Sequentiality of Group-Weighted Tree Automata*

Frederic Dörband[1], Thomas Feller[2], and Kevin Stier[3]

[1] Technische Universität Dresden, Faculty of Computer Science, Germany
frederic.doerband@tu-dresden.de
[2] Technische Universität Dresden, Faculty of Computer Science, Germany
thomas.feller@tu-dresden.de
[3] Universität Leipzig, Institute of Computer Science, Germany
stier@informatik.uni-leipzig.de

Abstract. We introduce the notion of group-weighted tree automata over commutative groups and characterise sequentialisability of such automata. In particular, we introduce a fitting notion for tree distance and prove the equivalence between sequentialisability, the so-called Lipschitz property, and the so-called twinning property.

Keywords: Weighted Tree Automata · Degree of Sequentiality · Deterministic Weighted Automata · Twinning Property.

1 Introduction

In theoretical computer science, automata theory arose as a very potent field of research. Besides having manifold applications in areas like natural language processing, model checking, and computational biology, automata are studied in a vast number of syntactical variations. The most prominent case of finite string automata has been extended to handle more complex input structures like pictures, trees, and forests (cf. [17,18]). Another direction of generalisation is to allow quantitative calculations rather than simple binary acceptance. Well-studied examples of such automata are weighted string automata and weighted tree automata over some weight structure S (cf. [8] for exhaustive references). Prominent weight structures include commutative semirings [1] and strong bimonoids [9]. In the present paper we consider so-called group-weighted tree automata (short: group-WTA), which are particular semiring-weighted tree automata. We have adapted the notion of group-weighted tree automata from [7], where group-weighted string automata are studied.

One of the major research fields in automata theory is the determinisation of automata. While this problem has a well-known solution for unweighted string

* Research of the first and third author was supported by the DFG through the Research Training Group QuantLA (GRK 1763). The second author was supported by the European Research Council (ERC) through the ERC Consolidator Grant No. 771779 (DeciGUT).

automata, very little results are known in the weighted setting. In fact, not every weighted automaton can be determinised [3, Example 5.9] and hence, the problem has shifted towards the question of a *characterisation* of determinisable weighted automata. Two recent approaches to this question involve maximal factorisations [5] and automata with set semantics [2,7]. Note however, that [2,7] deal with sequentiality rather than determinism, which makes a subtle difference (Remark 10).

The main goal of the present paper is to characterise *sequential* weighted tree languages (i.e. weighted tree languages accepted by sequential group-WTA) by the so-called Lipschitz property and the so-called twinning property, namely

Theorem 1. For every group-WTA $\mathscr{A}$ it holds that

$$\begin{aligned} [\![\mathscr{A}]\!] \text{ is sequential } &\iff [\![\mathscr{A}]\!] \text{ satisfies the Lipschitz property} \\ &\iff \mathscr{A} \text{ satisfies the twinning property }. \end{aligned}$$

Hereby, our paper generalises [2,6] from the string case to the tree case. Note however, that [2,6] are proven for free monoids rather than (infinitary commutative) groups as in our case. The idea for the proof of Theorem 1 is based on [2] and our proof applies the terminology and proof techniques given in [7]. Note that [2] provides merely an implication of the form "twinning property $\Longrightarrow$ sequential", whereas [7] provides a full characterisation of sequentiality. In fact, [7] proves a more general theorem for unions of k sequential automata and the present paper only covers the case $k = 1$. Moreover, [7] is based on [6], which first introduced an equivalence similar to Theorem 1.

The present paper executes the proof of Theorem 1 in the following way. In Section 2, we introduce some elementary technical machinery and our automaton model.

In Section 3, we first introduce the *Lipschitz property* of weighted tree languages, which essentially says that close trees have close values in $\mathbb{G}$ (with respect to a metric on the set of trees and the Cayley distance on $\mathbb{G}$). Second, we introduce the *twinning property* of group-WTA, which states that if the automaton can loop[4] on a context tree in two different states, then the weights of these loops are equal. Next, we prove two implications of Theorem 1, namely "sequential $\Longrightarrow$ Lipschitz" and "Lipschitz $\Longrightarrow$ twinning".

In Section 4, we prove the implication "twinning $\Longrightarrow$ sequential" by applying a construction similar to the well-known weighted power set construction.

In Section 5, we give a brief presentation of our endeavours to lift the cases $k > 1$ from [7]. Most importantly we show where the approach from [7] fails in the tree case.

We conclude this introductory section by comparing our results to the existing sequentialisation/determinisation results from the literature. The major references for our proofs are [2,6,7]. As stated above, our results generalise [2,6] to the case of group-weighted tree automata. Furthermore, we lift [7, case $k = 1$]

[4] A *loop* is a run on a context tree such that the state at the context variable is the same as the state at the root of the context.

from the string case to the tree case. Another major result in the theory of determinisation is given in [5, Theorem 5.2], which subsumes the determinisation results from [3,4,13,14,15]. Besides the fact that our paper is concerned with sequentiality rather than determinism, our class of weight structures is not subsumed by [5]. In particular, [5] provides a determinisation result only if either $\mathscr{A}$ is nonrecursive, the semiring S is locally finite, or S is extremal, none of which apply to our semirings of the form $\mathcal{P}_{\text{fin}}(\mathbb{G})$. Similarly, the determinisation result given in [16, Section 6] deals only with locally finite strong bimonoids and hence again does not subsume our results.

2 Preliminaries

We denote the set of *nonnegative integers* by $\mathbb{N}$ and the set of *positive integers* by $\mathbb{N}_+$. For every $k \in \mathbb{N}$, we denote the set $\{i \in \mathbb{N} \mid 1 \leq i \leq k\}$ by $[k]$. Note that $[0] = \emptyset$. For a set A we denote the *size* of A by $\#A$ and for every $k \in \mathbb{N}_+$ we denote by A^k the k-fold cartesian power of A.

An *alphabet* is a finite and non-empty set A and $A^* = \bigcup_{k\in\mathbb{N}} A^k$ is the set of all (finite) *words* over A, where $A^0 = \{\varepsilon\}$ contains solely the *empty word* ε. We denote by $|w|$ the *length* of the word $w \in A^*$. Given words $v, w \in A^*$, their concatenation is written $v.w$ or simply vw. We write $v \preceq w$ provided that there exists $u \in A^*$ such that $vu = w$. The relation $\preceq$ is in fact a partial order, called the *prefix order*.

A *ranked alphabet* is a pair (Σ, rk) consisting of an alphabet Σ and a mapping $\mathrm{rk}\colon \Sigma \to \mathbb{N}$ that assigns a *rank* to each symbol of Σ. We refer to the ranked alphabet (Σ, rk) by the set Σ whenever the map rk is clear from the context. Furthermore, for every $k \in \mathbb{N}$, we let $\Sigma^{(k)} = \{\sigma \in \Sigma \mid \mathrm{rk}(\sigma) = k\}$ and we write $\sigma^{(k)}$ to indicate that $\mathrm{rk}(\sigma) = k$. Moreover we define $\mathrm{maxrk}(\Sigma) \coloneqq \max(\mathrm{rk}(\Sigma))$.

Throughout the rest of this paper, we assume Σ to be an arbitrary ranked alphabet.

Given a set Z, the set of Σ-*trees* indexed by Z, denoted by $\mathrm{T}_\Sigma(Z)$, is the smallest set T such that $Z \subseteq \mathrm{T}$ and $\sigma(\xi_1, \ldots, \xi_s) \in \mathrm{T}$ for every $s \in \mathbb{N}$, $\sigma \in \Sigma^{(s)}$, and $\xi_1, \ldots, \xi_s \in \mathrm{T}$. We abbreviate $\mathrm{T}_\Sigma = \mathrm{T}_\Sigma(\emptyset)$ and call every subset $L \subseteq \mathrm{T}_\Sigma$ a *tree language*.

Next, we recall some common notions and notations for trees. In the following, let $\xi \in \mathrm{T}_\Sigma(Z)$. The set $\mathrm{pos}(\xi)$ of *positions* of ξ is defined inductively by $\mathrm{pos}(z) = \{\varepsilon\}$ for all $z \in Z$, and $\mathrm{pos}(\sigma(\xi_1, \ldots, \xi_s)) = \{\varepsilon\} \cup \{i.w \mid i \in [s],\, w \in \mathrm{pos}(\xi_i)\}$ for every $s \in \mathbb{N}$, $\sigma \in \Sigma^{(s)}$, and $\xi_1, \ldots, \xi_s \in T_\Sigma(Z)$. The *height* of ξ is defined by $\mathrm{height}(\xi) = \max_{w\in\mathrm{pos}(\xi)} |w|$, and the *size* of ξ is defined by $\mathrm{size}(\xi) = \#\mathrm{pos}(\xi)$. A *leaf* is a position $w \in \mathrm{pos}(\xi)$ such that $w.1 \notin \mathrm{pos}(\xi)$. We denote the set of leaves of ξ by $\mathrm{leaf}(\xi)$. Given a position $w \in \mathrm{pos}(\xi)$, the *label* of ξ at w is denoted by $\xi(w)$. The *subtree* of ξ at w, denoted $\xi|_w$, is defined for every $z \in Z$ by $z|_\varepsilon = z$ and for every $s \in \mathbb{N}$, $\sigma \in \Sigma^{(s)}$, and $\xi_1, \ldots, \xi_s \in \mathrm{T}_\Sigma(Z)$

by

$$\sigma(\xi_1,\dots,\xi_s)|_w = \begin{cases} \sigma(\xi_1,\dots,\xi_s) & \text{if } w = \varepsilon \\ \xi_i|_{w'} & \text{if } w = i.w' \text{ with } i \in \mathbb{N} \text{ and } w' \in \mathrm{pos}(\xi_i). \end{cases}$$

Let Y be a set. The set of *positions of ξ labeled by elements in Y*, denoted by $\mathrm{pos}_Y(\xi)$, is the set $\{w \in \mathrm{pos}(\xi) \mid \xi(w) \in Y\}$. Moreover, the *replacement* of the leaf $w \in \mathrm{leaf}(\xi)$ by the tree $\eta \in T_\Sigma(Z)$, denoted $\xi[\eta]_w$, is given for every $z \in Z$ by $z[\eta]_\varepsilon = \eta$ and for every $s \in \mathbb{N}$, $i \in [s]$, $\sigma \in \Sigma^{(s)}$, $\xi_1,\dots,\xi_s \in \mathrm{T}_\Sigma(Z)$, and $w' \in \mathrm{pos}(\xi_i)$ by $\sigma(\xi_1,\dots,\xi_s)[\eta]_{i.w'} = \sigma(\xi_1,\dots,\xi_{i-1},\xi_i[\eta]_{w'},\xi_{i+1},\dots,\xi_s)$.

We fix the set $X = \{x_1, x_2, \dots\}$ of *variables* (which we impose to be disjoint from any other set we consider), and $X_n = \{x_1,\dots,x_n\}$ for every $n \in \mathbb{N}_+$. A tree $\xi \in \mathrm{T}_\Sigma(X_1)$ is a *context*, if $\#\mathrm{pos}_{x_1}(\xi) = 1$. The set of all contexts is denoted by C_Σ.

Given a context $\zeta \in \mathrm{C}_\Sigma$ and a tree $\xi \in \mathrm{T}_\Sigma(Z)$, the *substitution* of ξ into ζ, denoted by $\zeta[\xi]$, is the tree $\zeta[\xi]_w$, where w is the unique position in $\mathrm{pos}_X(\zeta)$. Note that, given $\zeta, \zeta' \in \mathrm{C}_\Sigma$, also $\zeta[\zeta'] \in \mathrm{C}_\Sigma$. We write ζ^k for $\zeta[\zeta[\cdots\zeta[\zeta]\cdots]]$ containing the context ζ a total of k times.

Let $\xi_1, \xi_2 \in \mathrm{T}_\Sigma$ be two trees. A *pair-cut* between ξ_1 and ξ_2 is a triple $(\zeta_1, \zeta_2, \eta) \in \mathrm{C}_\Sigma \times \mathrm{C}_\Sigma \times \mathrm{T}_\Sigma$ such that $\zeta_i[\eta] = \xi_i$ for $i \in [2]$. In this case, we call η an *overlap* of ξ_1 and ξ_2. The set of pair-cuts between ξ_1 and ξ_2 is denoted $\mathrm{PairCut}(\xi_1, \xi_2)$. We moreover define the *distance* between ξ_1 and ξ_2 as

$$\mathrm{dist}(\xi_1,\xi_2) \coloneqq \mathrm{size}(\xi_1) + \mathrm{size}(\xi_2) - 2 \cdot \mathrm{maxoverlap}(\xi_1,\xi_2),$$

where $\mathrm{maxoverlap}(\xi_1,\xi_2)$ is the maximal size of an overlap of ξ_1 and ξ_2.

A *group* $(\mathbb{G}, \otimes, 1)$ is a set $\mathbb{G}$ with an associative operation $\otimes : \mathbb{G}^2 \to \mathbb{G}$, a neutral element $1 \in \mathbb{G}$ such that for all $\alpha \in \mathbb{G}$ there exists $\beta \in \mathbb{G}$ satisfying $\alpha \otimes \beta = \beta \otimes \alpha = 1$. We refer to β as the *inverse element* of α and denote it by α^{-1}. We simply write $\alpha\beta$ for $\alpha \otimes \beta$. A group is *commutative* if $\otimes$ is commutative. We call a group *infinitary* if for every $\alpha, \beta, \gamma \in \mathbb{G}$ with $\alpha\beta\gamma \neq \beta$, the set $\{\alpha^n \beta \gamma^n \mid n \in \mathbb{N}\}$ is infinite (cf. [7,10]). We define the *delay* of $\alpha \in \mathbb{G}$ and $\beta \in \mathbb{G}$, denoted $\mathrm{delay}(\alpha, \beta)$, by $\alpha^{-1}\beta$.

> Throughout the rest of this paper, we assume $\mathbb{G}$ to be a finitely generated, infinitary, commutative[5] group, 1 its neutral element and Γ to be a finite generating set of $\mathbb{G}$.

The *undirected Cayley graph* for $\mathbb{G}$ and Γ is the graph (V, E), where $V = \mathbb{G}$ is the set of vertices and for every $\alpha \in \mathbb{G}$ and $\beta \in \Gamma$, we have that $(\alpha, \alpha\beta), (\alpha\beta, \alpha) \in E$. The *Cayley distance* $d(\alpha, \beta)$ between $\alpha \in \mathbb{G}$ and $\beta \in \mathbb{G}$ is defined as the length of the shortest path between α and β in the undirected Cayley Graph. For $\alpha \in \mathbb{G}$ we define the *Γ-length* of α as the Cayley distance between 1 and α and denote it by $|\alpha|_\Gamma$.

[5] In fact, we do not require commutativity for the proof our results. However, in order to limit the notational complexity of the present paper, we require $\mathbb{G}$ to be commutative.

2.1 Weighted tree automata

Definition 2. A *(group-)weighted tree automaton* over Σ and $\mathbb{G}$ (short: group-WTA or simply WTA) is a tuple $(Q, \Sigma, \mathbb{G}, \mathrm{final}, T)$, where Q is a finite set of *states*, $\mathrm{final} \subseteq Q \times \mathbb{G}$ is the finite *final relation*, and T is a family $(T_\sigma \subseteq Q^s \times \mathbb{G} \times Q \mid s \geq 0, \sigma \in \Sigma^{(s)})$ of finite sets of *transitions*.

We call $q \in Q$ *final* if there exists $\alpha \in \mathbb{G}$ such that $(q, \alpha) \in \mathrm{final}$, which we depict as $q \xrightarrow{\alpha}$. For every $\sigma \in \Sigma^{(s)}$ and $t = (q_1, \dots, q_s, \alpha, q) \in T_\sigma$, we denote $\mathrm{out}(t) \coloneqq q$, $\mathrm{in}(t) \coloneqq (q_1, \dots, q_s)$, and $\mathrm{wt}(t) \coloneqq \alpha$. For notational convenience, we use the notation $\mathrm{out}(q) \coloneqq q$ and $\mathrm{wt}(q) \coloneqq 1$. To aid readability, we denote the fact that $(q_1, \dots, q_s, \alpha, q) \in T_\sigma$ by $\sigma(q_1, \dots, q_s) \xrightarrow{\alpha} q$.

Definition 3. Let $\mathscr{A} = (Q, \Sigma, \mathbb{G}, \mathrm{final}, T)$ be a WTA and let $\xi \in \mathrm{T}_\Sigma \cup \mathrm{C}_\Sigma$ be a tree or a context. A *run* of $\mathscr{A}$ on ξ is a map $\rho \colon \mathrm{pos}(\xi) \to T \cup Q$ such that

- for every $w \in \mathrm{pos}_\Sigma(\xi)$ we have $\rho(w) \in T_\sigma$ where $\sigma = \xi(w)$ and $\mathrm{in}(\rho(w)) = (\mathrm{out}(\rho(w1)), \dots, \mathrm{out}(\rho(ws)))$ where $s = \mathrm{rk}(\sigma)$, and
- for every $w \in \mathrm{pos}_X(\xi)$ we have $\rho(w) \in Q$.

We denote by $\mathrm{out}(\rho)$ the state $\mathrm{out}(\rho(\varepsilon))$ and if $\xi \in \mathrm{C}_\Sigma$ we denote by $\mathrm{in}(\rho)$ the state $\rho(w)$ where w is the unique position in $\mathrm{pos}_X(\xi)$. The *weight* of such a run ρ is $\mathrm{wt}(\rho) \coloneqq \prod_{w \in \mathrm{pos}(\xi)} \mathrm{wt}(\rho(w))$. Moreover, we say that ρ *contains* a state $q \in Q$ if there exists $w \in \mathrm{pos}(\xi)$ such that $q = \mathrm{out}(\rho(w))$. A run ρ is called *accepting* if $\mathrm{out}(\rho)$ is final.

Remark 4. We use the following notation for a run ρ of $\mathscr{A}$ on a tree or context ξ. Let $q \coloneqq \mathrm{out}(\rho)$ and $\alpha \coloneqq \mathrm{wt}(\rho)$. If $\xi \in T_\Sigma$, then we write $\xrightarrow{\xi|\rho|\alpha} q$. If $\xi \in C_\Sigma$, then we write $p \xrightarrow{\xi|\rho|\alpha} q$, where $p \coloneqq \mathrm{in}(\rho)$. Whenever we do not care about the name of the run, we simply write $\xrightarrow{\xi|\alpha} q$ and $p \xrightarrow{\xi|\alpha} q$, respectively. Furthermore, if $\xrightarrow{\xi|\alpha} q$ for some tree ξ and some weight α, then we call the state q *reachable*.

Remark 5. Throughout this paper, we assume that all considered WTA are *trim*. For a WTA $\mathscr{A}$, this condition means that every state appears in some accepting run. In particular, for every state p, there exists a context $\xi \in \mathrm{C}_\Sigma$, a final state q, and a run $p \xrightarrow{\xi|\alpha} q$.

Note moreover that, without loss of generality, the size of ξ is bounded. If the run on ξ contains a single state q' multiple times on a single branch (excluding the root of the tree), then we can replace the subtree at the topmost occurrence of q' with the subtree at the bottommost occurrence of q'. Therefore, we can assume $\mathrm{height}(\xi) \leq \#Q + 1$ and hence $\mathrm{size}(\xi) \leq \mathrm{maxrk}(\Sigma)^{\#Q+1}$.

Definition 6. Let $\mathscr{A} = (Q, \Sigma, \mathbb{G}, \mathrm{final}, T)$ be a WTA. The *weighted tree language accepted by* $\mathscr{A}$ is the relation $[\![\mathscr{A}]\!] \subseteq \mathrm{T}_\Sigma \times \mathbb{G}$ containing the pairs $(\xi, \beta\gamma)$ such that $\xrightarrow{\xi|\beta} q \xrightarrow{\gamma}$ for some $q \in Q$.

Two WTA $\mathscr{A}$ and $\mathscr{B}$ are called *equivalent* if they accept the same weighted tree language, that is, $[\![\mathscr{A}]\!] = [\![\mathscr{B}]\!]$.

Moreover, we define the constant

$$M_{\mathscr{A}} \coloneqq \max\{|\alpha|_\Gamma \mid (q_1,\dots,q_k,\alpha,q) \in \bigcup_{\sigma\in\Sigma} T_\sigma \text{ or } (q,\alpha)\in \text{final}\}.$$

That is, $M_{\mathscr{A}}$ is the maximal Γ-length of weights occurring in T or final.

We will now briefly compare group-WTA to semiring-WTA[6].

Remark 7. Consider the tuple $S = (\mathcal{P}_{\text{fin}}(\mathbb{G}), \cup, \cdot, \emptyset, \{1\})$, where $\mathcal{P}_{\text{fin}}(\mathbb{G})$ is the set of finite subsets of $\mathbb{G}$, $1 \in \mathbb{G}$ is the neutral group element, and $\cdot$ is the group operation lifted to finite sets. It is immediate that S is a semiring.

Let $\mathscr{A} = (Q, \Sigma, \mathbb{G}, \text{final}, T)$ be a group-WTA. In order to syntactically match the definition of group-WTA over $\mathbb{G}$ with the definition of semiring-WTA over S, we replace T_σ by the map $T_\sigma^{\text{sr}} \colon Q^s \times Q \to \mathcal{P}_{\text{fin}}(\mathbb{G})$ such that $(q_1,\dots,q_s,q) \mapsto \{\alpha \mid (q_1,\dots,q_s,\alpha,q) \in T_\sigma\}$. Furthermore we replace the final relation by $\text{final}^{\text{sr}} \colon Q \to \mathcal{P}_{\text{fin}}(\mathbb{G})$ such that $q \mapsto \{\beta \mid (q,\beta) \in \text{final}\}$. Denote the semiring-WTA $\mathscr{A}^{\text{sr}} = (Q, T^{\text{sr}}, \text{final}^{\text{sr}})$ and note that $[\![\mathscr{A}^{\text{sr}}]\!](\xi) = \{\alpha \mid (\xi,\alpha) \in [\![\mathscr{A}]\!]\}$. Therefore, up to this identification of maps and relations, group-WTA are particular semiring-WTA. However, the important difference is that each run of $\mathscr{A}$ calculates a *single* group element, whereas each run of $\mathscr{A}^{\text{sr}}$ calculates *multiple* aggregated group elements at once.

Definition 8. Let $\mathscr{A}_i = (Q_i, \Sigma, \mathbb{G}, \text{final}_i, T_i)$ be WTA for $i \in [2]$.

The *union* of $\mathscr{A}_1$ and $\mathscr{A}_2$, denoted $\mathscr{A}_1 \cup \mathscr{A}_2$, is the WTA $(Q_1 \cup Q_2, \Sigma, \mathbb{G}, \text{final}_1 \cup \text{final}_2, T_1 \cup T_2)$, where we (without loss of generality) assume that $Q_1 \cap Q_2 = \emptyset$. This definition naturally extends to finitely many WTA.

The *direct product* of $\mathscr{A}_1$ and $\mathscr{A}_2$, denoted $\mathscr{A}_1 \times \mathscr{A}_2$, is the WTA $(Q_1 \times Q_2, \Sigma, \mathbb{G}, \text{final}, T)$, where $\text{final} \coloneqq \{((q,p), \alpha\beta) \mid (q,\alpha) \in \text{final}_1 \wedge (p,\beta) \in \text{final}_2\}$ and

$$\begin{aligned} T_\sigma \coloneqq \{&((q_1,p_1),\dots,(q_s,p_s)), \alpha\beta, (q,p)) \\ &\mid (q_1,\dots,q_s,\alpha,q) \in (T_1)_\sigma \wedge (p_1,\dots,p_s,\beta,p) \in (T_2)_\sigma\}. \end{aligned}$$

Again, without loss of generality, we assume that $Q_1 \cap Q_2 = \emptyset$. This definition naturally extends to finitely many WTA.

Definition 9. Let $\mathscr{A} = (Q, \Sigma, \mathbb{G}, \text{final}, T)$ be a WTA. We call $\mathscr{A}$ *sequential* if for all $s \geq 0, \sigma \in \Sigma^{(s)}$, and $q_1,\dots,q_s \in Q$ there exist at most one $\alpha \in \mathbb{G}$ and $q \in Q$ such that $(q_1,\dots,q_s,\alpha,q) \in T_\sigma$.

A relation $R \subseteq T_\Sigma \times \mathbb{G}$ is called *sequential* if there exists a sequential WTA $\mathscr{A}$ such that $[\![\mathscr{A}]\!] = R$.

[6] As a reference, we use the definition of semiring-weighted tree automata from [5]. For a more thorough introduction to semirings confer [12] and for semiring-WTA we refer to [11].

Remark 10. Note that Definition 9 is highly similar to the definition of *deterministic* semiring-WTA [5, preceeding Example 3.1]. However, sequentiality forces the weight of transitions to be at most one *single* group element, whereas determinism merely forces the weight of transitions to be at most one *set* of group elements. This difference results in sequentiality being a properly more restrictive condition on the automaton than determinism.

Example 11. Let $\Sigma = \{\sigma^{(2)}, \alpha^{(0)}\}$ and $\mathbb{G} = (\mathbb{Z}, +, 0)$. Note that $\mathbb{G}$ is a commutative, finitely generated, infinitary group with finite generating set $\Gamma = \{1\}$. Define the WTA $\mathscr{A} \coloneqq (Q, \Sigma, \mathbb{G}, \mathrm{final}, T)$ where $Q \coloneqq \{q_\alpha, q_0, q_1\}$, $\mathrm{final} \coloneqq \{(q_0, 0)\}$ and T is defined by

$$\begin{aligned} T_\alpha \cup T_\sigma = \{\alpha \xrightarrow{0} q_\alpha, \qquad & \sigma(q_\alpha, q_\alpha) \xrightarrow{1} q_1, \qquad && \sigma(q_\alpha, q_\alpha) \xrightarrow{3} q_0, \\ & \sigma(q_\alpha, q_1) \xrightarrow{1} q_0, && \sigma(q_\alpha, q_0) \xrightarrow{1} q_1\}. \end{aligned}$$

Consider the context $\eta = \sigma(\alpha, x_1) \in \mathrm{C}_\Sigma$. One easily sees that all trees $\xi \in \mathrm{T}_\Sigma$ occurring in $[\![\mathscr{A}]\!]$ are of the form $\xi = \eta^\ell[\alpha]$ for some $\ell \geq 1$. In this case, if $\#\mathrm{pos}_{\{\sigma\}}(\xi) = 2n$ for some $n \in \mathbb{N}$ we have $(\xi, 2n) \in [\![\mathscr{A}]\!]$ and if $\#\mathrm{pos}_{\{\sigma\}}(\xi) = 2n+1$ we have $(\xi, 2n+3) \in [\![\mathscr{A}]\!]$. Clearly, $\mathscr{A}$ is not sequential.

3 Lipschitz and Twinning Property

In this section we formally introduce the two characterisations of sequentiality from Theorem 1, and prove the implications "sequential $\implies$ Lipschitz" in Theorem 13 and "Lipschitz $\implies$ twinning" in Theorem 16.

3.1 The Lipschitz Property

Definition 12. A relation $R \subseteq T_\Sigma \times \mathbb{G}$ satisfies the *Lipschitz property* if there exists $L \in \mathbb{N}$ such that for all pairs $(\xi_0, \alpha_0), (\xi_1, \alpha_1) \in R$ it holds that $d(\alpha_0, \alpha_1) \leq L \cdot (\mathrm{dist}(\xi_0, \xi_1) + 1)$.

Theorem 13. Let $R \subseteq T_\Sigma \times \mathbb{G}$ be a sequential relation. Then R satisfies the Lipschitz property.

The proof of Theorem 13 primarily uses the fact that a sequential WTA has a unique non-vanishing run weight on every overlap of two input trees ξ_1 and ξ_2.

3.2 The Twinning Property

Throughout the rest of this paper, we assume $\mathscr{A} = (Q, \Sigma, \mathbb{G}, \mathrm{final}, T)$ to be a WTA.

Definition 14. We say that $\mathscr{A}$ satisfies the *twinning property* if for all runs ρ_0 and ρ_1 of $\mathscr{A}$, states $q_0, q_1 \in Q$, trees $\xi \in \mathrm{T}_\Sigma$, and contexts $\zeta \in \mathrm{C}_\Sigma$, such that ρ_j $(j = 0, 1)$ equals

$$\xrightarrow{\xi|\alpha_j} q_j \xrightarrow{\zeta|\beta_j} q_j,$$

it holds that $\beta_0 = \beta_1$.

Example 15. We continue Example 11 by showing that $\mathscr{A}$ satisfies the twinning property. Let ρ_0 and ρ_1 be runs of $\mathscr{A}$ quantified as in Definition 14. Recall that $\zeta[\xi]$ has the form $\eta^\ell[\alpha]$. Moreover, non-empty runs cannot loop in the state q_α by definition of the transition relation. Therefore, we have that $\xi \neq \alpha$, whence $\xi = \eta^j[\alpha]$ for some $1 \leq j < \ell$. However, in this case the single non-deterministic choice already occurs in ξ and hence $\beta_0 = \beta_1 = \#\mathrm{pos}_{\{\sigma\}}(\zeta)$. In particular, this proves the twinning property.

Next we provide a WTA $\mathscr{B}$ over $\Delta = \{\sigma^{(2)}, \alpha^{(0)}, \beta^{(0)}\}$ and $\mathbb{G} = (\mathbb{Z}, +, 0)$ which does not satisfy the twinning property. Define $\mathscr{B} = (\widetilde{Q}, \Delta, \mathbb{G}, \widetilde{\mathrm{final}}, \widetilde{T})$, where $\tilde{Q} = \{q_\alpha, q_\beta\}$, $\widetilde{\mathrm{final}} = \{(q_\alpha, 0), (q_\beta, 0)\}$, and

$$\begin{aligned}
\widetilde{T}_\alpha \cup \widetilde{T}_\beta \cup \widetilde{T}_\sigma = \{\alpha \xrightarrow{1} q_\alpha, \quad & \beta \xrightarrow{0} q_\alpha, \quad \sigma(q_\alpha, q_\alpha) \xrightarrow{0} q_\alpha, \qquad \text{(counting } \alpha) \\
\alpha \xrightarrow{0} q_\beta, \quad & \beta \xrightarrow{1} q_\beta, \quad \sigma(q_\beta, q_\beta) \xrightarrow{0} q_\beta\}. \qquad \text{(counting } \beta)
\end{aligned}$$

One easily verifies the fact that for every $\xi \in \mathrm{T}_\Sigma$ there are exactly two runs of $\mathscr{B}$ on ξ and we obtain $(\xi, \#\mathrm{pos}_{\{\alpha\}}(\xi)), (\xi, \#\mathrm{pos}_{\{\beta\}}(\xi)) \in [\![\mathscr{B}]\!]$. Clearly, $\mathscr{B}$ is not sequential. Consider the tree $\xi = \sigma(\alpha, \beta)$, the context $\zeta = \sigma(\alpha, x_1)$, and the two runs of $\mathscr{B}$ on $\zeta[\xi]$

$$\xrightarrow{\xi|\alpha_0} q_\alpha \xrightarrow{\zeta|\beta_0} q_\alpha \qquad \text{and} \qquad \xrightarrow{\xi|\alpha_1} q_\beta \xrightarrow{\zeta|\beta_1} q_\beta.$$

By the definition of $\widetilde{T}$ we calculate the values $\beta_0 = 0 + 1 + 0$ and $\beta_1 = 0 + 0 + 0$. This proves $\beta_0 \neq \beta_1$, whence we obtain that $\mathscr{B}$ does not satisfy the twinning property.

Theorem 16. If $[\![\mathscr{A}]\!]$ satisfies the Lipschitz property, then $\mathscr{A}$ satisfies the twinning property.

The proof of Theorem 16 is done by contradiction. We take a witness of the non-satisfaction of the twinning property and pump the occurring loops. This makes the run weights diverge in $\mathbb{G}$ (using the fact that $\mathbb{G}$ is infinitary), which contradicts the Lipschitz property.

Remark 17. Note that Theorem 16 implies that, whenever $\mathscr{A}$ does not satisfy the twinning property, no equivalent automaton can satisfy the twinning property .

4 Sequentiality of the Twinning Property

This section executes the proof of the implication "twinning $\implies$ sequential" of Theorem 1. Given a WTA $\mathscr{A}$, we apply a construction similar to the well-known power set construction to $\mathscr{A}$. This yields a (not necessarily finite) sequential WTA $D_{\mathscr{A}}$. However, we prove that $D_{\mathscr{A}}$ is indeed finite if $\mathscr{A}$ satisfies the twinning property (see Corollary 25). The proof of Corollary 25 can be outlined as follows. First we show that all runs of $\mathscr{A}$ on a fixed input tree generate close weights with respect to the Cayley-distance (see Lemma 23). Next, the definition of $D_{\mathscr{A}}$

implies that every (reachable) state of $D_{\mathscr{A}}$ contains only weights that are close to the neutral element $1 \in \mathbb{G}$ (see Lemma 24), which implies that the set of reachable states of $D_{\mathscr{A}}$ is finite. We derive the fact that $D_{\mathscr{A}}$ is equivalent to $\mathscr{A}$ from the definition of $D_{\mathscr{A}}$ (see Theorem 26). We conclude this chapter by applying our construction to the automata from Examples 11 and 15.

Theorem 18. If $\mathscr{A}$ satisfies the twinning property, then $[\![\mathscr{A}]\!]$ is sequential.

Throughout this section, we assume $\mathscr{A}$ to satisfy the twinning property.

Definition 19. We define the infinite WTA[7] $D_{\mathscr{A}} = (Q', \Sigma, \mathbb{G}, \mathrm{final}', T')$ as follows. The states of $D_{\mathscr{A}}$ are $Q' := \mathcal{P}(Q \times \mathbb{G})$, the final relation is

$$\mathrm{final}' := \{(S, \alpha\beta) \mid S \in Q', \exists q \in Q \colon (q, \alpha) \in S \text{ and } (q, \beta) \in \mathrm{final}\},$$

and the transitions are constructed as follows. For every $\sigma \in \Sigma^{(s)}$ and $S_1, \ldots, S_s \in Q'$, consider the set

$$S := \{(q, \alpha_1 \cdots \alpha_s \beta) \mid \exists p_1, \ldots, p_s \in Q \colon \\ (\forall i \in [s] \colon (p_i, \alpha_i) \in S_i) \text{ and } (p_1, \ldots, p_s, \beta, q) \in T_\sigma\}$$

and fix an arbitrary element[8] $(p, \alpha) \in S$. We define the set

$$S' := \{(q, \alpha^{-1}\gamma) \mid (q, \gamma) \in S\}$$

and ultimately add $(S_1, \ldots, S_s, \alpha, S')$ to T'_σ.

Remark 20. Note that $D_{\mathscr{A}}$ is indeed sequential. This follows directly from the construction. Moreover, in Definition 19 we first calculate an intermediate successor state S, which is then shifted by a fixed value α occurring in some pair $(p, \alpha) \in S$. We call this shifting process the *factorisation* of S.

We will show in Corollary 25 that every reachable state S of $D_{\mathscr{A}}$ satisfies $\#S \leq K$ for a global constant K and hence after trimming $D_{\mathscr{A}}$, also the final relation final' is finite.

Lemma 21. Let $\xi \in \mathrm{T}_\Sigma$ and consider the (unique) run of $D_{\mathscr{A}}$ on ξ, $\xrightarrow{\xi|\alpha} S'$. It holds that

$$S' = \{(q, \beta) \mid \exists \text{ run } \xrightarrow{\xi|\delta} q \text{ of } \mathscr{A} \colon \alpha\beta = \delta\}.$$

Definition 22. We define the constant $N_{\mathscr{A}} := 2M_{\mathscr{A}} \mathrm{maxrk}(\Sigma)^{(\#Q^2+1)}$.

Lemma 23. For every tree $\xi \in \mathrm{T}_\Sigma$ and every two runs $\xrightarrow{\xi|\alpha} q$ and $\xrightarrow{\xi|\beta} p$ of $\mathscr{A}$ on ξ it holds that

$$d(\alpha, \beta) < N_{\mathscr{A}}.$$

[7] That is, a tuple satisfying the conditions of a WTA, except for finiteness.

[8] Formally, we have a globally fixed choice function $f \colon \mathcal{P}(Q \times \mathbb{G}) \to Q \times \mathbb{G}$ and then simply define $(p, \alpha) := f(S)$.

Note that the proof of Lemma 23 uses the fact that $\mathscr{A}$ satisfies the twinning property.

Lemma 24. Let S be a reachable state of $D_{\mathscr{A}}$ and let $(q,\alpha) \in S$. It holds that $|\alpha|_\Gamma \leq N_{\mathscr{A}}$.

Proof. The reachability of S implies the existence of a tree $\xi \in \mathrm{T}_\Sigma$ such that there exists a run $\xrightarrow{\xi|\delta} S$. If $\alpha = 1$, then we are done. Therefore, we assume that $\alpha \neq 1$ and hence there exists $(p,\beta) \in S$ such that $\beta = 1$. Therefore by Lemma 21, there are two runs of $\mathscr{A}$ on ξ, $\xrightarrow{\xi|\delta\alpha} q$ and $\xrightarrow{\xi|\delta} p$. By Lemma 23 it holds that $d(\delta\alpha,\delta) < N_{\mathscr{A}}$. The fact that $d(\delta\alpha,\delta) = |\alpha|_\Gamma$ implies that $|\alpha|_\Gamma < N_{\mathscr{A}}$. □

Corollary 25. The set of states of $D_{\mathscr{A}}$ is finite and hence $D_{\mathscr{A}}$ is a WTA.

Proof. Denote for every $N \in \mathbb{N}$ the (finite) set $\mathbb{G}_N \coloneqq \{g \in \mathbb{G} \mid |g|_\Gamma \leq N\}$.

By Lemma 24 every reachable state of $D_{\mathscr{A}}$ is an element of the finite set $\mathcal{P}(Q \times \mathbb{G}_{N_{\mathscr{A}}})$, which proves the claim. □

Theorem 26. $D_{\mathscr{A}}$ is equivalent to $\mathscr{A}$.

Proof. We first show that $[\![D_{\mathscr{A}}]\!] \subseteq [\![\mathscr{A}]\!]$. Let $(\xi,\alpha) \in [\![D_{\mathscr{A}}]\!]$ and let $\xrightarrow{\xi|\beta} S$ be a run of $D_{\mathscr{A}}$ on ξ and $(S,\gamma) \in \mathrm{final}'$ such that $\alpha = \beta\gamma$. Note that by the definition of final' there exist $q \in Q$, $(q,\beta') \in S$ and $(q,\gamma') \in \mathrm{final}$ such that $\gamma = \beta'\gamma'$.

By Lemma 21 there exists a run $\xrightarrow{\xi|\delta} q$ of $\mathscr{A}$ such that $\beta\beta' = \delta$. Hence, $\alpha = \beta\gamma = \beta\beta'\gamma' = \delta\gamma'$ and therefore $(\xi,\alpha) = (\xi,\delta\gamma') \in [\![\mathscr{A}]\!]$.

To prove the fact that $[\![\mathscr{A}]\!] \subseteq [\![D_{\mathscr{A}}]\!]$, we apply a similar argument. Let $(\xi,\alpha) \in [\![\mathscr{A}]\!]$. There is a unique run $\xrightarrow{\xi|\beta} S$ of $D_{\mathscr{A}}$ on ξ. By definition of $[\![\mathscr{A}]\!]$, there exist a run $\xrightarrow{\xi|\delta} q$ of $\mathscr{A}$ on ξ and a pair $(q,\gamma') \in \mathrm{final}$ such that $\delta\gamma' = \alpha$.

By Lemma 21 there exists an element $(q,\beta') \in S$ such that $\beta\beta' = \delta$ and hence by the definition of final' we obtain $(S,\beta'\gamma') \in \mathrm{final}'$. We obtain $\alpha = \delta\gamma' = \beta\beta'\gamma'$ and hence $(\xi,\alpha) \in [\![D_{\mathscr{A}}]\!]$. □

Proof of Theorem 18. We have seen that $D_{\mathscr{A}}$ is a sequential (Remark 20) WTA (Corollary 25) which is equivalent to $\mathscr{A}$ (Theorem 26). This proves the claim. □

Example 27. Recall the WTA $\mathscr{A}$ and $\mathscr{B}$ from Examples 11 and 15. We apply the construction from Definition 19 to both, $\mathscr{A}$ and $\mathscr{B}$, and obtain that $D_{\mathscr{A}}$ has a finite trim state space, whereas $D_{\mathscr{B}}$ has an infinite trim state space.

First we consider $D_{\mathscr{A}}$. Clearly, $T'_\alpha = \{(0,S_0)\}$, where $S_0 = \{(q_\alpha,0)\}$. By pointwise application of T_σ to (S_0,S_0) we obtain $\{(q_1,1),(q_0,3)\}$. We chose $(q_1,1)$ for the factorisation, which yields the new state $S_1 = \{(q_1,0),(q_0,2)\}$ and therefore we have constructed the transition $(S_0,S_0,1,S_1) \in T'_\sigma$. By continuing this process we arrive at

$$T'_\sigma = \{(S_0,S_0,1,S_1),(S_0,S_1,1,S_2),(S_0,S_2,1,S_1)\},$$

where $S_2 = \{(q_0, 0), (q_1, 2)\}$. The trim state space of $D_{\mathscr{A}}$ is $Q' = \{S_0, S_1, S_2\}$ and the final relation is $\mathrm{final}' = \{(S_1, 2), (S_2, 0)\}$.

Next we consider $D_{\mathscr{B}}$. Define $R_1 = \{(q_\alpha, 1), (q_\beta, 0)\}$ and note that $\widetilde{T}'_\alpha = \{(0, R_1)\}$. Pointwise application of $\widetilde{T}_\sigma$ to (R_1, R_1) yields $R_2 = \{(q_\alpha, 2), (q_\beta, 0)\}$, which is already normalised. Another pointwise application of $\widetilde{T}_\sigma$ to (R_1, R_2) results in $R_3 = \{(q_\alpha, 3), (q_\beta, 0)\}$, which is again normalised. One easily sees that repeatedly generating transitions of $D_{\mathscr{B}}$ like this yields an infinite set of reachable states of $D_{\mathscr{B}}$ and hence $D_{\mathscr{B}}$ is not a WTA. In Section 5 we will discuss the approach given in [7, case $k > 1$], which describes how to handle $D_{\mathscr{B}}$ in order to generate a finite union of sequential WTA which is equivalent to $\mathscr{B}$.

5 Outlook

In the present paper, we have successfully lifted the result from [7, case $k = 1$] to weighted tree automata. Recall that [7] characterises unions of k sequential automata. The natural next step is to lift the remaining cases $k > 1$. This section is designed to briefly demonstrate why a straightforward lift of [7, cases $k > 1$] to weighted tree automata fails.

The outline of the proof given in [7] goes as follows. Let $k \in \mathbb{N}$. The notions of k-sequential WTA, the k-Lipschitz property, and the k-branching twinning property are introduced and the directions "sequential $\implies$ Lipschitz" and "Lipschitz $\implies$ twinning" are proven similarly to our Theorems 13 and 16. For the direction "twinning $\implies$ sequential", the automaton $D_{\mathscr{A}}$ is introduced and its properties are studied. As we have seen in the second part of Example 27, $D_{\mathscr{A}}$ is in general infinite. However, if $\mathscr{A}$ satisfies the k-branching twinning property, [7] describes the following construction on $D_{\mathscr{A}}$, yielding a k-sequential automaton which is equivalent to $\mathscr{A}$. First, the set of states of $D_{\mathscr{A}}$ is restricted to a finite set. In fact, the set of reachable states S of $D_{\mathscr{A}}$ containing only "small" weights $|\alpha|_\Gamma < N_{\mathscr{A}}$ is denoted U and the set of states reachable from U in one step is denoted U'. Note that U and U' are finite. $D_{\mathscr{A}}$ is restricted to $U \cup U'$ and each state S in $U' \setminus U$ (i.e. the outer border of U) is replaced by a union of k sequential WTA. These sequential WTA are constructed by induction on k and depend on the state S. The resulting automaton $\bar{D}_{\mathscr{A}}$ can easily be divided into k sequential automata, which concludes the proof.

The tree case differs in the following way. Consider a symbol $\sigma \in \Sigma^{(2)}$ and consider two different states $S, S' \in U' \setminus U$. Surely, in $D_{\mathscr{A}}$ we can find a transition of the form $\sigma(S, S') \to S''$. However, the states S and S' are replaced by different automata in $\bar{D}_{\mathscr{A}}$. Therefore, a run ρ of $D_{\mathscr{A}}$ on a tree ξ ending in S (resp. S') translates into a run of $\bar{D}_{\mathscr{A}}$ on ξ ending in some state $q_S \notin Q'$ (resp. $q_{S'} \notin Q'$). Moreover, q_S and $q_{S'}$ are taken from disjoint sets. We have not been able to find the proper way to construct a transition of the form $\sigma(q_S, q_{S'}) \to q$.

Therefore, we leave the lift of [7, cases $k > 1$] as an open research question.

References

1. Athanasios Alexandrakis and Symeon Bozapalidis. Weighted grammars and Kleene's theorem. *Information Processing Letters*, 24(1):1–4, 1987.
2. Marie-Pierre Béal and Olivier Carton. Determinization of transducers over finite and infinite words. *Theoretical Computer Science*, 289(1):225–251, 2002.
3. Björn Borchardt. A Pumping Lemma and Decidability Problems for Recognizable Tree Series. *Acta Cybernetica*, 16(4):509–544, September 2004.
4. Björn Borchardt and Heiko Vogler. Determinization of finite state weighted tree automata. *Journal of Automata, Languages and Combinatorics*, 8(3):417–463, 2003.
5. Matthias Büchse, Jonathan May, and Heiko Vogler. Determinization of Weighted Tree Automata Using Factorizations. *Journal of Automata, Languages and Combinatorics*, 15(3/4):229–254, 2010.
6. Christian Choffrut. Une Caracterisation des Fonctions Sequentielles et des Fonctions Sous-Sequentielles en tant que Relations Rationnelles. *Theoretical Computer Science*, 5(3):325–337, 1977. `doi:10.1016/0304-3975(77)90049-4`.
7. Laure Daviaud, Ismaël Jecker, Pierre-Alain Reynier, and Didier Villevalois. Degree of Sequentiality of Weighted Automata. In *FOSSACS 2017*, volume 10203 of LNCS, pages 215–230, 2017. `doi:10.1007/978-3-662-54458-7\_13`.
8. Manfred Droste, Werner Kuich, and Heiko Vogler, editors. *Handbook of Weighted Automata*. EATCS Monographs in Theoretical Computer Science. Springer-Verlag, 2009.
9. Manfred Droste, Torsten Stüber, and Heiko Vogler. Weighted finite automata over strong bimonoids. *Information Sciences*, 180:156–166, 2010.
10. Emmanuel Filiot, Raffaella Gentilini, and Jean-François Raskin. Quantitative languages defined by functional automata. In *International Conference on Concurrency Theory*, pages 132–146. Springer, 2012.
11. Zoltán Fülöp and Heiko Vogler. Weighted tree automata and tree transducers. In M. Droste, W. Kuich, and H. Vogler, editors, *Handbook of Weighted Automata*, chapter 9, pages 313–403. Springer-Verlag, 2009.
12. Jonathan Golan. *Semirings and their Applications*. Kluwer Academic Publishers, Dordrecht, 1999.
13. Daniel Kirsten and Ina Mäurer. On the Determinization of Weighted Automata. *Journal of Automata, Languages and Combinatorics*, 10:287–312, January 2005.
14. Jonathan May and Kevin Knight. A Better N-Best List: Practical Determinization of Weighted Finite Tree Automata. In *Proceedings of the Human Language Technology Conference of the NAACL, Main Conference*, pages 351–358, New York City, USA, June 2006. Association for Computational Linguistics.
15. Mehryar Mohri. Finite-State Transducers in Language and Speech Processing. *Computational Linguistics*, 23(2):269–311, June 1997.
16. Dragica Radovanovic. Weighted tree automata over strong bimonoids. *Novi Sad Journal of Mathematics*, 40(3):89–108, 2010.
17. Grzegorz Rozenberg and Arto Salomaa, editors. *Handbook of Formal Languages, Vol. 1: Word, Language, Grammar*. Springer-Verlag, Berlin, Heidelberg, 1997.
18. Grzegorz Rozenberg and Arto Salomaa, editors. *Handbook of Formal Languages, Vol. 3: Beyond Words*. Springer-Verlag, Berlin, Heidelberg, 1997.

A Appendix

Lemma 28. For every $\alpha, \beta, \gamma \in \mathbb{G}$ it holds that $|\alpha\beta\gamma|_\Gamma \geq |\beta|_\Gamma - |\alpha|_\Gamma - |\gamma|_\Gamma$

Proof. We know that $|\delta|_\Gamma = |\delta^{-1}|_\Gamma$ for every $\delta \in \mathbb{G}$ by definition of the Cayley graph of $\mathbb{G}$. Hence by $|\beta|_\Gamma = |\alpha^{-1}\alpha\beta\gamma\gamma^{-1}|_\Gamma \leq |\alpha|_\Gamma + |\alpha\beta\gamma|_\Gamma + |\gamma|_\Gamma$ we obtain the desired inequality. □

The following lemma collects some straightforward formulas for delays.

Lemma 29 (Lemmas 1 and 2 from [7]). For all $\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma_0, \gamma_1 \in \mathbb{G}$ it holds that

1. $\mathrm{delay}(\alpha_0, \beta_0) = 1$ iff $\alpha_0 = \beta_0$,
2. $\mathrm{delay}(\alpha_0, \alpha_1) = \mathrm{delay}(\beta_0, \beta_1)$ implies $\mathrm{delay}(\alpha_0\gamma_0, \alpha_1\gamma_1) = \mathrm{delay}(\beta_0\gamma_0, \beta_1\gamma_1)$,
3. $d(\alpha_0, \beta_0) = |\mathrm{delay}(\alpha_0, \beta_0)|_\Gamma$.

We now provide the missing proofs for our paper.

Proof of Theorem 13. By assumption there exists a sequential WTA $\mathscr{A}$ such that $[\![\mathscr{A}]\!] = R$.

Let $(\xi_0, \alpha_0), (\xi_1, \alpha_1) \in [\![\mathscr{A}]\!]$ and let (ζ_0, ζ_1, η) be a pair-cut between ξ_0 and ξ_1 with a maximal overlap η. We know that there exist runs ρ_i on ξ_i and elements $(q_i, \beta_i) \in \mathrm{final}$ for $i \in \{0, 1\}$, such that

$$\begin{aligned}\alpha_i = \mathrm{wt}(\rho_i) \cdot \beta_i &= \prod_{w \in \mathrm{pos}(\xi_i)} \mathrm{wt}(\rho_i(w)) \cdot \beta_i \\ &= \prod_{w \in \mathrm{pos}(\zeta_i)} \mathrm{wt}(\rho_i(w)) \cdot \prod_{w \in \mathrm{pos}(\eta)} \mathrm{wt}(\rho_i(v_i w)) \cdot \beta_i,\end{aligned}$$

where $v_i \in \mathrm{pos}(\zeta_i)$ is the unique position of x_1 in ζ_i. Because $\mathscr{A}$ is sequential, we have that $\mathrm{wt}(\rho_0(v_0 w)) = \mathrm{wt}(\rho_1(v_1 w))$ for every $w \in \mathrm{pos}(\eta)$. Thus we deduce

$$\begin{aligned}d(\alpha_0, \alpha_1) &= d\Big(\prod_{w \in \mathrm{pos}(\zeta_0)} \mathrm{wt}(\rho_0(w)) \cdot \beta_0, \prod_{w \in \mathrm{pos}(\zeta_1)} \mathrm{wt}(\rho_1(w)) \cdot \beta_1\Big) \\ &\leq M_{\mathscr{A}}(\mathrm{size}(\zeta_0) + \mathrm{size}(\zeta_1)) + 2M_{\mathscr{A}} \leq 2M_{\mathscr{A}}(\mathrm{dist}(\xi_0, \xi_1) + 1).\end{aligned}$$

This proves the Lipschitz property. □

Lemma 30. If $\mathscr{A}$ does not satisfy the twinning property, then for all $L \in \mathbb{N}$, there are two runs of $\mathscr{A}$, $\xrightarrow{\xi|\alpha_0} p_0$ and $\xrightarrow{\xi|\alpha_1} p_1$ (for some $\xi \in \mathrm{T}_\Sigma$ and $p_0, p_1 \in Q$), such that it holds that $d(\alpha_0, \alpha_1) > L$.

Proof. The idea of this proof is as follows. We quantify an instance of the twinning property that is not satisfied and pump the loop contexts until the weights arrive at a high Cayley-distance. This uses the fact that $\mathbb{G}$ is infinitary.

Let $L \in \mathbb{N}$. As $\mathscr{A}$ does not satisfy the twinning property, there are runs ρ_0 and ρ_1 of $\mathscr{A}$, states $q_0, q_1 \in Q$, a tree $\xi \in \mathrm{T}_\Sigma$, and contexts $\zeta \in \mathrm{C}_\Sigma$, such that ρ_j $(j = 0, 1)$ equals

$$\xrightarrow{\xi|\alpha_j} q_j \xrightarrow{\zeta|\beta_j} q_j$$

and it holds that $\beta_0 \neq \beta_1$, which implies $\mathrm{delay}(\alpha_0, \alpha_1) \neq \mathrm{delay}(\alpha_0\beta_0, \alpha_1\beta_1)$.

Note that $\mathbb{G}$ is infinitary and hence we obtain that $\{\mathrm{delay}(\alpha_0\beta_0^N, \alpha_1\beta_1^N) \mid N \in \mathbb{N}\}$ is an infinite set. However, the generating set Γ of $\mathbb{G}$ is finite and hence, there exists $N \in \mathbb{N}$ such that $d(\alpha_0\beta_0^N, \alpha_1\beta_1^N) > L$.

Consider the tree $\xi' := \zeta^N[\xi]$ and for every $j \in \{0, 1\}$ the run ρ'_j obtained from ρ_j by pumping the loop parts according to the value N. It surely holds that the weight of ρ'_j satisfies $\mathrm{wt}(\rho'_j) = \alpha_j\beta_j^N$. Therefore we have found two runs of $\mathscr{A}$ on ξ', namely ρ'_0 and ρ'_1, such that $d(\mathrm{wt}(\rho'_0), \mathrm{wt}(\rho'_1)) > L$, which proves the claim. □

Proof of Theorem 16. Assume that $\mathscr{A}$ does not satisfy the twinning property, let $L \in \mathbb{N}$, and let $L' := L(2N+1) + 2NM_{\mathscr{A}} + 2M_{\mathscr{A}}$, where $N := \mathrm{maxrk}(\Sigma)^{\#Q+1}$. Recall from Remark 5 that N is an upper bound for the size of minimal contexts that extend runs into accepting runs.

By Lemma 30 there exist two runs of $\mathscr{A}$ on some tree $\xi \in \mathrm{T}_\Sigma$, $\xrightarrow{\xi|\alpha_0} p_0$ and $\xrightarrow{\xi|\alpha_1} p_1$, for some $p_0, p_1 \in Q$, such that it holds that $d(\alpha_0, \alpha_1) > L'$.

By Remark 5 there exist contexts $\zeta_0, \zeta_1 \in \mathrm{C}_\Sigma$ such that $\mathrm{size}(\zeta_j) \leq N$, pairs $(q_j, \gamma_j) \in \mathrm{final}$, and runs $\xrightarrow{\xi|\alpha_j} p_j \xrightarrow{\zeta_j|\beta_j} q_j$ (for $j \in \{0, 1\}$). Then, $(\zeta_j[\xi], \alpha_j\beta_j\gamma_j) \in [\![\mathscr{A}]\!]$ for every $j \in \{0, 1\}$. Moreover, we apply Lemma 28 to obtain

$$d(\alpha_0\beta_0\gamma_0, \alpha_1\beta_1\gamma_1) > L' - 2NM_{\mathscr{A}} - 2M_{\mathscr{A}} = L(2N+1) \geq L(\mathrm{dist}(\xi_0, \xi_1) + 1).$$

As L was arbitrary, this proves that $[\![\mathscr{A}]\!]$ does not satisfy the Lipschitz property. □

Proof of Lemma 21. We prove the claim by induction on the structure of ξ. Let $\sigma \in \Sigma^{(s)}$, $\xi_1, \ldots, \xi_s \in \mathrm{T}_\Sigma$, such that $\xi = \sigma(\xi_1, \ldots, \xi_s)$ and assume the claim is true for $\xi_1, \ldots, \xi_s$. For every $i \in [s]$, let S_i be the unique state such that $\xrightarrow{\xi_i|\alpha_i} S_i$ and let $(S_1, \ldots, S_s, \hat{\alpha}, S') \in T'_\sigma$. We therefore know that $\alpha = \alpha_1 \cdots \alpha_s\hat{\alpha}$ and

$$S_i = \{(q, \beta) \mid \exists \text{ run} \xrightarrow{\xi_i|\delta} q \text{ of } \mathscr{A} \colon \alpha_i\beta = \delta\}.$$

Hence by definition of T' it holds that

$$\begin{aligned}
S' &= \{(q, \hat{\alpha}^{-1}\beta_1 \cdots \beta_s\gamma) \mid \exists p_1, \ldots, p_s \in Q\colon \\
&\qquad (\forall i \in [s]\colon (p_i, \beta_i) \in S_i) \text{ and } (p_1, \ldots, p_s, \gamma, q) \in T_\sigma\} \\
&= \{(q, \hat{\alpha}^{-1}\beta_1 \cdots \beta_s\gamma) \mid \exists p_1, \ldots, p_s \in Q\colon \\
&\qquad (\forall i \in [s]\colon \exists \text{ run} \xrightarrow{\xi_i|\delta_i} p_i \text{ of } \mathscr{A}\colon \alpha_i\beta_i = \delta_i) \text{ and } (p_1, \ldots, p_s, \gamma, q) \in T_\sigma\} \\
&= \{(q, \hat{\alpha}^{-1}\beta_1 \cdots \beta_s\gamma) \mid \exists \text{ run} \xrightarrow{\xi|\delta} q \text{ of } \mathscr{A}\colon \alpha_1\beta_1 \cdots \alpha_s\beta_s\gamma = \delta\} \\
&= \{(q, \beta) \mid \exists \text{ run} \xrightarrow{\xi|\delta} q \text{ of } \mathscr{A}\colon \alpha_1 \cdots \alpha_s\hat{\alpha}\beta = \delta\} \\
&= \{(q, \beta) \mid \exists \text{ run} \xrightarrow{\xi|\delta} q \text{ of } \mathscr{A}\colon \alpha\beta = \delta\},
\end{aligned}$$

which proves the claim. □

Proof of Lemma 23. We prove the claim by induction on the size of ξ.

If $\mathrm{size}(\xi) \leq \mathrm{maxrk}(\Sigma)^{(\#Q^2+1)}$, the claim holds by definition of $N_{\mathscr{A}}$.

Now assume $\mathrm{size}(\xi) > \mathrm{maxrk}(\Sigma)^{(\#Q^2+1)}$ and let $\rho_0 :\xrightarrow{\xi|\alpha} q$ and $\rho_1 :\xrightarrow{\xi|\beta} p$ be two runs of $\mathscr{A}$ on ξ.

Consider the automaton $\mathscr{A} \times \mathscr{A}$ and the run[9] $\rho_0 \times \rho_1$ of $\mathscr{A} \times \mathscr{A}$ on ξ. Note that the number of states of $\mathscr{A} \times \mathscr{A}$ is $\#Q^2$ and hence by Remark 5, ξ is large enough such that $\rho_0 \times \rho_1$ contains a loop.

In particular, there exist $\zeta, \zeta' \in \mathrm{C}_\Sigma$ and $\xi_0 \in \mathrm{T}_\Sigma$ and states $q', p' \in Q$ such that $\xi = \zeta'[\zeta[\xi_0]]$, $\mathrm{size}(\zeta) > 1$, and we can decompose ρ_0 and ρ_1 such that

$$\rho_0 = \xrightarrow{\xi_0|\alpha_0} q' \xrightarrow{\zeta|\alpha_1} q' \xrightarrow{\zeta'|\alpha_2} q, \qquad \text{and} \qquad \rho_1 = \xrightarrow{\xi_0|\beta_0} p' \xrightarrow{\zeta|\beta_1} p' \xrightarrow{\zeta'|\beta_2} p.$$

By the twinning property we obtain

$$\begin{aligned}
d(\alpha, \beta) &= |\mathrm{delay}(\alpha_0\alpha_1\alpha_2, \beta_0\beta_1\beta_2)|_\Gamma = |\alpha_2^{-1}\mathrm{delay}(\alpha_0\alpha_1, \beta_0\beta_1)\beta_2|_\Gamma \\
&= |\alpha_2^{-1}\mathrm{delay}(\alpha_0, \beta_0)\beta_2|_\Gamma = d(\alpha_0\alpha_2, \beta_0\beta_2),
\end{aligned}$$

hence we can apply the induction hypothesis to the tree $\zeta'[\xi_0]$, which proves the claim. □

[9] Defined in the obvious way as the position-wise direct product of ρ_0 and ρ_1.