co-reachable. Consequently, there exists no such state and hence no transition which processes $\langle \sigma, 0 \rangle$. For the given fta M we can easily construct a related deterministic aa-wta and the previously proved statements guarantee that $L(S(M)) = L(M') = L(\psi)$. One final observation yields that $L \colon \mathcal{A}\langle\!\langle T_\Sigma \rangle\!\rangle \to T_\Delta$ is injective. Thus $L(S(M)) = L(\psi)$ yields that $S(M) = \psi$. ◻

The theorem admits a very important corollary. Namely, it can be observed that every minimal deterministic aa-wta M recognizing a given tree series ψ yields a minimal deterministic fta recognizing $L(\psi)$. In the opposite direction, every minimal deterministic fta recognizing $L(\psi)$ where ψ is implementable, yields a minimal deterministic aa-wta recognizing ψ.

Corollary 12 (of Theorem 11). *Let $\mathcal{A}$ be a cancellative semiring. For every deterministic aa-wta M there exists a unique (up to isomorphism) minimal deterministic aa-wta recognizing $S(M)$.*

Let us shortly describe a minimization procedure. Let M be a deterministic aa-wta. Then $S(M)$ is implementable and by the proof of Theorem 11 we can obtain an deterministic fta N

recognizing $L(S(M))$. Then we minimize N to obtain the unique minimal deterministic fta N' recognizing $L(S(M))$. Finally, we can construct a deterministic aa-wta M' recognizing $S(M)$ again using the notion of relatedness from the proof of Theorem 11.

We can imagine that the established relation between aa-wta and fta can also be exploited in the learning task of [8]. There the underlying semiring is a semifield (and hence cancellative).

4.2 A Myhill-Nerode congruence for cancellative semirings

In this section we consider general deterministic wta over certain cancellative semirings. The main problem is the implementability condition; it is crucial to the condition given in the previous section that the series is subtree-closed. In the general setting, subtree-closedness cannot be assumed.

A more careful analysis shows that the implementation of $\equiv_\psi$ [3, 7] uses inverses in an essential manner. Here we present a more refined version of the Myhill-Nerode congruence. Let $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$. Let $\cong_\psi \subseteq T_\Sigma \times T_\Sigma$ be defined for every $t, u \in T_\Sigma$ by $t \cong_\psi u$ if and only if there exist $a, b \in A \setminus \{0\}$ such that for every $C \in C_\Sigma$ there exists a $d \in A$ with

$$(\psi, C[t]) = d \cdot a \quad \text{and} \quad (\psi, C[u]) = d \cdot b \ .$$

This relation has several drawbacks as we will see next (it is, in general, no equivalence relation), however, we can already see that $\equiv_\psi$ is coarser than $\cong_\psi$.

Lemma 13. *Let $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$. In general, $\equiv_\psi$ is coarser than $\cong_\psi$. Provided that $\mathcal{A}$ is a semifield, then $\equiv_\psi$ and $\cong_\psi$ coincide.*

Proof. Let $\mathcal{A} = (A, +, \cdot, 0, 1)$. Moreover, let $t, u \in T_\Sigma$ such that $t \cong_\psi u$. Thus there exist $a, b \in A \setminus \{0\}$ such that for every context $C \in C_\Sigma$ there exists a $d \in A$ with

$$(\psi, C[t]) = d \cdot a \quad \text{and} \quad (\psi, C[u]) = d \cdot b \ .$$

Thus we also have that $b \cdot (\psi, C[t]) = a \cdot (\psi, C[u])$, and consequently, $t \equiv_\psi u$. For the second statement, suppose that $\mathcal{A}$ is a semifield and $t \equiv_\psi u$. Thus there exist $a, b \in A \setminus \{0\}$ such that for every $C \in C_\Sigma$

$$a \cdot (\psi, C[t]) = b \cdot (\psi, C[u]) \ .$$

Hence $(\psi, C[t]) = (\psi, C[u]) \cdot (a^{-1} \cdot b)$ and $(\psi, C[u]) = (\psi, C[t]) \cdot (b^{-1} \cdot a)$. Clearly, $a^{-1} \cdot b$ and $b^{-1} \cdot a$ are both nonzero, and consequently, $t \cong_\psi u$. ☐

Now let us investigate when $\cong_\psi$ is actually a congruence. A similar analysis was already done in [2] for weighted automata over strings. However, we slightly adapted the notions of greedy factorization and minimal residue (cf. [2, Section 4]).

Lemma 14. *The relation $\cong_\psi$ is reflexive for every $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$.*

Proof. Let $t \in T_\Sigma$. We need to prove that there exists an $a \in A \setminus \{0\}$ such that for every $C \in C_\Sigma$ there exists $d \in A$ with $(\psi, C[t]) = d \cdot a$. To this end, we let $a = 1$ and $d = (\psi, C[t])$. ☐

Clearly, $\cong_\psi$ is symmetric, so it remains to investigate transitivity. For this, we need an additional property. The semiring $\mathcal{A}$ allows *greedy factorization* if for every $a, b \in A$ there exist $a', b' \in A$ such that for every $c, d \in A$ there exists an $e \in A$ such that $a \cdot c = b \cdot d \neq 0$ implies $c = a' \cdot e$ and $d = b' \cdot e$. A similar property was already defined in [2].

Intuitively, the property demands that when a and b are divisors of a common element h, then there should be elements a' and b', that depend only on a and b and not on h, such that when cancelling a and a' from h we obtain the same element as we would obtain by cancelling b and b'. In this sense it represents a confluency property. It does not matter whether we first cancel a or b; we can later find elements a' and b', which depend solely on the cancelled elements a and b, that we can cancel to obtain a common element.

In semifields the property is trivially fulfilled because if we set $a' = b$ and $b' = a$ and $e = c \cdot b^{-1}$ then $a \cdot c = b \cdot d \neq 0$ implies $c = b \cdot c \cdot b^{-1}$ and $d = a \cdot c \cdot b^{-1}$. The first part of the conclusion is trivial and the second part is given by the hypothesis.

Let us try to give another example in order to explain the property. Suppose that $\mathcal{A}$ is a cancellative semiring with the additional property that a *least common multiple* (lcm) is defined for every two elements (e.g., the semiring of natural numbers fulfils these restrictions). We can then set $a' = \operatorname{lcm}(a, b)/a$ and $b' = \operatorname{lcm}(a, b)/b$ and $e = (a \cdot c)/\operatorname{lcm}(a, b)$ provided that $a \cdot c = b \cdot d$ otherwise set $e = 1$. Since the semiring is cancellative and $a | \operatorname{lcm}(a, b)$ and $b | \operatorname{lcm}(a, b)$ and $\operatorname{lcm}(a, b) | a \cdot c$ (because $a | a \cdot c$ and $b | a \cdot c$), the elements a', b', and e are uniquely determined. We thus obtain that $a \cdot c = b \cdot d \neq 0$ implies that $c = (\operatorname{lcm}(a, b)/a) \cdot ((a \cdot c)/\operatorname{lcm}(a, b))$ and $d = (\operatorname{lcm}(a, b)/b) \cdot ((a \cdot c)/\operatorname{lcm}(a, b))$. The first part of the conclusion is again trivial and the second part yields $b \cdot d = a \cdot c$, which holds by the hypothesis.

Lemma 15. *Let $\mathcal{A}$ be a zero-divisor free semiring that allows greedy factorization. Then $\cong_\psi$ is transitive for every $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$.*

Thus we successfully showed that $\cong_\psi$ is an equivalence relation. The only remaining step is to show that $\cong_\psi$ is even a congruence. Fortunately, this is rather easy.

Lemma 16. *Let $\mathcal{A}$ be a zero-divisor free semiring that allows greedy factorization. Then $\cong_\psi$ is a congruence for every $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$.*

Now let us proceed with the implementation of the congruence by some deterministic and complete wta. We prepare this by presenting conditions that imply that we can successfully implement a congruence. We chose to rephrase Conditions (MN1) and (MN2) from [7] in order to improve readability. In essence, we can already see the automaton in that modified definition of Conditions (MN1) and (MN2).

Definition 17. *Let $\cong \subseteq T_\Sigma \times T_\Sigma$ be a congruence and $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$. We say that $\cong$* respects *ψ if there exists a mapping $F\colon [T_\Sigma] \to A$ and a mapping $c\colon T_\Sigma \to A \setminus \{0\}$ such that*

- *$(\psi, t) = F([t]) \cdot c(t)$ for every $t \in T_\Sigma$; and*
- *for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $T_1, \ldots, T_k \in [T_\Sigma]$ there exists an $a \in A$, denoted by $b_\sigma(T_1, \ldots, T_k)$, such that*

$$c(\sigma(t_1, \ldots, t_k)) = a \cdot c(t_1) \cdot \ldots \cdot c(t_k)$$

for every $t_i \in T_i$ with $i \in [1, k]$.

Next we state that every series $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$ that is respected by some congruence with finite index can be recognized by a deterministic wta. Thus, the previous definition establishes sufficient conditions so that the congruence is implementable.

Lemma 18. *Let $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$. Moreover, let $\cong$ be a congruence with finite index that respects ψ. Then $\psi \in \mathcal{A}^{\mathrm{rec}}_{\mathrm{det}}\langle\langle T_\Sigma \rangle\rangle$.*

Proof. Since $\cong$ respects ψ, there exist $F\colon [T_\Sigma] \to A$, $c\colon T_\Sigma \to A \setminus \{0\}$, and $b_\sigma\colon [T_\Sigma]^k \to A$ for every $k \in \mathbb{N}$ and $\sigma \in \Sigma_k$ such that the conditions of Definition 17 hold. We construct the wta $M_\cong = ([T_\Sigma], \Sigma, \mathcal{A}, F, \mu)$ where

$$\mu_k(\sigma)_{[t_1]\cdots[t_k],[\sigma(t_1,\ldots,t_k)]} = b_\sigma([t_1], \ldots, [t_k])$$

for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \dots, t_k \in T_\Sigma$ and the remaining entries in μ are 0. Clearly, $M_\cong$ is deterministic. The proof of $S(M) = \psi$ is straightforward. □

In fact, the "respects" property is necessary and sufficient, which can be seen in the next theorem.

Theorem 19. *Let $\mathcal{A}$ be a zero-divisor free semiring, and let $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$. The following are equivalent:*

1. *There exists a congruence relation with finite index that respects ψ.*
2. *ψ is deterministically recognizable.*

Proof. The implication 1 to 2 is proved in Lemma 18. It remains to show that 2 implies 1. Let $M = (Q, \Sigma, \mathcal{A}, F, \mu)$ be a deterministic and complete wta such that $S(M) = \psi$. Clearly, $\equiv_M$ is a congruence with finite index by Lemma 3.

Finally, it remains to show that $\equiv_M$ respects ψ. Let $G\colon [T_\Sigma] \to A$ and $c\colon T_\Sigma \to A \setminus \{0\}$ be defined by $G([t]) = F_{R_M(t)}$ and $c(t) = h_\mu(t)_{R_M(t)}$ for every $t \in T_\Sigma$. It is easily verified that both mappings are well-defined. First we need to prove that $(\psi, t) = G([t]) \cdot c(t)$ for every $t \in T_\Sigma$.

$$G([t]) \cdot c(t) = F_{R_M(t)} \cdot h_\mu(t)_{R_M(t)} = (S(M), t) = (\psi, t)$$

because M is deterministic. We observe that for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \dots, t_k \in T_\Sigma$

$$\begin{aligned}
& c(\sigma(t_1, \dots, t_k)) \\
&= h_\mu(\sigma(t_1, \dots, t_k))_{R_M(\sigma(t_1,\dots,t_k))} \\
&= \mu_k(\sigma)_{R_M(t_1)\cdots R_M(t_k), R_M(\sigma(t_1,\dots,t_k))} \cdot h_\mu(t_1)_{R_M(t_1)} \cdot \ldots \cdot h_\mu(t_k)_{R_M(t_k)} \\
&= \mu_k(\sigma)_{R_M(t_1)\cdots R_M(t_k), R_M(\sigma(t_1,\dots,t_k))} \cdot c(t_1) \cdot \ldots \cdot c(t_k)
\end{aligned}$$

which proves that $\equiv_M$ respects ψ. □

In analogy to Theorem 4 we can show that $\cong_{S(M)}$ is coarser than $\equiv_M$ for every deterministic and complete wta over a zero-divisor free semiring. Thus, the only remaining question is whether $\cong_\psi$ respects ψ. If this would be true and $\cong_\psi$ would have finite index, then $\cong_\psi$ would be implementable and thus a minimal deterministic and complete wta would be found.

Open problem: Find suitable conditions on ψ and $\mathcal{A}$ so that $\cong_\psi$ respects ψ!

4.3 A Myhill-Nerode theorem for all-accepting wta

In this section we show how we can use the approach of the previous section to derive a Myhill-Nerode theorem for deterministic aa-wta.

Let $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$ be a tree series over the cancellative semiring $\mathcal{A}$. We define $\simeq_\psi \subseteq T_\Sigma \times T_\Sigma$ by $t \simeq_\psi u$ if and only if there exist $a, b \in A \setminus \{0\}$ such that for every $C \in C_\Sigma$ there exists a $d \in A$ with

$$(\psi, C[t]) = d \cdot a \quad \text{and} \quad (\psi, C[u]) = d \cdot b \quad \text{and} \quad d \in \{0, 1\} \text{ if } C = \square \ .$$

Lemma 20. *If ψ is implementable and $\mathcal{A}$ allows greedy factorization, then $\simeq_\psi$ is a congruence.*

Proof. The proof can be obtained by reconsidering the proofs of Lemmata 14, 15, and 16. □

Let us consider the open problem for deterministic aa-wta over cancellative semirings.

Theorem 21. *Let $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$ be implementable with $\mathcal{A}$ a cancellative semiring that allows greedy factorization. Then $\simeq_\psi$ respects ψ.*

Proof. By Lemma 20, $\simeq_\psi$ is a congruence. Thus we need to show that there exist mappings $F\colon [T_\Sigma] \to A$ and $c\colon T_\Sigma \to A \setminus \{0\}$ such that the conditions of Definition 17 are met. For every $t \in T_\Sigma$ let

$$F([t]) = \begin{cases} 1 & \text{if } t \in \operatorname{supp}(\psi) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad c(t) = \begin{cases} (\psi, t) & \text{if } t \in \operatorname{supp}(\psi) \\ 1 & \text{otherwise.} \end{cases}$$

We first verify that F is well-defined. Let $t \simeq_\psi u$. We need to prove that $t \in \operatorname{supp}(\psi)$ if and only if $u \in \operatorname{supp}(\psi)$. Since $t \simeq_\psi u$ there exist $a, b \in A \setminus \{0\}$ such that for every context $C \in C_\Sigma$ there exists $d \in A$ with

$$(\psi, C[t]) = d \cdot a \quad \text{and} \quad (\psi, C[u]) = d \cdot b \quad \text{and} \quad d \in \{0, 1\} \text{ if } C = \square \ .$$

Now consider the context $C = \square$. Thus $(\psi, t) = d \cdot a$ and $(\psi, u) = d \cdot b$ with $d \in \{0, 1\}$. Depending on d either (i) $(\psi, t) = 0 = (\psi, u)$ or (ii) $t, u \in \operatorname{supp}(\psi)$, which proves that F is well-defined. It remains to verify the properties of Definition 17. First, for every $t \in T_\Sigma$

$$F([t]) \cdot c(t) = \begin{cases} 1 \cdot (\psi, t) & \text{if } t \in \operatorname{supp}(\psi) \\ 0 & \text{otherwise} \end{cases} = (\psi, t) \ .$$

Second, let $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T_\Sigma$. We need to show that there exists a $b_\sigma([t_1], \ldots, [t_k])$ such that $c(\sigma(t_1, \ldots, t_k)) = b_\sigma([t_1], \ldots, [t_k]) \cdot c(t_1) \cdot \ldots \cdot c(t_k)$. Since ψ is implementable, we can define $b_\sigma([t_1], \ldots, [t_k]) = (\psi, \sigma(t_1, \ldots, t_k))/((\psi, t_1) \cdot \ldots \cdot (\psi, t_k))$. We should first verify that this is independent of the representatives. Thus, let $u_1, \ldots, u_k \in T_\Sigma$ be such that $t_i \simeq_\psi u_i$ for every $i \in [1, k]$. Then there exist $a_i, b_i \in A \setminus \{0\}$ such that for every context $C \in C_\Sigma$ there exists $d_i \in A$ with

$$(\psi, C[t_i]) = d_i \cdot a_i \quad \text{and} \quad (\psi, C[u_i]) = d_i \cdot b_i \quad \text{and} \quad d_i \in \{0, 1\} \text{ if } C = \square$$

for every $i \in [1, k]$. Now if $(\psi, t_i) = 0$ then also $(\psi, C[t]) = 0$ because ψ is implementable. The same argument holds for u_i and $C[u_i]$. Suppose that there exist $i \in [1, k]$ such that $(\psi, t_i) = 0$. Then $(\psi, \sigma(t_1, \ldots, t_k))/((\psi, t_1) \cdot \ldots \cdot (\psi, t_k)) = 0$ and since $(\psi, u_i) = 0$ by $t_i \simeq_\psi u_i$ also $(\psi, \sigma(u_1, \ldots, u_k))/((\psi, u_1) \cdot \ldots \cdot (\psi, u_k)) = 0$. Now suppose that $(\psi, t_i) \neq 0$ for every $i \in [1, k]$. It is immediately clear that $a_i = (\psi, t_i)$ and $b_i = (\psi, u_i)$ by considering the context $\square$. Consequently,

$$\begin{aligned}
&(\psi, \sigma(t_1, \ldots, t_k))/\Big(\prod_{i=1}^{k} (\psi, t_i)\Big) \\
&= (\psi, \sigma(u_1, t_2, \ldots, t_k))/\Big((\psi, u_1) \cdot \prod_{i=2}^{k} (\psi, t_i)\Big) && \text{(via the context } \sigma(\square, t_2, \ldots, t_k)) \\
&= \ldots \\
&= (\psi, \sigma(u_1, \ldots, u_{k-1}, t_k))/\Big(\prod_{i=1}^{k-1} (\psi, u_i) \cdot (\psi, t_k)\Big) \\
&= (\psi, \sigma(u_1, \ldots, u_k))/\Big(\prod_{i=1}^{k} (\psi, u_i)\Big) && \text{(via the context } \sigma(u_1, \ldots, u_{k-1}, \square))
\end{aligned}$$

□

Let us now derive a MYHILL-NERODE theorem for deterministic aa-wta. In [8] such a theorem is shown for the case that the underlying semiring is a semifield. We extend this result to certain cancellative semirings.

Corollary 22. *Let $\psi \in \mathcal{A}\langle\langle T_\Sigma \rangle\rangle$ be implementable with $\mathcal{A}$ a cancellative semiring that allows greedy factorization. The following are equivalent:*

1. *$\simeq_\psi$ has finite index.*
2. *There exists a congruence with finite index that respects ψ.*
3. *ψ is deterministically recognizable.*
4. *ψ is recognized by some deterministic aa-wta M.*

Proof. $1 \to 2$ was shown in Theorem 21. The equivalence of 2 and 3 is due to Theorem 19. Moreover, we already remarked that $\simeq_\psi$ is coarser than $\equiv_M$, which shows $4 \to 1$. It remains to show $3 \to 4$. This can be shown by a straightforward construction that normalizes the final weights to 1. In general, this is only possible in a semifield, but due to the implementability of ψ, it can also be performed in the cancellative semiring $\mathcal{A}$. □

Clearly, the above corollary shows that the tree series that can be recognized by deterministic aa-wta are exactly the implementable tree series that can be recognized by deterministic wta. Moreover, it can be shown that the deterministic aa-wta that can be constructed from the deterministic wta using the final weight normalization mentioned in the proof of Corollary 22 is indeed the minimal deterministic aa-wta recognizing ψ.

References

1. Kozen, D.: On the Myhill-Nerode theorem for trees. Bulletin of the EATCS **47** (1992) 170–173
2. Eisner, J.: Simpler and more general minimization for weighted finite-state automata. In: HLT-NAACL. (2003)
3. Borchardt, B.: The Myhill-Nerode theorem for recognizable tree series. In: Proc. 7th Int. Conf. Develop. in Lang. Theory. Volume 2710 of LNCS., Springer (2003) 146–158
4. Hebisch, U., Weinert, H.J.: Semirings—Algebraic Theory and Applications in Computer Science. World Scientific (1998)
5. Golan, J.S.: Semirings and their Applications. Kluwer Academic, Dordrecht (1999)
6. Knight, K., Graehl, J.: An overview of probabilistic tree transducers for natural language processing. In: Proc. 6th Int. Conf. Comput. Ling. and Int. Text Proc. Volume 3406 of LNCS., Springer (2005)
7. Borchardt, B.: The Theory of Recognizable Tree Series. PhD thesis, Technische Universität Dresden (2005)
8. Drewes, F., Vogler, H.: Learning deterministically recognizable tree series. J. Autom. Lang. Combin. (2007) to appear.
9. Berstel, J., Reutenauer, C.: Rational Series and Their Languages. Volume 12 of EATCS Monographs on Theoret. Comput. Sci. Springer (1988)
10. Bozapalidis, S.: Equational elements in additive algebras. Theory Comput. Systems **32**(1) (1999) 1–33
11. Borchardt, B., Vogler, H.: Determinization of finite state weighted tree automata. J. Autom. Lang. Combin. **8**(3) (2003) 417–463
12. Gécseg, F., Steinby, M.: Tree Automata. Akadémiai Kiadó, Budapest (1984)
13. Gécseg, F., Steinby, M.: Tree languages. In Rozenberg, G., Salomaa, A., eds.: Handbook of Formal Languages. Volume 3. Springer (1997) 1–68