Deciding the Twins Property for Weighted Tree Automata over Extremal Semifields

Matthias Büchse and **Anja Fischer**
Department of Computer Science
Technische Universität Dresden
01062 Dresden, Germany
Matthias.Buechse@tu-dresden.de

Abstract

It has remained an open question whether the twins property for weighted tree automata is decidable. This property is crucial for determinizing such an automaton, and it has been argued that determinization improves the output of parsers and translation systems. We show that the twins property for weighted tree automata over extremal semifields is decidable.

1 Introduction

In natural-language processing (NLP), language and translation are often modeled using some kind of grammar, automaton or transducer, such as a probabilistic context-free grammar, a synchronous context-free grammar, a weighted tree automaton, or a tree transducer, among others (May and Knight, 2006; Petrov et al., 2006; Chiang, 2007; Graehl, Knight and May, 2008; Zhang et al., 2008; Pauls and Klein, 2009). In statistical NLP, the structure of the grammar is extracted heuristically from a large corpus of example sentences or sentence pairs, and the rule weights are estimated using methods from statistics or machine learning.

In general, a grammar such as those named above will be *ambiguous*, i.e., offering several ways of deriving the same object (sentence or sentence pair). While the derivation of an object is crucial to the intrinsics of a system, it is neither relevant to the user nor observed in the corpus. Hence, we speak of *spurious ambiguity* (Li, Eisner and Khudanpur, 2009).

As a consequence, the true importance of an object can only be assessed by aggregating all its derivations. Unfortunately, this proves computationally intractable in almost all cases: for instance, finding the best string of a probabilistic regular grammar is NP hard (Sima'an, 1996; Casacuberta and de la Higuera, 2000). Finding the best derivation, on the other hand, is possible in polynomial time (Eppstein, 1998; Huang and Chiang, 2005), and thus, most NLP systems approximate the importance of an object by its best derivation (Li, Eisner and Khudanpur, 2009).

There is, however, a line of research that deals with the costly aggregating approach, and it is closely related to determinization techniques from automata theory.

For instance, May and Knight (2006) argue that the output of a parser or syntax-based translation system can be represented by a weighted tree automaton (wta), which assigns a weight to each parse tree. Under some circumstances, the wta can be determinized, yielding an equivalent, but unambiguous wta, which offers at most one derivation for each object. Then the weight of an object is equal to the weight of its derivation, and the aforementioned polynomial-time algorithms deliver exact results.

The caveat of the determinization approach is that deterministic weighted automata are strictly less powerful than their general counterparts, i.e., not every automaton can be determinized. Büchse, May and Vogler (2010) give a review of known sufficient conditions under which determinization is possible. One of these conditions requires that (i) the weights are calculated in an extremal semiring, (ii) there is a maximal factorization, and (iii) the wta has the twins property.[1]

[1] Items (i) and (iii) guarantee that the wta only computes weight vectors that are scalar multiples of a finite number

Proc. EACL 2012 Workshop on Applications of Tree Automata Techniques in Natural Language Processing, pages 11–20, Avignon, France, April 24 2012.

Regarding (i), we note that in an extremal semiring the weight of a parse tree is equal to the weight of its best derivation. It follows that, while the determinized wta will have at most one derivation per parse tree, its weight will be the weight of the best derivation of the original wta. The benefit of determinization reduces to removing superfluous derivations from the list of best derivations.

Regarding (ii), the factorization is used in the determinization construction to distribute the weight computation in the determinized automaton between its transition weights and its state behavior. A maximal factorization exists for every zero-sum free semifield.

Regarding (iii), the question whether the twins property is decidable has remained open for a long time, until Kirsten (2012)[2] gave an affirmative answer for a particular case: weighted string automata over the tropical semiring. He also showed that the decision problem is PSPACE-complete.

In this paper, we close one remaining gap by adapting and generalizing Kirsten's proof: we show that the twins property is decidable for wta over extremal semifields (Theorem 3.1). We proceed by recalling the concepts related to determinizing wta, such as ranked trees, semirings, factorizations, wta themselves, and the twins property (Sec. 2). Then we show our main theorem, including two decision algorithms (Sec. 3). Finally, we conclude the paper with a discussion and some open questions (Sec. 4).

2 Preliminaries

2.1 Ranked Trees

A *ranked alphabet* is a tuple (Σ, rk) where Σ is an alphabet, i.e., a finite set, and $rk\colon \Sigma \to \mathbb{N}$ assigns an *arity* to every symbol $\sigma \in \Sigma$. Throughout this paper we will identify (Σ, rk) with Σ. For every $k \in \mathbb{N}$ the set $\Sigma^{(k)} = \{\sigma \in \Sigma \mid rk(\sigma) = k\}$ contains all symbols of arity k.

Let H be a set and Σ a ranked alphabet. The set $T_\Sigma(H)$ of *trees over* Σ *indexed by* H is defined inductively as the smallest set T such that: (i) $H \subseteq T$ and (ii) $\sigma(\xi_1, \ldots, \xi_k) \in T$ for every $k \in \mathbb{N}$, $\sigma \in \Sigma^{(k)}$, and $\xi_1, \ldots, \xi_k \in T$. We write T_Σ instead of $T_\Sigma(\emptyset)$.

For every $\xi \in T_\Sigma(H)$, we define the set $\mathrm{pos}(\xi) \subseteq \mathbb{N}^*$ of *positions of* ξ by

(i) if $\xi \in H$, then $\mathrm{pos}(\xi) = \{\varepsilon\}$;

(ii) if $\xi = \sigma(\xi_1, \ldots, \xi_k)$, then $\mathrm{pos}(\xi) = \{\varepsilon\} \cup \{i \cdot w \mid i \in \{1, \ldots, k\}, w \in \mathrm{pos}(\xi_i)\}$.

The mapping $\mathrm{ht}\colon T_\Sigma(H) \to \mathbb{N}$ maps each tree ξ to its *height*, i.e., the length of a longest position of ξ. We denote the *label of* ξ *at position* w by $\xi(w)$, the *subtree of* ξ *rooted at* w by $\xi|_w$, and the tree obtained by *replacing the subtree of* ξ *rooted at position* w *with* ξ', $\xi' \in T_\Sigma(H)$, by $\xi[\xi']_w$.

A Σ-*context* is a tree in $T_\Sigma(\{z\})$ that contains exactly one occurrence of the special symbol z. The set of all Σ-contexts is denoted by C_Σ. Let $\xi \in T_\Sigma \cup C_\Sigma$ and $\zeta \in C_\Sigma$. Then the *concatenation of* ξ *and* ζ, denoted by $\xi \cdot \zeta$, is obtained from ζ by replacing the leaf z by ξ. If $\xi \in T_\Sigma$, then so is $\xi \cdot \zeta$, and likewise for $\xi \in C_\Sigma$.

2.2 Semirings

A *semiring* (Hebisch and Weinert, 1998; Golan, 1999) is a quintuple $\mathcal{S} = (S, +, \cdot, 0, 1)$ where S is a set, $+$ and $\cdot$ are binary, associative operations on S, called *addition* and *multiplication*, respectively, $+$ is commutative, $\cdot$ distributes over $+$ from both sides, 0 and 1 are elements of S, 0 is neutral with respect to $+$, 1 is neutral with respect to $\cdot$, and 0 is absorbing with respect to $\cdot$ (i.e., $s \cdot 0 = 0 = 0 \cdot s$).

Let $\mathcal{S} = (S, +, \cdot, 0, 1)$ be a semiring. In notation, we will identify $\mathcal{S}$ with S. We call S *commutative* if the multiplication is commutative; a *semifield* if it is commutative and for every $a \in S \setminus \{0\}$ there is an $a^{-1} \in S$ such that $a \cdot a^{-1} = 1$; *zero-sum free* if $a + b = 0$ implies $a = b = 0$; *zero-divisor free* if $a \cdot b = 0$ implies $a = 0$ or $b = 0$; and *extremal* (Mahr, 1984) if $a + b \in \{a, b\}$. We note that every extremal semiring is also zero-sum free and every semifield is zero-divisor free.

Example 2.1 We present four examples of semirings. The *Boolean semiring* $\mathbb{B} = (\{0,1\}, \vee, \wedge, 0, 1)$, with disjunction and conjunction, is an extremal semifield. The *formal-language semiring* $(\mathcal{P}(\Sigma^*), \cup, \cdot, \emptyset, \{\varepsilon\})$ over an alphabet Σ, with union and language concatenation, is neither commutative nor extremal, but zero-divisor free and zero-sum free. The *tropical semiring* $(\mathbb{R} \cup \{\infty\}, \min, +, \infty, 0)$, with minimum and conventional addition, is

of vectors corresponding to a set of height-bounded trees, while Item (ii) ensures that the latter vectors suffice as the states of the constructed deterministic wta; cf. (Büchse, May and Vogler, 2010, Lm. 5.9 and Lm. 5.8, respectively).

[2]A manuscript with the same content has been available on Daniel Kirsten's website for a year from Sept. 2010 on.

an extremal semifield. The *Viterbi semiring* $([0,1], \max, \cdot, 0, 1)$ is a commutative, extremal, zero-divisor-free semiring, but not a semifield. □

Let Q be a set. The set S^Q contains all mappings $u\colon Q \to S$, or, equivalently, all Q-vectors over S. Instead of $u(q)$ we also write u_q to denote the q-component of a vector $u \in S^Q$. The Q-vector mapping every q to 0 is denoted by $\tilde{0}$. For every $q \in Q$ we define $e_q \in S^Q$ such that $(e_q)_q = 1$, and $(e_q)_p = 0$ for every $p \neq q$.

2.3 Factorizations

We use the notion of a factorization as defined in (Kirsten and Mäurer, 2005).

Let Q be a nonempty finite set. A pair (f, g) is called a *factorization of dimension* Q if $f\colon S^Q \setminus \{\tilde{0}\} \to S^Q$, $g\colon S^Q \setminus \{\tilde{0}\} \to S$, and $u = g(u) \cdot f(u)$ for every $u \in S^Q \setminus \{\tilde{0}\}$. A factorization (f, g) is called *maximal* if for every $u \in S^Q$ and $a \in S$, we have that $a \cdot u \neq \tilde{0}$ implies $f(a \cdot u) = f(u)$.

Example 2.2 Let Q be a nonempty finite set. We show three factorizations of dimension Q.

If S is an arbitrary semiring, $g(u) = 1$ and $f(u) = u$ constitute the *trivial factorization*. It is not maximal in general.

If S is a zero-sum free semifield, such as the tropical semiring or the semifield of non-negative reals, then $g(u) = \sum_{q \in Q} u_q$ and $f(u) = \frac{1}{g(u)} \cdot u$ constitute a factorization (Büchse, May and Vogler, 2010, Lemma 4.2). It is maximal: $f(a \cdot u) = \frac{1}{g(a \cdot u)} \cdot (a \cdot u) = \frac{1}{a \cdot g(u)} \cdot a \cdot u = f(u)$.

As shown in (Büchse, May and Vogler, 2010, Lemma 4.4) a maximal factorization only exists if S is zero-divisor free or $|Q| = 1$.

2.4 Weighted Tree Automata

A weighted tree automaton (Ésik and Kuich, 2003) is a finite-state machine that represents a *weighted tree language*, i.e., a mapping $\varphi\colon T_\Sigma \to S$. It assigns a weight to every tree based on weighted transitions.

Formally, a *weighted tree automaton (wta)* is a tuple $\mathcal{A} = (Q, \Sigma, S, \delta, \nu)$ such that Q is a nonempty finite set (of *states*), Σ is a ranked alphabet, S is a semiring, δ is the *transition mapping*, mapping *transitions* $(q_1 \cdots q_k, \sigma, q)$ into S where $q_1, \ldots, q_k, q \in Q$ and $\sigma \in \Sigma^{(k)}$, and $\nu \in S^Q$ maps every state to its *root weight*.

A wta $\mathcal{A}$ is *bottom-up deterministic* if for every $(q_1 \cdots q_k, \sigma)$, there is at most one q such that $\delta(q_1 \cdots q_k, \sigma, q) \neq 0$.

Example 2.3 Let $\mathcal{A} = (Q, \Sigma, S, \delta, \nu)$ be the wta where $\Sigma = \{\alpha^{(0)}, \gamma^{(1)}, \sigma^{(2)}\}$, S is the arctic semiring $(\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$, δ is given by the directed functional hypergraph in Fig. 1, and $\nu = (0, -\infty)$. Each node in the hypergraph (drawn as circle) corresponds to a state, and each hyperedge (drawn as box with arbitrarily many ingoing arcs and exactly one outgoing arc) represents a weighted transition. Ingoing arcs of a hyperedge are meant to be read counter-clockwise, starting from the outgoing arc. The final weight 0 of q_1 is indicated by an additional arc. Transitions not shown have the weight $-\infty$. □

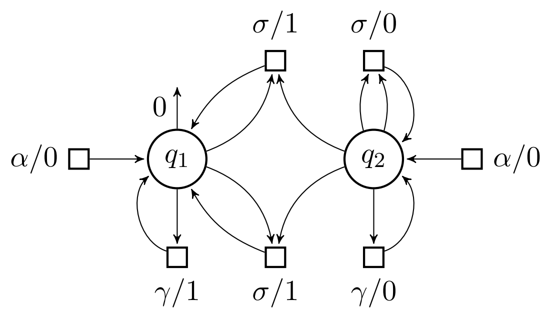

Figure 1: Hypergraph representation of wta $\mathcal{A}$.

Typically, wta are given initial-algebra semantics (Goguen et al., 1977). In this paper, we use the equivalent run semantics (Fülöp and Vogler, 2009, Sec. 3.2) as it constitutes the basis for our proofs. In this setting, every node of a given tree is decorated with a state; this decoration is called a *run*. The label of a node, its state, and the states of its successors comprise a transition. The weight of a run is given by the product of the weights of all these transitions (under δ), calculated in the semiring S. Roughly speaking, the weight of a tree is then the sum of the weights of all runs on that tree, again calculated in S.

Now we formalize the notions of a run and its weight. For our proofs, we will need runs and their weights to be as easily composable and decomposable as trees and contexts. Therefore, we will consider trees indexed by semiring elements and even Q-vectors over S. Let H be a set, $\xi \in T_\Sigma(H)$, and $q \in Q$. The *set* $R^q_{\mathcal{A}}(\xi)$ *of all runs on* ξ *that end in state* q *at the root of* ξ is

$$R^q_{\mathcal{A}}(\xi) = \{(\xi, \kappa) \mid \kappa\colon \mathrm{pos}(\xi) \to Q, \kappa(\varepsilon) = q\} .$$

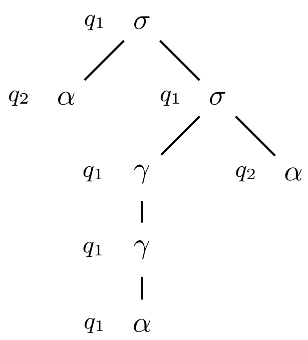

Figure 2: A tree together with a run.

We will denote the pair (ξ, κ) just by κ and indicate ξ by stating $\kappa \in R^q_{\mathcal{A}}(\xi)$. We will also omit the subscript $\mathcal{A}$. We set $R(\xi) = \bigcup_{q \in Q} R^q(\xi)$.

Let $w \in \mathrm{pos}(\xi)$ and $\kappa \in R^q(\xi)$. The following notions are defined in the obvious way: (i) $\kappa|_w \in R^{\kappa(w)}(\xi|_w)$, (ii) $\kappa[\kappa']_w \in R^q(\xi[\xi']_w)$ for every $\xi' \in T_\Sigma(H)$ and $\kappa' \in R^{\kappa(w)}(\xi')$, and (iii) $\kappa \cdot \kappa' \in R^{q'}(\xi \cdot \zeta)$ for every $q' \in Q$, $\zeta \in C_\Sigma$, and $\kappa' \in R^{q'}(\zeta)$ that maps the z-labelled position to q. We will abuse the above notation in two ways: (i) we write $\kappa[z]_w$ to denote $\kappa[\kappa']_w$ where κ' is the only element of $R^{\kappa(w)}(z)$, and (ii) for every $s \in S$, we write $s \cdot \kappa$ to denote the run on $s \cdot \zeta$ which coincides with κ.

Let $\xi \in T_\Sigma(S \cup S^Q)$ and $\kappa \in R(\xi)$. We define the *weight* $\langle\kappa\rangle_{\mathcal{A}} \in S$ *of* κ as follows (omitting the subscript $\mathcal{A}$): if $\xi \in S$, then $\langle\kappa\rangle = \xi$; if $\xi \in S^Q$, then $\langle\kappa\rangle = \xi_{\kappa(\varepsilon)}$; if $\xi = \sigma(\xi_1, \dots, \xi_k)$, then $\langle\kappa\rangle = \langle\kappa|_1\rangle \cdot \ldots \cdot \langle\kappa|_k\rangle \cdot \delta(\kappa(1) \cdots \kappa(k), \sigma, \kappa(\varepsilon))$.

We define the mapping $[\![.]\!]_{\mathcal{A}} \colon T_\Sigma(S^Q) \to S^Q$ such that $[\![\xi]\!]_{\mathcal{A}}(q) = \sum_{\kappa \in R^q(\xi)} \langle\kappa\rangle$. Again, we will often omit the subscript $\mathcal{A}$. If we have a factorization (f, g), we will shorten $f([\![\xi]\!])$ to $f[\![\xi]\!]$. We will often use relationships such as $\langle\kappa \cdot \kappa'\rangle = \langle\langle\kappa\rangle \cdot \kappa'\rangle$ and $[\![\xi \cdot \zeta]\!] = [\![[\![\xi]\!] \cdot \zeta]\!]$.

The weighted tree language *run-recognized by* $\mathcal{A}$ is the mapping $\varphi_{\mathcal{A}} \colon T_\Sigma \to S$ such that for every $\xi \in T_\Sigma$ we have $\varphi_{\mathcal{A}}(\xi) = \sum_{q \in Q} [\![\xi]\!]_q \cdot \nu_q$.

Example 2.4 (Ex. 2.3 contd.) Figure 2 shows a tree together with a run κ. We compute $\langle\kappa\rangle$ (recall that we use the arctic semiring):

$$\begin{aligned}\langle\kappa\rangle &= \langle\kappa|_1\rangle + \langle\kappa|_2\rangle + \delta(q_2 q_1, \sigma, q_1)\\ &= \delta(\varepsilon, \alpha, q_2) + \delta(\varepsilon, \alpha, q_1) + \delta(q_1, \gamma, q_1)\\ &\quad + \delta(q_1, \gamma, q_1) + \delta(\varepsilon, \alpha, q_2)\\ &\quad + \delta(q_1 q_2, \sigma, q_1) + \delta(q_2 q_1, \sigma, q_1)\\ &= 0 + 0 + 1 + 1 + 0 + 1 + 1 = 4 .\end{aligned}$$

It can be shown that $[\![\xi]\!]_{q_1} = \mathrm{ht}(\xi)$ and $[\![\xi]\!]_{q_2} = 0$, and thus, that $\varphi_{\mathcal{A}} = \mathrm{ht}$. □

For every $\xi \in T_\Sigma(S^Q)$ and $\kappa \in R(\xi)$ we call κ *victorious (on* ξ*)* if $\langle\kappa\rangle = [\![\xi]\!]_{\kappa(\varepsilon)}$. The following observations are based on (Büchse, May and Vogler, 2010, Obs. 5.11 and 5.12).

Observation 2.5 *Let* S *be an extremal semiring. For every* $\xi \in T_\Sigma(S^Q)$ *and* $q \in Q$ *there is a* $\kappa \in R^q(\xi)$ *such that* κ *is victorious.*

Observation 2.6 *Let* $\xi \in T_\Sigma(S^Q)$*,* $w \in \mathrm{pos}(\xi)$*, and* $\kappa \in R(\xi)$ *victorious. Then we obtain* $\langle\kappa\rangle = [\![(\langle\kappa|_w\rangle \cdot e_{\kappa(w)}) \cdot \xi[z]_w]\!]_{\kappa(\varepsilon)}$.

PROOF.

$$\begin{aligned}&[\![(\langle\kappa|_w\rangle \cdot e_{\kappa(w)}) \cdot \xi[z]_w]\!]_{\kappa(\varepsilon)}\\ &= \textstyle\sum_{\kappa' \in R^{\kappa(\varepsilon)}\left((\langle\kappa|_w\rangle \cdot e_{\kappa(w)}) \cdot \xi[z]_w\right)} \langle\kappa'\rangle\\ &= \textstyle\sum_{\kappa' \in R^{\kappa(\varepsilon)}(\xi[z]_w), \kappa'(w) = \kappa(w)} \langle\langle\kappa|_w\rangle \cdot \kappa'\rangle\\ &= \textstyle\sum_{\kappa' \in R^{\kappa(\varepsilon)}(\xi[z]_w), \kappa'(w) = \kappa(w)} \langle\kappa|_w \cdot \kappa'\rangle\\ &= \langle\kappa\rangle .\end{aligned}$$

For the last equation, we note that the summands on the left-hand side form a subset of $\{\langle\nu\rangle \mid \nu \in R^{\kappa(\varepsilon)}(\xi)\}$, which contains $\langle\kappa\rangle$. Since S is extremal and $\langle\kappa\rangle = [\![\xi]\!]_{\kappa(\varepsilon)}$, the equation holds. ■

2.5 Twins Property

We define two binary relations SIBLINGS$(\mathcal{A})$ and TWINS$(\mathcal{A})$ over Q as follows. Let $p, q \in Q$. Then

- $(p, q) \in$ SIBLINGS$(\mathcal{A})$ iff there is a tree $\xi \in T_\Sigma$ such that $[\![\xi]\!]_p \neq 0$ and $[\![\xi]\!]_q \neq 0$.
- $(p, q) \in$ TWINS$(\mathcal{A})$ iff for every context $\zeta \in C_\Sigma$ we have that $[\![e_p \cdot \zeta]\!]_p \neq 0$ and $[\![e_q \cdot \zeta]\!]_q \neq 0$ implies $[\![e_p \cdot \zeta]\!]_p = [\![e_q \cdot \zeta]\!]_q$.

The wta $\mathcal{A}$ is said to have the *twins property* if SIBLINGS$(\mathcal{A}) \subseteq$ TWINS$(\mathcal{A})$.

Example 2.7 We cover two examples.

First, consider the wta from Ex. 2.3. Its two states are siblings as witnessed by the tree $\xi = \alpha$. However, they are not twins, as witnessed by the context $\zeta = \gamma(z)$: $[\![e_{q_1} \cdot \gamma(z)]\!]_{q_1} = 1$, whereas $[\![e_{q_2} \cdot \gamma(z)]\!]_{q_2} = 0$.

Second, consider the wta over the Viterbi semiring shown Fig. 3. Its two states are siblings as witnessed by the tree $\xi = \alpha$. Furthermore, they are twins because their transitions are symmetric. Hence, this wta has the twins property. □

The following observation shows that we can enumerate SIBLINGS$(\mathcal{A})$ in finite time.

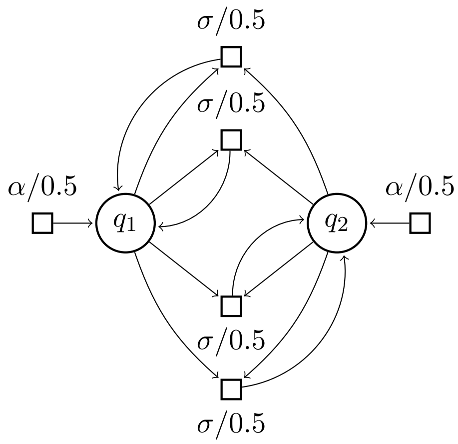

Figure 3: Siblings and twins.

Observation 2.8 *If S is zero-sum free, we have* $\text{SIBLINGS}(\mathcal{A}) = \text{SIB}(\mathcal{A})$ *where* $\text{SIB}(\mathcal{A})$ *is defined like* $\text{SIBLINGS}(\mathcal{A})$*, with the additional condition that* $\text{ht}(\xi) < |Q|^2$.

PROOF. The direction $\supseteq$ is trivial. We show $\subseteq$ by contradiction. Let $p, q \in Q$ and $\xi \in T_\Sigma$ such that (i) $[\![\xi]\!]_p \neq 0$ and $[\![\xi]\!]_q \neq 0$, and (ii) $(p, q) \notin \text{SIB}(\mathcal{A})$. We assume that ξ is smallest, and we show that we find a smaller counterexample.

By (ii), we have (iii) $\text{ht}(\xi) \geq |Q|^2$. By (i), there are $\kappa_p \in R^p(\xi)$ and $\kappa_q \in R^q(\xi)$ such that (iv) $\langle \kappa_p \rangle \neq 0$ and $\langle \kappa_q \rangle \neq 0$.

By (iii), there are positions w_1, w_2 such that w_1 is above w_2, $\kappa_p(w_1) = \kappa_p(w_2)$, and $\kappa_q(w_1) = \kappa_q(w_2)$. Cutting out the slice between w_1 and w_2, we construct the tree $\xi' = \xi[\xi|_{w_2}]_{w_1}$. Moreover, we construct the runs κ'_p and κ'_q accordingly, i.e., $\kappa'_x = \kappa_x[\kappa_x|_{w_2}]_{w_1}$.

We have that $\langle \kappa'_p \rangle \neq 0$, $\langle \kappa'_q \rangle \neq 0$, because otherwise (iv) would be violated. Since S is zero-sum free, we obtain $[\![\xi']\!]_p \neq 0$, $[\![\xi']\!]_q \neq 0$. ∎

3 Decidability of the Twins Property

This section contains our main theorem:

Theorem 3.1 *The twins property of wta over extremal semifields is decidable.*

The following subsections provide the infrastructure and lemmata needed for the proof of the theorem. Henceforth, we assume that S is an extremal semifield. As noted in Ex. 2.2, there is a maximal factorization (f, g).

3.1 Rephrasing the Twins Relation

In the definition of $\text{TWINS}(\mathcal{A})$, we deal with two vectors $[\![e_p \cdot \zeta]\!]$ and $[\![e_q \cdot \zeta]\!]$ for each $\zeta \in C_\Sigma$. In the following we concatenate these vectors into one, which enables us to use a factorization. To this end, we construct a wta $\mathcal{A} \cup \bar{\mathcal{A}}$ that runs two instances of $\mathcal{A}$ in parallel, as shown in Fig. 4.

Let $\mathcal{A} = (Q, \Sigma, S, \delta, \nu)$ a wta and $\bar{\mathcal{A}} = (\bar{Q}, \Sigma, S, \bar{\delta}, \bar{\nu})$ be the wta obtained from $\mathcal{A}$ by renaming states via $q \mapsto \bar{q}$. We construct the wta $\mathcal{A} \cup \bar{\mathcal{A}} = (Q \cup \bar{Q}, \Sigma, S, \delta', \nu')$ where δ' coincides with δ and $\bar{\delta}$ on the transitions of $\mathcal{A}$ and $\bar{\mathcal{A}}$, respectively; it maps all other transitions to 0; and ν' coincides with ν and $\bar{\nu}$ on Q and $\bar{Q}$, respectively.

For every $p, q \in Q$ we define the set $T_{p,q} \subseteq S^{Q \cup \bar{Q}}$ by $T_{p,q} = \{[\![(e_p + e_{\bar{q}}) \cdot \zeta]\!]_{\mathcal{A} \cup \bar{\mathcal{A}}} \mid \zeta \in C_\Sigma\}$; note that $e_p, e_{\bar{q}} \in S^{Q \cup \bar{Q}}$. With this definition, we observe the following trivial equivalence.

Observation 3.2 *Let* $p, q \in Q$*. Then* $(p, q) \in \text{TWINS}(\mathcal{A})$ *iff for every* $u \in T_{p,q}$ *we have that*

$$u_p \neq 0 \textit{ and } u_{\bar{q}} \neq 0 \textit{ implies } u_p = u_{\bar{q}}.$$

For every pair $(p, q) \in \text{SIBLINGS}(\mathcal{A})$, a vector $u \in S^{Q \cup \bar{Q}}$ is called a *critical vector (for* (p, q)*)* if it does not fulfill the centered implication of Obs. 3.2. Any critical vector in $T_{p,q}$ thereby witnesses $(p, q) \notin \text{TWINS}(\mathcal{A})$. Consequently, $\mathcal{A}$ has the twins property iff $T_{p,q}$ contains no critical vector for any $(p, q) \in \text{SIBLINGS}(\mathcal{A})$. Deciding the twins property thus amounts to searching for a critical vector.

3.2 Compressing the Search Space

In this subsection we approach the decidability of the twins property by compressing the search space for critical vectors. First we show that the vectors in $T_{p,q}$ are scalar multiples of a finite number of vectors.

Lemma 3.3 *Let S be a commutative, extremal semiring. Assume that $\mathcal{A}$ has the twins property. Then there is a finite set* $S' \subseteq S^{Q \cup \bar{Q}}$ *such that for every* $(p, q) \in \text{SIBLINGS}(\mathcal{A})$ *we have*

$$T_{p,q} \subseteq S \cdot S'.$$

PROOF. We construct sets $S', S'' \subseteq S^{Q \cup \bar{Q}}$ and show the following inclusions:

$$T_{p,q} \subseteq S \cdot S'' \subseteq S \cdot S'. \qquad (*)$$

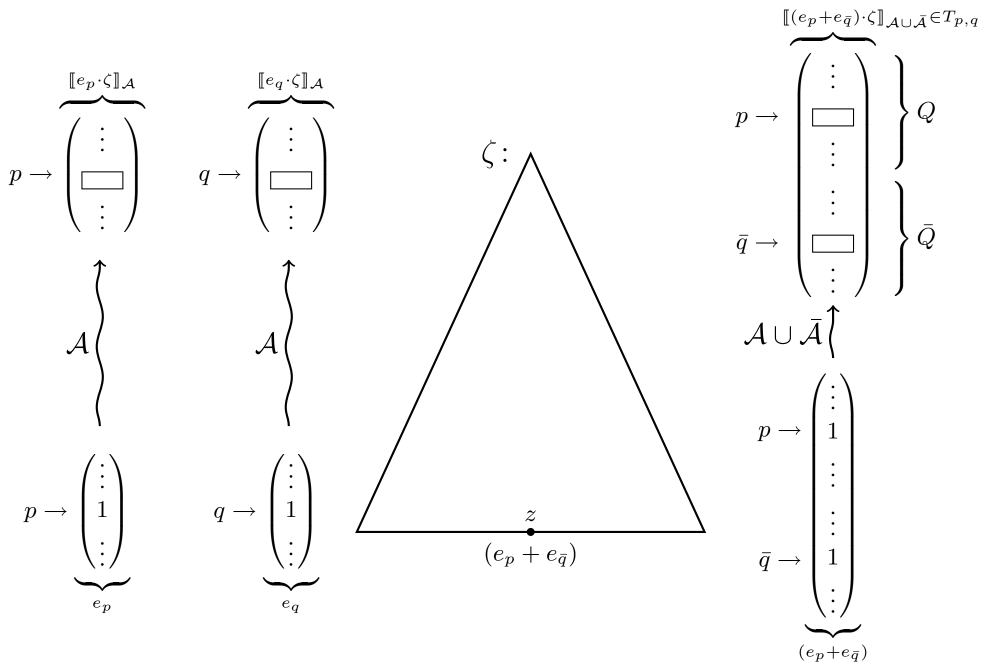

Figure 4: Moving from parallel execution of $\mathcal{A}$ (left-hand side) to the union wta $\mathcal{A} \cup \bar{\mathcal{A}}$ (right-hand side).

To this end, we consider each entry in each vector to be induced by an according (victorious) run. In this spirit we define for every $p, q \in Q$ and $\zeta \in C_\Sigma$ the set $C_{p,q}(\zeta) \subseteq R((e_p + e_{\bar{q}}) \cdot \zeta)^{Q \cup \bar{Q}}$ of vectors of runs of $\mathcal{A} \cup \bar{\mathcal{A}}$ as follows: $\kappa \in C_{p,q}(\zeta)$ iff (i) $\kappa_r \in R^r((e_p + e_{\bar{q}}) \cdot \zeta)$ for every $r \in Q \cup \bar{Q}$ and (ii) for every pair $w_1, w_2 \in \mathrm{pos}(\zeta)$ with w_1 above w_2 and $\kappa_r(w_1) = \kappa_r(w_2)$ we have that $\kappa_r|_{w_1}$ is victorious on $((e_p + e_{\bar{q}}) \cdot \zeta)|_{w_1}$. We map each vector of runs to the corresponding weight vector as follows. For every $Q' \subseteq Q \cup \bar{Q}$ let $\gamma_{Q'} : R((e_p + e_{\bar{q}}) \cdot \zeta)^{Q \cup \bar{Q}} \to S^{Q \cup \bar{Q}}$ be the mapping such that for every κ and $q' \in Q \cup \bar{Q}$:

$$\gamma_{Q'}(\kappa)_{q'} = \begin{cases} \langle \kappa_{q'} \rangle & \text{if } q' \in Q' \\ 0 & \text{otherwise.} \end{cases}$$

We set $S'' = \{\gamma_{Q'}(\kappa) \mid (p,q) \in \textsc{Siblings}(\mathcal{A}), \zeta \in C_\Sigma, \kappa \in C_{p,q}(\zeta), Q' \subseteq Q \cup \bar{Q}\}$. The set S' is defined in the same way, with the additional condition that $\mathrm{ht}(\zeta) < 2|Q|^{2|Q|}$.

The first inclusion of $(*)$ can be proved in the same way as (Büchse, May and Vogler, 2010, Lemma 5.14). Here we show the second inclusion by contradiction. To this end, let $s \in S$, $(p,q) \in \textsc{Siblings}(\mathcal{A})$, $\zeta \in C_\Sigma$, $\kappa \in C_{p,q}(\zeta)$, and $Q' \subseteq Q \cup \bar{Q}$ such that $s \cdot \gamma_{Q'}(\kappa) \notin S \cdot S'$, and thus $\mathrm{ht}(\zeta) \geq 2|Q|^{2|Q|}$. We can assume that $\langle \kappa_r \rangle \neq 0$ for every $r \in Q'$ because otherwise we could adjust Q' without harm. Finally, we assume that ζ is smallest.

We will construct a new context ζ' and a corresponding vector $\kappa' \in C_{p,q}(\zeta')$ such that ζ' is smaller than ζ and $s \cdot \gamma_{Q'}(\kappa) = s \cdot s' \cdot \gamma_{Q'}(\kappa')$ for some $s' \in S$. Then, if the right-hand side is in $S \cdot S'$, so is the left-hand side. By contraposition, this shows that ζ was not a smallest counterexample, yielding the contradiction.

First, let w be the position in ζ labelled z. We show that we are able to find a pair (w_1, w_2) of positions such that w_1 is above w_2, $\kappa_r(w_1) = \kappa_r(w_2)$ for every r, and either both or none of w_1 and w_2 are above w. To this end, we distinguish two cases (cf. Fig. 5).

(a) If $|w| \leq |Q|^{2|Q|}$, then the length of the common prefix of w and any path of length at least $2|Q|^{2|Q|}$ can be at most $|Q|^{2|Q|}$. Hence, on such a path remain at least $|Q|^{2|Q|} + 1$ positions that are not above w. By the pidgeonhole principle, we find said pair (w_1, w_2).

(b) If $|w| > |Q|^{2|Q|}$, then we find the pair immediately on the path to the position labelled z.

Second, we pick a pair (w_1, w_2) such that the position w_1 has minimal length. Cutting out the slice between the positions w_1 and w_2 yields the smaller context $\zeta' = \zeta[\zeta|_{w_2}]_{w_1}$. We construct κ' accordingly, i.e., $\kappa'_r = \kappa_r[\kappa_r|_{w_2}]_{w_1}$ for every $r \in Q \cup \bar{Q}$. We have that $\kappa' \in C_{p,q}(\zeta')$; for this we

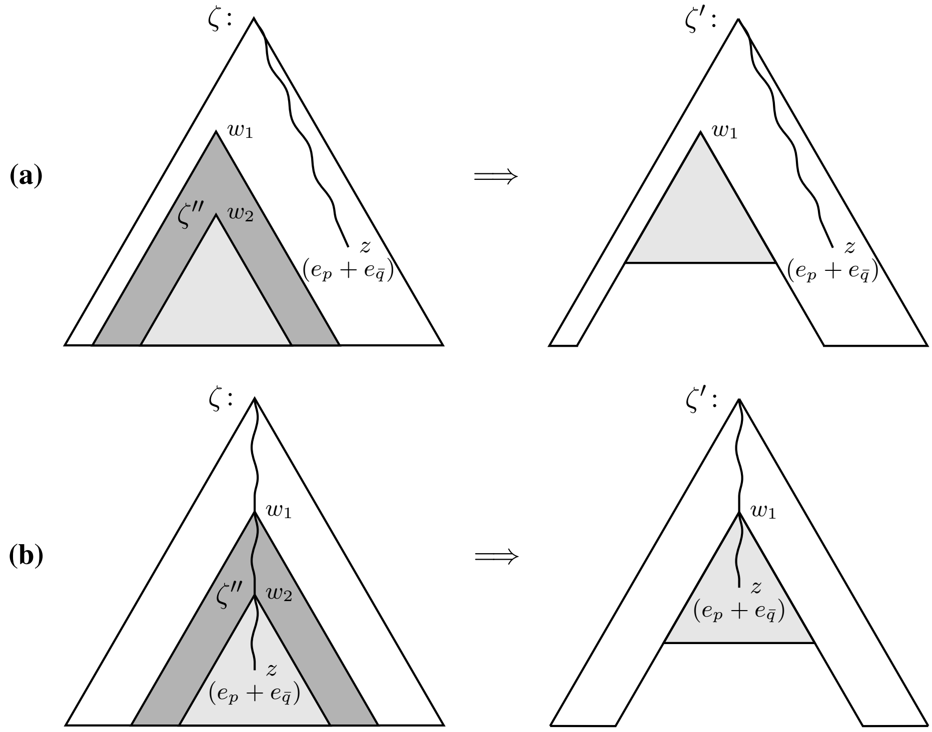

Figure 5: Two cases for the construction of $\zeta' = \zeta[\zeta|_{w_2}]_{w_1}$.

need that we chose w_1 with minimal length.

Third, we use the twins property to show that there is an $s' \in S$ such that $s \cdot \gamma_{Q'}(\kappa) = s \cdot s' \cdot \gamma_{Q'}(\kappa')$. If $Q' = \emptyset$, we set $s' = 0$, and the proof is done. Otherwise we choose some $r' \in Q'$ and set $s' = [\![e_{\kappa_{r'}(w_2)} \cdot \zeta'']\!]_{\kappa_{r'}(w_1)}$ where $\zeta'' = \zeta[z]_{w_2}|_{w_1}$ is the slice we have cut out. We prove that $\gamma_{Q'}(\kappa) = s' \cdot \gamma_{Q'}(\kappa')$. To this end, let $r \in Q'$, $p' = \kappa_r(w_1) = \kappa_r(w_2)$, and $q' = \kappa_{r'}(w_1) = \kappa_{r'}(w_2)$. Then

$$\begin{aligned}
\gamma_{Q'}(\kappa)_r &= \langle \kappa_r \rangle = \langle \langle \kappa_r|_{w_1} \rangle \cdot \kappa_r[z]_{w_1} \rangle \\
&= \langle [\![(\langle \kappa_r|_{w_2} \rangle \cdot e_{p'}) \cdot \zeta'']\!]_{p'} \cdot \kappa_r[z]_{w_1} \rangle \quad \text{(Obs. 2.6)} \\
&= \langle \kappa_r|_{w_2} \rangle \cdot [\![e_{p'} \cdot \zeta'']\!]_{p'} \cdot \langle 1 \cdot \kappa_r[z]_{w_1} \rangle \\
&\qquad \text{(commutativity)} \\
&= \langle \kappa_r|_{w_2} \rangle \cdot [\![e_{q'} \cdot \zeta'']\!]_{q'} \cdot \langle 1 \cdot \kappa_r[z]_{w_1} \rangle \quad (\dagger) \\
&= s' \cdot \langle \langle \kappa_r|_{w_2} \rangle \cdot \kappa_r[z]_{w_1} \rangle \quad \text{(commutativity)} \\
&= s' \cdot \langle \kappa'_r \rangle = s' \cdot \gamma_{Q'}(\kappa')_r \ .
\end{aligned}$$

At (†) we have used the twins property. We show that this is justified. First, we show that $(p', q') \in \textsc{Siblings}(\mathcal{A} \cup \bar{\mathcal{A}})$. To this end, we distinguish two cases.

If z occurs in $\zeta|_{w_2}$: by $(p, q) \in \textsc{Siblings}(\mathcal{A})$ we obtain a tree ξ such that $[\![\xi]\!]_p \neq 0$ and $[\![\xi]\!]_q \neq 0$. By our assumption we have $\langle \kappa_r \rangle \neq 0$, $\langle \kappa_{r'} \rangle \neq 0$, and thus, $\langle \kappa_r|_{w_2} \rangle \neq 0$, $\langle \kappa_{r'}|_{w_2} \rangle \neq 0$. Since S is extremal, and thus, zero-sum free, we obtain $[\![\xi \cdot \zeta|_{w_2}]\!]_{p'} \neq 0$, $[\![\xi \cdot \zeta|_{w_2}]\!]_{q'} \neq 0$.

If z does not occur in $\zeta|_{w_2}$: we derive in a similar fashion that $\langle \kappa_r|_{w_2} \rangle \neq 0$, $\langle \kappa_{r'}|_{w_2} \rangle \neq 0$, and thus, $[\![\zeta|_{w_2}]\!]_{p'} \neq 0$, $[\![\zeta|_{w_2}]\!]_{q'} \neq 0$.

Second, by the twins property, we have that $(p', q') \in \textsc{Twins}(\mathcal{A} \cup \bar{\mathcal{A}})$. Using again that $\langle \kappa_r \rangle \neq 0$, $\langle \kappa_{r'} \rangle \neq 0$, we derive $\langle \kappa_r[z]_{w_2}|_{w_1} \rangle \neq 0$, $\langle \kappa_{r'}[z]_{w_2}|_{w_1} \rangle \neq 0$. Hence, $[\![e_{p'} \cdot \zeta'']\!]_{p'} \neq 0$, $[\![e_{q'} \cdot \zeta'']\!]_{q'} \neq 0$. Consequently, we have (†). ■

We note that $u \in S^{Q \cup \bar{Q}}$, $u \neq \tilde{0}$, is a critical vector iff $f(u)$ is a critical vector. Hence, applying the factorization to $T_{p,q}$ for every $(p, q) \in \textsc{Siblings}(\mathcal{A})$ results in a compressed search space for critical vectors. It follows from the preceding lemma that the resulting search space is finite.

Lemma 3.4 *Let (f, g) be a maximal factorization of dimension $Q \cup \bar{Q}$. Assume that $\mathcal{A}$ has the twins property. For every $(p, q) \in \textsc{Siblings}(\mathcal{A})$ the set $f(T_{p,q} \setminus \{\tilde{0}\})$ is finite.*

PROOF. By Lemma 3.3 there is a finite set S' with

$$f(T_{p,q} \setminus \{\tilde{0}\}) \subseteq f(S \cdot S') \subseteq f(S') \ ,$$

where we used that (f, g) is maximal. Since S' is finite, so is $f(T_{p,q} \setminus \{\tilde{0}\})$. ■

Algorithm 1 Decision algorithm

Require: $\mathcal{A} = (Q, \Sigma, S, \delta, \nu)$ a wta, S commutative, extremal, (f, g) maximal factorization
Ensure: print "yes" iff $\mathcal{A}$ has the twins property

1: compute SIBLINGS($\mathcal{A}$)
2: **for** $(p, q) \in$ SIBLINGS($\mathcal{A}$) in parallel **do**
3:     **for** $u \in f(T_{p,q} \setminus \{\tilde{0}\})$ **do**
4:         **if** u is a critical vector **then**
5:             print "no" and terminate
6: print "yes"

3.3 Two Decision Algorithms

In this section we consider two decision algorithms. The first one is part of the following proof.

PROOF (OF THM. 3.1). Algorithm 1 proceeds as follows. First, it enumerates SIBLINGS($\mathcal{A}$). This is possible as shown by Obs. 2.8. Second, for each $(p, q) \in$ SIBLINGS($\mathcal{A}$) in parallel, it enumerates $f(T_{p,q} \setminus \{\tilde{0}\})$, checking for critical vectors. For this step, we distinguish two cases.

Either $\mathcal{A}$ has the twins property. Then, by Lemma 3.4, $f(T_{p,q} \setminus \{\tilde{0}\})$ is finite, and the algorithm will terminate without finding any critical vector, in which case it outputs "yes".

Or $\mathcal{A}$ does not have the twins property, but then, by Obs. 3.2, the algorithm is guaranteed to find a critical vector at some point, in which case it outputs "no". Note that the parallel processing (line 2) is critical in this case because there may be $(p, q) \in$ SIBLINGS($\mathcal{A}$) such that $f(T_{p,q} \setminus \{\tilde{0}\})$ is infinite, but does not contain a critical vector. ∎

Note that Algorithm 1 basically enumerates the set $\bigcup_{(p,q)\in\text{SIBLINGS}(\mathcal{A})} f(T_{p,q} \setminus \{\tilde{0}\})$. In principle, this can be done by enumerating C_Σ and computing $f[\![(e_p + e_{\bar{q}}) \cdot \zeta]\!]$ for each $\zeta \in C_\Sigma$. However, the computation of weights already done for subcontexts of ζ is not reused in this approach.

In the following we show an alternative procedure (Algorithm 2) that does not enumerate C_Σ explicitly but works on weight vectors instead, thereby avoiding redundant calculation. This procedure maintains a pair of subsets of $S^{Q\cup\bar{Q}}$. It begins with $(\emptyset, \emptyset)$ and keeps adding vectors by applying a monotone operation F until either the second component contains a critical vector or no new vectors are added.

To this end, we define the unary operation F over pairs of subsets of $S^{Q\cup\bar{Q}}$ by $(T, C) \mapsto$

Algorithm 2 Improved decision algorithm

Require: $\mathcal{A} = (Q, \Sigma, S, \delta, \nu)$ a wta, S commutative, extremal, (f, g) maximal factorization
Ensure: print "yes" iff $\mathcal{A}$ has the twins property

compute SIBLINGS($\mathcal{A}$)
$(T, C) \leftarrow (\emptyset, \emptyset)$
repeat
    $(T', C') \leftarrow (T, C)$
    $(T, C) \leftarrow F(T', C')$ ▷ uses SIBLINGS($\mathcal{A}$)
until C contains critical vector or $C = C'$
if critical vector has been found **then**
    print "no"
else
    print "yes"

(T', C') where T' and C' contain exactly the following elements:

(F1) for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $u_1, \ldots, u_k \in T$, if $[\![\sigma(u_1, \ldots, u_k)]\!] \neq \tilde{0}$, then $f[\![\sigma(u_1, \ldots, u_k)]\!] \in T'$,

(F2) for every $(p, q) \in$ SIBLINGS($\mathcal{A}$), we have $f(e_p + e_{\bar{q}}) \in C'$,

(F3) for every $k \geq 1$, $\sigma \in \Sigma^{(k)}$, $i \in \{1, \ldots, k\}$, $u_i \in C$, and $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_k \in T$, if $[\![\sigma(u_1, \ldots, u_k)]\!] \neq \tilde{0}$, then $f[\![\sigma(u_1, \ldots, u_k)]\!] \in C'$.

Kleene's fixpoint theorem (Wechler, 1992, Sec. 1.5.2, Theorem 7) yields that F has a least fixpoint (where we use the pointwise subset order), and that it can be calculated by the saturation procedure outlined above. In the forthcoming Lemma 3.6, we show that said fixpoint contains the desired set $\bigcup_{(p,q)\in\text{SIBLINGS}(\mathcal{A})} f(T_{p,q} \setminus \{\tilde{0}\})$. This implies both the correctness of our procedure and its termination, by the same line of reasoning as for Algorithm 1. As a preparation we recall two auxiliary statements.

Observation 3.5 *Let S be commutative and (f, g) maximal. Then for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $\xi_1, \ldots, \xi_k \in T_\Sigma(S^Q)$, we have that $[\![\sigma(\xi_1, \ldots, \xi_k)]\!] = [\![\sigma([\![\xi_1]\!], \ldots, [\![\xi_k]\!])]\!]$ and $f[\![\sigma([\![\xi_1]\!], \ldots, [\![\xi_k]\!])]\!] = f[\![\sigma(f[\![\xi_1]\!], \ldots, f[\![\xi_k]\!])]\!]$.*

PROOF. By (Fülöp and Vogler, 2009, Sec 3.2) and (Büchse, May and Vogler, 2010, Lemma 5.5), respectively. ∎

Lemma 3.6 *Let (T^f, C^f) be the least fixpoint of F. Then (i) $T^f = f([\![T_\Sigma]\!] \setminus \{\tilde{0}\})$ and (ii) $C^f = \bigcup_{(p,q)\in\text{SIBLINGS}(\mathcal{A})} f(T_{p,q} \setminus \{\tilde{0}\})$.*

PROOF. In this proof we will often use Obs. 3.5.

For "$\subseteq$" of Statement (i), we refer to (Büchse, May and Vogler, 2010, Lemma 5.8).

We prove "$\supseteq$" of Statement (i) by contradiction. To this end, let $\xi \in T_\Sigma$ a smallest tree such that $[\![\xi]\!] \neq \tilde{0}$ and $f[\![\xi]\!] \notin T^f$. By definition of T_Σ, there are $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $\xi_1, \ldots, \xi_k \in T_\Sigma$ such that $\xi = \sigma(\xi_1, \ldots, \xi_k)$. We derive

$$\begin{aligned} f[\![\sigma(\xi_1, \ldots, \xi_k)]\!] &= f[\![\sigma([\![\xi_1]\!], \ldots, [\![\xi_k]\!])]\!] \\ &= f[\![\sigma(f[\![\xi_1]\!], \ldots, f[\![\xi_k]\!])]\!] \,. \end{aligned}$$

Now either $f[\![\xi_i]\!] \in T^f$ for every $i \in \{1, \ldots, k\}$, but then so is $f[\![\xi]\!]$, or ξ was not the smallest counterexample.

For "$\subseteq$" of Statement (ii), we show that $\big(T^f, \bigcup_{(p,q)\in\text{SIBLINGS}(\mathcal{A})} f(T_{p,q} \setminus \{\tilde{0}\})\big)$ is a prefixpoint of F. It is easy to see that (F1) and (F2) hold. Now let $k, \sigma, i, u_1, \ldots, u_k$ as in (F3) such that $[\![\sigma(u_1, \ldots, u_k)]\!] \neq \tilde{0}$. Hence, $u_1, \ldots, u_k \neq \tilde{0}$. By (i) there are $\xi_1, \ldots, \xi_{i-1}, \xi_{i+1}, \ldots, \xi_k$ such that $u_j = f[\![\xi_j]\!]$ for $j \neq i$. Moreover there are $(p, q) \in \text{SIBLINGS}(\mathcal{A})$ and $\zeta_i \in C_\Sigma$ such that $u_i = f[\![(e_p + e_{\bar{q}}) \cdot \zeta_i]\!]$. We derive

$$\begin{aligned} & f[\![\sigma(u_1, \ldots, u_k)]\!] \\ &= f[\![\sigma(f[\![\xi_1]\!], \ldots, f[\![(e_p + e_{\bar{q}}) \cdot \zeta_i]\!], \ldots, f[\![\xi_k]\!])]\!] \\ &= f[\![\sigma([\![\xi_1]\!], \ldots, [\![(e_p + e_{\bar{q}}) \cdot \zeta_i]\!], \ldots, [\![\xi_k]\!])]\!] \\ &= f[\![\sigma(\xi_1, \ldots, (e_p + e_{\bar{q}}) \cdot \zeta_i, \ldots, \xi_k)]\!] \\ &= f[\![(e_p + e_{\bar{q}}) \cdot \sigma(\xi_1, \ldots, \zeta_i, \ldots, \xi_k)]\!] \,, \end{aligned}$$

which, by definition, is in $f(T_{p,q} \setminus \{\tilde{0}\})$.

We prove "$\supseteq$" of (ii) by contradiction. To this end, let $(p, q) \in \text{SIBLINGS}(\mathcal{A})$ and $\zeta \in C_\Sigma$ a smallest context such that $f[\![(e_p + e_{\bar{q}}) \cdot \zeta]\!] \in f(T_{p,q} \setminus \{\tilde{0}\}) \setminus C^f$. Hence, $[\![(e_p + e_{\bar{q}}) \cdot \zeta]\!] \neq \tilde{0}$. By (F2), we obtain $\zeta \neq z$. Hence, there are $k \geq 1$, $\sigma \in \Sigma^{(k)}$, $i \in \{1, \ldots, k\}$, $\xi_1, \ldots, \xi_{i-1}, \xi_{i+1}, \ldots, \xi_k \in T_\Sigma$, and $\zeta_i \in C_\Sigma$ such that $\zeta = \sigma(\xi_1, \ldots, \xi_{i-1}, \zeta_i, \xi_{i+1}, \ldots, \xi_k)$. We have that $[\![\xi_j]\!] \neq \tilde{0}$ for $j \neq i$, and $[\![(e_p + e_{\bar{q}}) \cdot \zeta_i]\!] \neq \tilde{0}$. We derive

$$\begin{aligned} & f[\![(e_p + e_{\bar{q}}) \cdot \zeta]\!] \\ &= f[\![(e_p + e_{\bar{q}}) \cdot \sigma(\xi_1, \ldots, \zeta_i, \ldots, \xi_k)]\!] \\ &= f[\![\sigma(\xi_1, \ldots, (e_p + e_{\bar{q}}) \cdot \zeta_i, \ldots, \xi_k)]\!] \\ &= f[\![\sigma([\![\xi_1]\!], \ldots, [\![(e_p + e_{\bar{q}}) \cdot \zeta_i]\!], \ldots, [\![\xi_k]\!])]\!] \\ &= f[\![\sigma(f[\![\xi_1]\!], \ldots, f[\![(e_p + e_{\bar{q}}) \cdot \zeta_i]\!], \ldots, f[\![\xi_k]\!])]\!] \end{aligned}$$

By (i), we have that $f[\![\xi_j]\!] \in T^f$. Now either $f[\![(e_p + e_{\bar{q}}) \cdot \zeta_i]\!] \in C^f$, but then so is $f[\![(e_p + e_{\bar{q}}) \cdot \zeta]\!]$, or ζ was not the smallest counterexample. ∎

4 Discussion and Further Research

The notion that the twins property can be decided by searching for critical vectors in a compressed search space is due to Kirsten (2012). We have generalized his work in two ways: (i) We allow arbitrary extremal semifields instead of the tropical semiring. To this end, we use the notion of a maximal factorization, which is implicit in his work. (ii) We consider weighted tree automata instead of weighted string automata. This makes the proof more complex, as we have to distinguish between contexts and trees.

Kirsten's result that deciding the twins property is PSPACE-hard directly transfers to our setting, giving a lower bound on the complexity of our algorithms. In addition, he shows that the problem is PSPACE-complete by giving an algorithm that is in PSPACE. We did not investigate whether this result can be transferred to our setting as well.

To check for critical vectors, Algorithm 1 does not need all components from the vectors in $T_{p,q}$ but only the p- and $\bar{q}$-components; thus in the proof of Lemma 3.3 the height restriction $\text{ht}(\zeta) \leq 2|Q|^{2|Q|}$ for S' can ultimately be lowered to $\text{ht}(\zeta) \leq 2|Q|^2$. It is an open question which of the two algorithms performs better in practice.

For further research, it would be interesting to investigate sufficient properties for determinizability that do not require the semifield to be extremal. Then the determinized wta could truly aggregate the weights of the original runs.

Acknowledgments

The authors wish to thank Heiko Vogler for stimulating remarks on a draft, as well as the anonymous referees for pointing out mistakes. The first author was financially supported by DFG VO 1011/6-1.

References

Büchse, Matthias and May, Jonathan and Vogler, Heiko. 2010. Determinization of Weighted Tree Automata using Factorizations. *J. Autom. Lang. Comb.*, 15(3/4).

Casacuberta, Francisco and de la Higuera, Colin. 2000. Computational complexity of problems on probabilistic grammars and transducers. In *Proc. ICGI*, LNCS. Springer.

Chiang, David. 2007. Hierarchical phrase-based translation. *Comp. Ling.*, 33(2):201–228.

Droste, Manfred and Kuich, Werner and Vogler, Heiko, editors. 2009. *Handbook of Weighted Automata*. EATCS Monographs in Theoretical Computer Science. Springer.

Eppstein, David. 1998. Finding the k shortest paths. *SIAM Journal on Computing*, 28(2):652–673.

Ésik, Zoltán and Kuich, Werner. 2003. Formal tree series. *J. Autom. Lang. Comb.*, 8(2):219–285.

Fülöp, Zoltán and Vogler, Heiko. 2009. Weighted tree automata and tree transducers. In (Droste, Kuich and Vogler, 2009), Chapter 9.

Goguen, Joseph. A. and Thatcher, James W. and Wagner, Eric G. and Wright, Jesse B. 1977. Initial Algebra Semantics and Continuous Algebras. *J. ACM*, 24(1):68–95.

Golan, Jonathan Samuel. 1999. *Semirings and their Applications*. Kluwer Academic Publishers.

Graehl, Jonathan and Knight, Kevin and May, Jonathan. 2008. Training tree transducers. *Comp. Ling.*, 34(3):391–427.

Hebisch, Udo and Weinert, Hanns Joachim. 1998. *Semirings: Algebraic Theory and Applications in Computer Science*, Series in Algebra, volume 5. World Scientific.

Huang, Liang and Chiang, David. 2005. Better k-best parsing. In *Proc. IWPT*, pp. 53–64. ACL.

Kirsten, Daniel. 2012. Decidability, Undecidability, and PSPACE-Completeness of the Twins Property in the Tropical Semiring. *Theoretical Computer Science*, 420:56–63.

Kirsten, Daniel and Mäurer, Ina. 2005. On the determinization of weighted automata. *J. Autom. Lang. Comb.*, 10:287–312.

Li, Zhifei and Eisner, Jason and Khudanpur, Sanjeev. 2009. Variational decoding for statistical machine translation. In *Proc. ACL-IJCNLP*, pp. 593–601. ACL.

Mahr, Bernd. 1984. Iteration and summability in semirings. *Annals of Discrete Mathematics*, 19:229–256.

May, Jonathan and Knight, Kevin. 2006. A better N-best list: practical determinization of weighted finite tree automata. In *Proc. NAACL-HLT*, pp. 351–358. ACL.

Pauls, Adam and Klein, Dan. 2009. k-best A* parsing. In *Proc. ACL-IJCNLP*, pp. 958–966. ACL.

Petrov, Slav and Barrett, Leon and Thibaux, Romain and Klein, Dan. 2006. Learning accurate, compact, and interpretable tree annotation. In *Proc. COLING-ACL*, pp. 433-440.

Sima'an, Khalil. 1996. Computational complexity of probabilistic disambiguation by means of tree-grammars. In *Proc. COLING*, pp. 1175–1180. ACL.

Wechler, Wolfgang. 1992. *Universal Algebra for Computer Scientists*, volume 25 of *Monogr. Theoret. Comput. Sci. EATCS Ser.* Springer.

Zhang, Min and Jiang, Hongfei and Aw, Aiti and Li, Haizhou and Tanm, Chew Lim and Li, Sheng 2008. A tree sequence alignment-based tree-to-tree translation model. In *Proc. ACL*, pp. 559–567. ACL.