A KLEENE Theorem for Weighted Tree Automata over Distributive Multioperator Monoids *

Zoltán Fülöp
Department of Computer Science
University of Szeged
Árpád tér 2., H-6720 Szeged, Hungary
fulop@inf.u-szeged.hu

Andreas Maletti and Heiko Vogler
Faculty of Computer Science
Dresden University of Technology
D-01062 Dresden, Germany
{maletti,vogler}@tcs.inf.tu-dresden.de

September 17, 2007

Abstract: KLEENE's theorem on the equivalence of recognizability and rationality for formal tree series over distributive multioperator monoids is proved. As a consequence of this, KLEENE's theorem for weighted tree automata over arbitrary, i.e., not necessarily commutative, semirings is derived.

1 Introduction

KLEENE's theorem on the equivalence of recognizability and rationality of languages [14] has been extended to various discrete structures such as, e.g., trees [25], trace monoids [21], and pictures [13]. This equivalence (or a slight modification of it) has also been proved for the weighted counterparts where the weights are taken from some semiring, i.e., for formal power series in non-commuting variables [24], formal power series of trees [2, 15, 10, 23, 22, 7], formal power series in partially commuting variables [6], and picture series [19, 20, 4].

Here we focus our attention on formal power series of trees, for short: tree series. These are mappings from the set T_Σ of trees over some ranked alphabet Σ to some monoid $\underline{A}$ of which the elements are called weights. Given a semiring $\underline{K}$, the concept of a weighted tree automaton (for short: wta) over $\underline{K}$ can be defined. A wta over $\underline{K}$ recognizes a tree series over the additive monoid of $\underline{K}$. The various versions of KLEENE's result for tree series differ in the classes of semirings over which recognizability and rationality are defined. In [2, 15] the equivalence between recognizability and rationality is proved for semirings that are commutative, complete, and continuous; the latter two properties are needed in order to solve systems of equations. This result is generalized in [23, 22, 7] in the sense that completeness and continuity can be dropped from the list of restrictions on the semiring; however, commutativity remains as requirement. It is needed in the proof of the fact that the class of recognizable tree series is closed under concatenation (cf. Lemma 6.5 of [7]).

In this paper we will prove KLEENE's result for tree series which are recognized by wta over distributive multioperator monoids (cf. Theorem 8.3). As a consequence of this, we can prove KLEENE's result for wta over *arbitrary* (i.e., not necessarily commutative) semirings (cf. Theorem 8.10).

*This research was supported by the Hungarian Scientific Foundation (OTKA) under Grant T46686 and DAAD-MÖB and DFG GK 334/3

Now let us briefly recall the concept of a wta over multioperator monoids from [17]. A multioperator monoid (for short: M-monoid) is an algebraic structure $\underline{A} = (A, \oplus, \mathbf{0}, \Omega)$ where $(A, \oplus, \mathbf{0})$ is a commutative monoid and Ω is a set of operations on A. It is distributive if every operation of Ω distributes over $\oplus$ and has $\mathbf{0}$ as annihilator. In a wta M over some M-monoid every transition is associated with an operation of Ω. Moreover, if the transition is made at a k-ary input symbol, then the associated operation has arity k. Then, given an input tree t and a run r on t, i.e., a choice of some transition at every node of t, the operations which are associated to transitions occurring in r, are composed according to the structure of t. Eventually, M computes for every run r on t a value in A (and not an operation), because the leaves of t are associated with nullary operations. This value is called the weight of r. If there is more than one run on t, then the weights of the runs are summed up by $\oplus$ to obtain the weight of t, which is denoted by $(S(M), t)$. In this way M recognizes a tree series $S(M)$ such that, for every $t \in T_\Sigma$, the value of $S(M)$ for t is $(S(M), t)$. In [17], wta over a distributive M-monoid $\underline{A}$, called $\underline{A}$-wta, are investigated. The origin of this automaton concept goes back to [2, 15] where it is required that the operations in Ω are multilinear mappings. That means, besides distributivity, it is required that factors can be pulled out of their arguments. Under this requirement, wta over M-monoids are equivalent to wta over commutative semirings (cf. Theorem 8.6 of [12]). In [17] this requirement is dropped, i.e., for distributive M-monoids it is not required that factors can be pulled out of arguments of operations.

It turns out that wta over M-monoids as investigated in [17] are not sufficiently general to prove KLEENE's result. To obtain such a result, we have to add variables which may label leaves in the input trees. Every variable is associated with a unary operation, and thus, variables are handled differently from nullary symbols of the input ranked alphabet. As a consequence, such a wta M recognizes a so called *uniform tree valuation* (again denoted by $S(M)$), i.e., it maps a given input tree t to an operation on $\underline{A}$ (again denoted by $(S(M), t)$) of which the arity is equal to the number of occurrences of variables in t. In fact, we will first prove KLEENE's result for wta with variables over M-monoids (cf. Theorem 8.2). Therefore we will consider recognizability and rationality of uniform tree valuations instead of tree series.

In order to understand the use of commutativity in [7] and to explain why it can be avoided in the general framework of wta with variables over M-monoids, let us first recall briefly the way in which a wta M over some semiring $\underline{K} = (K, \oplus, \odot, \mathbf{0}, \mathbf{1})$ computes a weight $(S(M), t) \in K$ for a given input tree t. With every transition of M a weight in K is associated. Then the weight of a run on t is simply the $\odot$-product of the weights of the chosen transitions. Clearly, if K is not commutative, then we have to prescribe an order on the factors of this product; let us call this order for the time being the *product order*. Finally, if there is more than just one run on t, then $(S(M), t)$ is the $\oplus$-sum of the weights of all the possible runs on t. Now, assume that there are two recognizable tree series S_1 and S_2 (recognized by wta M_1 and M_2, respectively) and a nullary input symbol α; then, roughly speaking, the α-concatenation of S_1 and S_2 is the tree series S such that for every tree t, the weight (S, t) is the sum of the products $(S_1, s) \odot (S_2, t_1) \odot \cdots \odot (S_2, t_k)$ where the sum ranges over all the tuples $(s, t_1, \ldots, t_k)$ such that t is obtained by substituting in s for the ith occurrence of α the tree t_i. Now it becomes clear that there is no product order which can be used by M_1 and M_2 and by the wta that recognizes S: every such product order can be corrupted. And this is the reason why commutativity is assumed in [7].

In wta with variables over M-monoids the problem with the product order disappears. In fact, the concatenation of S_1 and S_2 now takes place at a variable z and not at some nullary input symbol α; moreover, S_1 and S_2 are uniform tree valuations. Then, the z-concatenation of S_1 and S_2 is the uniform tree valuation S such that (S, t) is the sum of all operations $(S_1, s) \circ ((S_2, t_1), \ldots, (S_2, t_k))$ where the sum is taken over all tuples $(s, t_1, \ldots, t_k)$ defined as above (since the sum now operates on operations, this has to be defined appropriately). Moreover, $\circ$ is the composition of operations according to the occurrences of z in s (cf. Definition 5.1).

There is one more technical problem that we want to recall from [25]. If a tree automaton M is analyzed, then the corresponding rational expression has to use the states of M as extra symbols, viz. as concatenation points. Basically this necessity appears because tree concatenation *replaces* leaves of trees. This does not cause any problem, because for a tree language $L \subseteq T_\Sigma$ which is accepted by an automaton M, also $L \subseteq T_\Sigma(Q)$ holds (where Q is the set of states of M, and $T_\Sigma(Q)$ is the set of trees over Σ of which the leaves may also be labeled by elements taken from Q), and the rational expression η is constructed such that it uses symbols from Q but its semantics $[\![\eta]\!]$ is equal to L, i.e., disregards trees that contain symbols of Q. Clearly, the same need for extra concatenation symbols occurs in the case of wta, both, over semirings and M-monoids. Here however, the tree series $S(M)$ which is recognized by M, is a mapping of type $T_\Sigma \to A$ (i.e., a set of pairs from $T_\Sigma \times A$) whereas the semantics $[\![\eta]\!]$ of the corresponding rational expression has the type $T_\Sigma(Q) \to A$, and $[\![\eta]\!]$ maps trees in $T_\Sigma(Q) \setminus T_\Sigma$ to $\mathbf{0}$; thus $S(M) \subset [\![\eta]\!]$ (in this sense Theorem 5.2 of [7] and, in particular, Equation (†) in the proof of that theorem contains a flaw). Of course, there is an easy way out: we simply lift the type of $S(M)$ such that $\mathrm{lift}_Q(S(M))\colon T_\Sigma(Q) \to A$ where every tree in $T_\Sigma(Q) \setminus T_\Sigma$ is mapped to $\mathbf{0}$. Then we have $\mathrm{lift}_Q(S(M)) = [\![\eta]\!]$. We will formally define this lifting in Section 8 and obtain as our main result for wta over some distributive M-monoid $\underline{A}$ the following KLEENE theorems (where $\underline{A}$ has to fulfill mild closure properties):

- for wta with variables (cf. Theorem 8.2):

$$\mathrm{lift}(\mathrm{Rec}(\Sigma, \mathrm{fin}, \underline{A})) = \mathrm{lift}(\mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A})),$$

- and for wta without variables (cf. Theorem 8.3):

$$\mathrm{Rec}(\Sigma, \emptyset, \underline{A}) = \mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A})|_{T_\Sigma},$$

where, as in [10, 7], we define $\mathrm{Rec}(\Sigma, \mathrm{fin}, \underline{A}) = \bigcup_{Z \text{ finite set}} \mathrm{Rec}(\Sigma, Z, \underline{A})$ and similarly for $\mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A})$; the expression $\mathrm{Rec}(\Sigma, Z, \underline{A})$ (and $\mathrm{Rat}(\Sigma, Z, \underline{A})$) denotes the class of all recognizable (respectively, rational) uniform tree valuations over Σ, Z, and $\underline{A}$. Moreover, as usual, for a class Φ of functions, $\Phi|_C$ is the class of restrictions $f|_C$ for $f \in \Phi$.

Finally, by simulating the semiring $\underline{K}$ by an appropriate distributive M-monoid $\underline{D}(\underline{K})$ and by applying Theorem 8.3, we obtain the following KLEENE result for wta over an *arbitrary* semiring $\underline{K}$ (cf. Theorem 8.10):

$$\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{K}) = \mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{D}(\underline{K}))|_{T_\Sigma}$$

where $\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{K})$ denotes the class of all tree series that are recognizable over $\underline{K}$. Thus, in particular, we can express a tree series that is recognized by a wta over a semiring $\underline{K}$, by a rational expression over the distributive M-monoid $\underline{D}(\underline{K})$. We will also show that, when restricted to a commutative semiring $\underline{K}$, our rational expressions over $\underline{D}(\underline{K})$ naturally correspond to the rational expressions of [7]. In this way, we reprove the KLEENE theorem of [7].

The proof of KLEENE's theorem for wta with variables over distributive M-monoids follows the usual main line as taken, e.g., in [25, 7]. However, the technical framework of the present paper is more involved, because we now have to deal with uniform tree valuations rather than tree series. In particular, the proof employs normal form results which, in their turn, require that the set Ω of operations of the M-monoid is appropriately closed.

In Section 2 we recall the notions of operations, trees, Ω-algebras, monoids, semirings, M-monoids, and tree series. In Section 3 we introduce wta with variables over M-monoids. In particular, we define the concept of a uniform tree valuation and some useful operations on uniform tree valuations. For wta we define a run semantics and an inductive semantics and prove that they are equivalent if the M-monoid is distributive. In Section 4 we prove normal forms for wta that concern uniqueness of particular states. In Section 5 we introduce rational operations over M-monoids. In Section 6 we analyze a wta and construct a corresponding rational expression; here we use the notion of a decomposition of a run. And in Section 7 we synthesize wta from rational expressions; there we use the closure of the class of recognizable uniform tree valuations under relabelings. In Section 8 we present the Kleene result for wta with variables over M-monoids, and we apply this result to the case of semirings and obtain Kleene's result for arbitrary semirings.

2 Preliminaries

2.1 Sets, mappings, and operations

For a set A we denote by $\mathcal{P}(A)$ the power set of A and by A^* the set of *strings* over A. The symbol ε denotes the *empty string*. We denote the set of nonnegative integers by $\mathbb{N}$, and for $k \in \mathbb{N}$, we denote the set $\{1, \ldots, k\}$ by $[k]$. Thus $[0] = \emptyset$.

We use the *lexicographic ordering on* $\mathbb{N}^*$ and write $v <_{\text{lex}} w$ whenever v is smaller than w with respect to the usual lexicographic ordering on sequences of natural numbers. We extend this relation to nonempty sets $U, V \subseteq \mathbb{N}^*$ by defining $U <_{\text{lex}} V$ if $u <_{\text{lex}} v$ for every $u \in U$ and $v \in V$. Note that $<_{\text{lex}}$ is does not yield a partial order on $\mathcal{P}(\mathbb{N}^*) \setminus \{\emptyset\}$.

Let $f\colon A \to B$ be a function and $C \subseteq A$. The *restriction of f to C* is the mapping $f|_C\colon C \to B$ that is defined for every $c \in C$ by $f|_C(c) = f(c)$. For a class $\Phi \subseteq \{f \mid f\colon A \to B\}$ of functions we extend the restriction by $\Phi|_C = \{f|_C \mid f \in \Phi\}$.

Let A be nonempty. For every $k \geq 0$, we denote the set of all k-ary operations over A and the set of all operations over A by $\mathrm{Ops}^k(A)$ and $\mathrm{Ops}(A)$, respectively. For $\Omega \subseteq \mathrm{Ops}(A)$ we let $\Omega^{(k)} = \Omega \cap \mathrm{Ops}^k(A)$. For $\omega \in \mathrm{Ops}^k(A)$ and $C \subseteq A$, we abbreviate $\omega|_{C^k}$ by $\omega|_C$ and we lift up this abbreviation to sets of mappings, i.e., we set $\Omega|_C = \{\omega|_C \mid \omega \in \Omega\}$. As usual we identify every nullary operation $f\colon A^0 \to A$ over A with the element $f() \in A$. Thus we do not distinguish between $\mathrm{Ops}^0(A)$ and A. Moreover, we denote by id_A the *identity function over A*, which is defined by $\mathrm{id}_A(a) = a$ for every $a \in A$.

2.2 Ranked alphabets and trees

A *ranked set* is a pair (Ω, rk) where Ω is a set and $\mathrm{rk}\colon \Omega \to \mathbb{N}$ (called rank function). Usually we denote the ranked set simply by Ω. For every $k \geq 0$ we denote the set $\{\omega \in \Omega \mid \mathrm{rk}(\omega) = k\}$ by $\Omega^{(k)}$. In the rest of the paper we assume that $\Omega^{(0)} \neq \emptyset$ for every ranked set Ω that we consider. Given two ranked sets (Ω, rk) and (Ω', rk') such that $\Omega \cap \Omega' = \emptyset$ we write $\Omega \cup \Omega'$ to denote the ranked set $(\Omega \cup \Omega', \mathrm{rk}'')$ where $\mathrm{rk}''(\omega) = \mathrm{rk}(\omega)$ for every $\omega \in \Omega$ and $\mathrm{rk}''(\omega') = \mathrm{rk}'(\omega')$ for every $\omega' \in \Omega'$. A ranked set (Σ, rk) with Σ finite is called a *ranked alphabet*.

Let Σ be a ranked alphabet and H an arbitrary set. The set of ΣH*-trees*, denoted by $T_\Sigma(H)$, is defined as usual inductively as the set of well-formed expressions over Σ which may have elements of H as leaf labels. If $H = \emptyset$, then we write T_Σ rather than $T_\Sigma(H)$.

We define the *set of positions* in a tree by means of the mapping $\mathrm{pos}\colon T_\Sigma(H) \to \mathcal{P}(\mathbb{N}^*)$, which is inductively defined as follows: (i) if $t \in \Sigma^{(0)} \cup H$, then $\mathrm{pos}(t) = \{\varepsilon\}$, and (ii) if $t = \sigma(t_1, \ldots, t_k)$ for some $\sigma \in \Sigma^{(k)}$, $k \geq 1$, and $t_1, \ldots, t_k \in T_\Sigma(H)$, then $\mathrm{pos}(t) = \{\varepsilon\} \cup \{iw \mid i \in [k], w \in \mathrm{pos}(t_i)\}$. We write $\mathrm{ht}(t)$ for $\max\{|w| \mid w \in \mathrm{pos}(t)\}$ and call it the *height* of t. The *size* of t is the cardinality of $\mathrm{pos}(t)$.

For every tree $t \in T_\Sigma(H)$ and each of its positions $w \in \mathrm{pos}(t)$, we define the *label of t at w*, the *subtree of t at w*, and the *replacement in t at w by a tree s*, denoted by $t(w)$, $t|_w$, and $t[w \leftarrow s]$, respectively, by induction as follows:

- if $t \in \Sigma^{(0)} \cup H$ (and thus $w = \varepsilon$), then $t(\varepsilon) = t|_\varepsilon = t$ and $t[\varepsilon \leftarrow s] = s$; and
- if $t = \sigma(t_1, \ldots, t_k)$ for some $\sigma \in \Sigma^{(k)}$ with $k \geq 1$ and $t_1, \ldots, t_k \in T_\Sigma(H)$, then
  - $t(\varepsilon) = \sigma$ and $t|_\varepsilon = t$ and $t[\varepsilon \leftarrow s] = s$; and
  - $t(w) = t_i(v)$ and $t|_w = t_i|_v$ and $t[w \leftarrow s] = \sigma(t_1, \ldots, t_{i-1}, t_i[v \leftarrow s], t_{i+1}, \ldots, t_k)$ whenever $w = iv$ for some $1 \leq i \leq k$ and $v \in \mathrm{pos}(t_i)$.

We abbreviate several replacements $t[w_1 \leftarrow s_1][w_2 \leftarrow s_2] \cdots [w_n \leftarrow s_n]$ in incomparable (with respect to the prefix-order) positions $w_1, \ldots, w_n$ of t to $t[w_1 \leftarrow s_1, \ldots, w_n \leftarrow s_n]$.

Let $V \subseteq \Sigma \cup H$. The set $\{w \in \mathrm{pos}(t) \mid t(w) \in V\}$ of all positions of t that are labelled with an element of V is denoted by $\mathrm{pos}_V(t)$. The cardinality of the set $\mathrm{pos}_V(t)$ is denoted by $|t|_V$. We write $\mathrm{pos}_v(t)$ and $|t|_v$ if $V = \{v\}$.

Let $a \in H$ and $r = |t|_a$. Moreover, let $s_1, \ldots, s_r \in T_\Sigma(H)$. Define $t[a \leftarrow (s_1, \ldots, s_r)] = t[w_1 \leftarrow s_1, \ldots, w_r \leftarrow s_r]$ where $\{w_1, \ldots, w_r\} = \mathrm{pos}_a(t)$ and $w_1 <_{\mathrm{lex}} \cdots <_{\mathrm{lex}} w_r$.

For a finite set Q, we write $\langle \Sigma, Q\rangle$ for $\Sigma \times Q$ and define the ranked alphabet $(\langle \Sigma, Q\rangle, \mathrm{rk}')$ by letting $\mathrm{rk}'(\langle \sigma, q\rangle) = \mathrm{rk}(\sigma)$ for every $\sigma \in \Sigma$ and $q \in Q$. We define the two projection mappings $\pi_1\colon \langle \Sigma, Q\rangle \to \Sigma$ and $\pi_2\colon \langle \Sigma, Q\rangle \to Q$ in the obvious way. We will also use these notations and projections for an arbitrary set H instead of the ranked alphabet Σ. We extend π_1 to a mapping $\pi_1\colon T_{\langle \Sigma, Q\rangle}(\langle H, Q\rangle) \to T_\Sigma(H)$ such that, for every $t \in T_{\langle \Sigma, Q\rangle}(\langle H, Q\rangle)$, the tree $\pi_1(t)$ is the ΣH-tree obtained from t by dropping q at every node of the form $\langle \sigma, q\rangle$ or $\langle h, q\rangle$ with $\sigma \in \Sigma$ and $h \in H$.

2.3 Ω-algebras, monoids, semirings, and M-monoids

Let Ω be a ranked set. An *Ω-algebra* (A, Ω_A) consists of a nonempty set A and a family

$$\Omega_A = (\omega_A : A^m \to A \mid m \geq 0,\, \omega \in \Omega^{(m)})$$

of operations on A. If the meaning is clear from the context, then we do not make a distinction between ω and ω_A and simply drop A from ω_A. Also we identify the ranked set Ω with the family Ω_A. If Ω is the finite set $\{f_1, \ldots, f_k\}$, then we also denote the Ω-algebra by $(A, f_1, \ldots, f_k)$.

A *monoid* is an algebra $(A, \otimes, \mathbf{1})$ where $\otimes$ is a binary, associative operation over A and $\mathbf{1}$ is the neutral element with respect to $\otimes$. A monoid is *commutative* if $\otimes$ is commutative.

A *semiring* is an algebra $\underline{K} = (K, \oplus, \odot, \mathbf{0}, \mathbf{1})$ where $(K, \oplus, \mathbf{0})$ is a commutative monoid (called the underlying additive monoid), $(K, \odot, \mathbf{1})$ is a monoid (called the underlying multiplicative monoid), and the distributivity laws (d1-SR) and (d2-SR) and the absorption law (a-SR) hold, i. e., for every $a, b, c \in K$:

$$a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c) \tag{d1-SR}$$

$$(a \oplus b) \odot c = (a \odot c) \oplus (b \odot c) \tag{d2-SR}$$

$$a \odot \mathbf{0} = \mathbf{0} \odot a = \mathbf{0}. \tag{a-SR}$$

A semiring is *commutative* if the underlying multiplicative monoid is commutative. We adopt the convention that $\odot$ has a higher binding priority than $\oplus$ and drop the parentheses around products.

Let $(A, \oplus, \mathbf{0})$ be a commutative monoid. For every $k \geq 0$, we define the operation $\mathbf{0}^k \colon A^k \to A$ by $\mathbf{0}^k(a_1, \ldots, a_k) = \mathbf{0}$ for every $a_1, \ldots, a_k \in A$. Moreover, let $\omega \colon A^k \to A$ be a k-ary operation on A. We say that ω is *distributive* (with respect to $(A, \oplus, \mathbf{0})$) if for every $a_1, \ldots, a_k \in A$, $1 \leq i \leq k$, and $a, a' \in A$, the distributivity law (d-M) and the absorption law (a-M) hold, i.e.,

$$\begin{aligned} &\omega(a_1, \ldots, a_{i-1}, a \oplus a', a_{i+1}, \ldots, a_k) \\ &= \omega(a_1, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_k) \oplus \omega(a_1, \ldots, a_{i-1}, a', a_{i+1}, \ldots, a_k) \end{aligned} \tag{d-M}$$

$$\mathbf{0} = \omega(a_1, \ldots, a_{i-1}, \mathbf{0}, a_{i+1}, \ldots, a_k). \tag{a-M}$$

A *multioperator monoid* (shortly, M-monoid) is an algebra $\underline{A} = (A, \oplus, \mathbf{0}, \Omega)$, where

- $(A, \oplus, \mathbf{0})$ is a commutative monoid (called the underlying monoid),
- (A, Ω) is an Ω-algebra, and
- $\mathrm{id}_A \in \Omega^{(1)}$ and $\mathbf{0}^k \in \Omega^{(k)}$ for every $k \geq 0$ (note that this condition slightly restricts the general notion but simplifies the following development).

An M-monoid $\underline{A}$ is *distributive* if for every $\omega \in \Omega$ the operation is distributive (with respect to $(A, \oplus, \mathbf{0})$). Clearly, $\mathbf{0}^k$ and id_A are distributive for every $k \geq 0$. We call a distributive M-monoid a DM-monoid. We note that in [10] and [17] a DM-monoid $\underline{A}$ is called distributive Ω-algebra. Also we note that in [5] a distributive M-monoid with an idempotent addition is called distributive Ω-magma.

2.4 Tree series

Let $\underline{A} = (A, \oplus, \mathbf{0})$ be a commutative monoid and Σ a ranked alphabet. A *tree series (over Σ and $\underline{A}$)* is a mapping $\varphi \colon T_\Sigma \to A$. For every $t \in T_\Sigma$, the element $\varphi(t) \in A$ is called the *coefficient* of t, and it is denoted by (φ, t). If there exists an $a \in A$ such that for every $t \in T_\Sigma$, we have $(\varphi, t) = a$, then φ is a *constant* and also denoted by $\widetilde{a}$. The *set of all tree series over Σ and $\underline{A}$* is denoted by $A\langle\langle T_\Sigma \rangle\rangle$.

The *support of a tree series* $\varphi \in A\langle\langle T_\Sigma \rangle\rangle$ is the set $\mathrm{supp}(\varphi) = \{t \in T_\Sigma \mid (\varphi, t) \neq \mathbf{0}\}$. Moreover, φ is called a *monomial* if $\mathrm{supp}(\varphi)$ is empty or a singleton. A monomial φ is denoted by $a.t$ if $(\varphi, t) = a$ and $(\varphi, s) = \mathbf{0}$ for every tree $s \neq t$.

In fact, $(A\langle\langle T_\Sigma \rangle\rangle, \oplus, \widetilde{\mathbf{0}})$ is a commutative monoid where, for every tree series $\varphi, \psi \in A\langle\langle T_\Sigma \rangle\rangle$ and $t \in T_\Sigma$, we define $(\varphi \oplus \psi, t) = (\varphi, t) \oplus (\psi, t)$. Let I be an index set and $(\varphi_i \in A\langle\langle T_\Sigma \rangle\rangle \mid i \in I)$ a family of tree series. The family is *locally finite* if for every $t \in T_\Sigma$, the set $I_{\mathrm{supp}}(t) = \{i \in I \mid t \in \mathrm{supp}(\varphi_i)\}$ is finite. Now, if the family of tree series is locally finite, then we can define the sum $\bigoplus_{i \in I} \varphi_i \in A\langle\langle T_\Sigma \rangle\rangle$ by $((\bigoplus_{i \in I} \varphi_i), t) = \bigoplus_{i \in I_{\mathrm{supp}}(t)} (\varphi_i, t)$ for every $t \in T_\Sigma$. It is easy to see that, for every tree series $\varphi \in A\langle\langle T_\Sigma \rangle\rangle$, the equation $\varphi = \bigoplus_{t \in T_\Sigma} (\varphi, t).t$ holds, because the family $((\varphi, t).t \mid t \in T_\Sigma)$ of monomials is locally finite.

3 Weighted tree automata over M-monoids

In this section we define the concept of weighted tree automata with variables over some M-monoid. As already indicated in the Introduction, such an automaton M recognizes a mapping $S(M)$ of the type $T_\Sigma(Z) \to \mathrm{Ops}(A)$ where Z is the finite set of variables used by M, and A is the carrier set of the M-monoid of M. Moreover, for every input tree $t \in T_\Sigma(Z)$ the arity of $S(M)(t)$ is equal to the number of occurrences of variables in t. We will call a mapping with this property a *uniform tree valuation* and we denote the set of all uniform tree valuations over Σ, Z, and $\underline{A}$ by $\mathrm{Uvals}(\Sigma, Z, \underline{A})$. In the next subsection we will formally define the concept of a uniform tree valuation and some useful operations on $\mathrm{Ops}(A)$.

Throughout this paper, let Σ be a ranked alphabet, Z a finite set (of variables), and $\underline{A} = (A, \oplus, \mathbf{0}, \Omega)$ an M-monoid. We abbreviate id_A to simply id.

3.1 Uniform tree valuations

Uniform tree valuations are a slight generalization of tree series. In fact, if $Z = \emptyset$, then $\mathrm{Uvals}(\Sigma, Z, \underline{A}) = \underline{A}\langle\langle T_\Sigma \rangle\rangle$. For this reason we will henceforth use the notation introduced for tree series in Section 2.4 also for uniform tree valuations. So, let $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$. We write (ψ, t) instead of $\psi(t)$. Moreover, we use $\widetilde{\mathbf{0}}$ to denote the uniform tree valuation with $(\widetilde{\mathbf{0}}, t) = \mathbf{0}^{|t|_Z}$ for every $t \in T_\Sigma(Z)$. When speaking about a uniform tree valuation ψ, the arity of (ψ, t) is fixed for every $t \in T_\Sigma(Z)$. Thus, we will write $(\psi, t) = \mathbf{0}$ when formally we mean $(\psi, t) = \mathbf{0}^{|t|_Z}$. The *support* of ψ is defined to be the set $\mathrm{supp}(\psi) = \{t \in T_\Sigma(Z) \mid (\psi, t) \neq \mathbf{0}\}$. Finally, given $t \in T_\Sigma(Z)$ and $\omega \in \mathrm{Ops}^{|t|_Z}(A)$ we write $\omega.t$ for the uniform tree valuation such that $(\omega.t, t) = \omega$ and $(\omega.t, s) = \mathbf{0}$ for every $s \in T_\Sigma(Z)$ with $s \neq t$. Such a uniform tree valuation is also called a *monomial*. For the summation of two uniform tree valuations, we first need a summation of operations. We return to uniform tree valuations at the end of this section.

We proceed by considering two partial operations on $\mathrm{Ops}(A)$, which can, in particular, be used to construct uniform tree valuations. In addition, we state some simple properties of those operations.

Definition 3.1. Let $k \geq 0$.

- Let $\omega_1, \omega_2 \in \mathrm{Ops}^k(A)$. The *sum of* ω_1 *and* ω_2 is the k-ary operation $\omega_1 \oplus \omega_2$ that is defined, for every $\vec{a} \in A^k$, by $(\omega_1 \oplus \omega_2)(\vec{a}) = \omega_1(\vec{a}) \oplus \omega_2(\vec{a})$.
- Let $\omega \in \mathrm{Ops}^k(A)$ and $\omega_j \in \mathrm{Ops}^{l_j}(A)$ with $l_j \geq 0$ for every $1 \leq j \leq k$. The *composition of* ω *with* $(\omega_1, \ldots, \omega_k)$ is the $(l_1 + \cdots + l_k)$-ary operation $\omega(\omega_1, \ldots, \omega_k)$ that is defined by
$$\big(\omega(\omega_1, \ldots, \omega_k)\big)(\vec{a_1}, \ldots, \vec{a_k}) = \omega(\omega_1(\vec{a_1}), \ldots, \omega_k(\vec{a_k}))$$
for every $\vec{a_j} \in A^{l_j}$ with $1 \leq j \leq k$.

By the above definitions, $(\mathrm{Ops}^k(A), \oplus, \mathbf{0}^k)$ is a commutative monoid for every $k \geq 0$, which for $k = 0$ is isomorphic to the monoid $(A, \oplus, \mathbf{0})$. Next, let us observe left-distributivity, right-distributivity, and associativity of composition.

Observation 3.2. Let $k \geq 0$ and $\omega, \omega' \in \mathrm{Ops}^k(A)$. Moreover, for every $1 \leq j \leq k$ let $l_j \geq 0$ and $\omega_j \in \mathrm{Ops}^{l_j}(A)$. Then

$$(\omega \oplus \omega')(\omega_1, \ldots, \omega_k) = \omega(\omega_1, \ldots, \omega_k) \oplus \omega'(\omega_1, \ldots, \omega_k).$$

Moreover, $\mathbf{0}^k(\omega_1, \ldots, \omega_k) = \mathbf{0}^{l_1 + \cdots + l_k}$.

Observation 3.3. Let $k \geq 0$ and $\omega \in \mathrm{Ops}^k(A)$ be distributive. Moreover, for every $1 \leq j \leq k$ let $l_j \geq 0$ and $\omega_j \in \mathrm{Ops}^{l_j}(A)$. Finally, let $1 \leq i \leq k$ and $\nu, \nu' \in \mathrm{Ops}^{l_i}(A)$. Then

$$\omega(\omega_1, \ldots, \omega_{i-1}, \nu \oplus \nu', \omega_{i+1}, \ldots, \omega_k)$$
$$= \omega(\omega_1, \ldots, \omega_{i-1}, \nu, \omega_{i+1}, \ldots, \omega_k) \oplus \omega(\omega_1, \ldots, \omega_{i-1}, \nu', \omega_{i+1}, \ldots, \omega_k).$$

Moreover, $\mathbf{0}^{l_1+\cdots+l_k} = \omega(\omega_1, \ldots, \omega_{i-1}, \mathbf{0}^{l_i}, \omega_{i+1}, \ldots, \omega_k)$.

Observation 3.4. Let $k \geq 0$, $\omega \in \mathrm{Ops}^k(A)$, and let $l_j \geq 0$, $\omega_j \in \mathrm{Ops}^{l_j}(A)$ for every $1 \leq j \leq k$. Finally, let $\omega_{j,i} \in \mathrm{Ops}(A)$ for every $1 \leq j \leq k$ and $1 \leq i \leq l_j$. Then

$$\begin{aligned}&\big(\omega(\omega_1, \ldots, \omega_k)\big)(\omega_{1,1}, \ldots, \omega_{1,l_1}, \ldots, \omega_{k,1}, \ldots, \omega_{k,l_k})\\ &= \omega\big(\omega_1(\omega_{1,1}, \ldots, \omega_{1,l_1}), \ldots, \omega_k(\omega_{k,1}, \ldots, \omega_{k,l_k})\big).\end{aligned}$$

Let us return to uniform tree valuations and define the sum of two uniform tree valuations $\psi_1, \psi_2 \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$. The *sum of* ψ_1 *and* ψ_2 is the uniform tree valuation that we denote by $\psi_1 \oplus^{\mathrm{u}} \psi_2$ and define by $(\psi_1 \oplus^{\mathrm{u}} \psi_2, t) = (\psi_1, t) \oplus (\psi_2, t)$ for every $t \in T_\Sigma(Z)$. In fact, this summation will be one of the rational operations.

We note that $(\mathrm{Uvals}(\Sigma, Z, \underline{A}), \oplus^{\mathrm{u}}, \widetilde{\mathbf{0}})$ is a commutative monoid; for $Z = \emptyset$ it is the commutative monoid $(A\langle\!\langle T_\Sigma \rangle\!\rangle, \oplus, \widetilde{\mathbf{0}})$. Let I be an index set and $(\psi_i \in \mathrm{Uvals}(\Sigma, Z, \underline{A}) \mid i \in I)$ a family of uniform tree valuations. The family is *locally finite* if for every $t \in T_\Sigma(Z)$, the set $I_{\mathrm{supp}}(t) = \{i \in I \mid t \in \mathrm{supp}(\psi_i)\}$ is finite. Now, if the family of uniform tree valuations is locally finite, then we can define the sum $\bigoplus^{\mathrm{u}}_{i \in I} \psi_i \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ by $((\bigoplus^{\mathrm{u}}_{i \in I} \psi_i), t) = \bigoplus_{i \in I_{\mathrm{supp}}(t)} (\psi_i, t)$ for every $t \in T_\Sigma(Z)$. It is easy to see that, for every uniform tree valuation $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$, the equation $\psi = \bigoplus^{\mathrm{u}}_{t \in T_\Sigma(Z)} (\psi, t).t$ holds, because the family $((\psi, t).t \mid t \in T_\Sigma(Z))$ of monomials is locally finite.

3.2 The automaton model

As mentioned in the Introduction, the weighted tree automaton model of [17] processes trees of T_Σ. We will now extend this model to one that is able to process trees of $T_\Sigma(Z)$. In the theory of unweighted tree automata this is usually achieved by stating that $T_\Sigma(Z)$ is essentially $T_{\Sigma \cup Z}$ where all the elements of Z are treated as nullary symbols. For our purposes the variables are special. They act as nullary symbols for all purposes of constructing trees, but they are assigned a *unary* operation (i.e., one of $\Omega^{(1)}$) by the automaton because in our intention they are placeholders. While a tree replaces the variable it is substituted for, the weight of the tree to be substituted is processed by the unary operation associated to the variable. Often this operation will be the identity operation, which would correspond to a replacement.

Let us define the syntax of our extended model first. Basically, we introduce a new component that will handle variables. In the absence of variables this component becomes meaningless.

Definition 3.5. A *weighted tree automaton with variables* is a tuple $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ where

- Q is a finite, nonempty set of *states*,
- Σ is a ranked alphabet of *input symbols*,
- Z is a finite set of *variables*,

- $\underline{A} = (A, \oplus, \mathbf{0}, \Omega)$ is an M-monoid,
- $F \in (\Omega^{(1)})^Q$ is a *final distribution function*,
- $\mu = (\mu_k \colon \Sigma^{(k)} \to (\Omega^{(k)})^{Q^k \times Q} \mid k \geq 0)$ is a family of *transition mappings*, where we view Q^k as the set of strings over Q of length k, and
- $\nu \colon Z \to (\Omega^{(1)})^Q$ is a *variable assignment*.

Such an automaton M is said to be over Σ, Z, and $\underline{A}$. In order to save parentheses, we will write $\mu_k(\sigma)_{q_1 \dots q_k, q}$, F_q, and $\nu(z)_q$ rather than $\mu_k(\sigma)(q_1 \dots q_k, q)$, $F(q)$, and $\nu(z)(q)$, respectively. We commonly call F_q the q-entry of F and use the same terminology also with μ and ν. In the sequel, we will abbreviate 'weighted tree automaton with variables' by wta (and we use that also for the plural). Since the arity of an entry in $\mu_k(\sigma)$ with $k \geq 0$ and $\sigma \in \Sigma^{(k)}$ is fixed to k, for every $q, q_1, \dots, q_k \in Q$ we will write $\mu_k(\sigma)_{q_1 \dots q_k, q} = \mathbf{0}$ when we mean that $\mu_k(\sigma)_{q_1 \dots q_k, q} = \mathbf{0}^k$. We apply analogous conventions also to F and ν. Moreover, we will say that there exists a σ-transition (in M) from $q_1 \dots q_k$ into q, whenever $\mu_k(\sigma)_{q_1 \dots q_k, q} \neq \mathbf{0}$. Analogously, we will say that there exists a z-transition into q whenever $\nu(z)_q \neq \mathbf{0}$, and that $\nu(z)_q$ is the weight of the z-transition into q.

We will introduce both a run semantics and an inductive semantics for wta. Our reference semantics will be the run semantics but we will resort to the inductive semantics in several proofs. We show that if the underlying M-monoid of a wta M is distributive, then there is no difference between the run semantics and the inductive semantics of M. Let us start with the run semantics.

For the rest of this section, $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ stands for an arbitrary wta.

Definition 3.6. Let $t \in T_\Sigma(Z)$, $X \subseteq Z$, $P \subseteq Q$, and $q \in Q$.

- The *set R_M of all runs of M* is the set $T_{\langle \Sigma, Q \rangle}(\langle Z, Q \rangle)$.
- The *set $R_M^{X,P}$ of all runs of M that use only states from P at symbols of Σ and at variables in $Z \setminus X$ (for short: using P outside X)* is

  $$\{r \in R_M \mid \forall w \in \mathrm{pos}(r) \setminus \{\varepsilon\} \colon r(w) \in \langle \Sigma \cup Z \setminus X, P \rangle \cup \langle X, Q \rangle\} \ .$$

  (Note that the root of r is not restricted.)
- The *set $R_M^{X,P}(t)$ of all runs of M on t using P outside X* is $\{r \in R_M^{X,P} \mid \pi_1(r) = t\}$.
- The *set $R_M^{X,P}(t, q)$ of all q-runs of M on t using P outside X* is $\{r \in R_M^{X,P}(t) \mid \pi_2(r(\varepsilon)) = q\}$.

If $X = Z$ or $P = Q$, then we drop the corresponding superscript(s) from $R_M^{X,P}$, $R_M^{X,P}(t)$, and $R_M^{X,P}(t, q)$. We note that the two parameters X and P provide the flexibility that is needed in Sections 6 and 7. Using these notions of runs, we now define a semantics based on runs for wta.

Definition 3.7.

- The *weight mapping* $c_M \colon R_M \to \mathrm{Ops}(A)$ is defined by induction as follows. For every $z \in Z$ and $q \in Q$ we define $c_M(\langle z, q \rangle) = \nu(z)_q$, and for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, $q \in Q$, and $r_1, \dots, r_k \in R_M$ we define

  $$c_M(\langle \sigma, q \rangle(r_1, \dots, r_k)) = \mu_k(\sigma)_{q_1 \dots q_k, q}(c_M(r_1), \dots, c_M(r_k)),$$

where $q_i = \pi_2(r_i(\varepsilon))$ for every $1 \leq i \leq k$ and $\mu_k(\sigma)_{q_1\ldots q_k,q}(c_M(r_1),\ldots,c_M(r_k))$ is a composition of operations, as defined in Definition 3.1. Note that for every $t \in T_\Sigma(Z)$ and $r \in R_M(t)$ we have $c_M(r) \in \mathrm{Ops}^{|t|_Z}(A)$.

- Let $q \in Q$ be a state. *The uniform tree valuation recognized by M in state q* is the mapping $S(M)_q\colon T_\Sigma(Z) \to \mathrm{Ops}(A)$, which is defined for every $t \in T_\Sigma(Z)$ by

$$(S(M)_q, t) = \bigoplus_{r \in R_M(t,q)} c_M(r).$$

- Finally, *the uniform tree valuation recognized by M* is the mapping $S(M)\colon T_\Sigma(Z) \to \mathrm{Ops}(A)$ defined for every $t \in T_\Sigma(Z)$ by

$$(S(M), t) = \bigoplus_{q \in Q} F_q((S(M)_q, t)).$$

Thus $S(M)_q$ and $S(M)$ are uniform tree valuations.

Definition 3.8. Let $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$. We say that ψ is *recognizable*, if there exists a wta M over Σ, Z, and $\underline{A}$ such that $S(M) = \psi$. The class of all recognizable uniform tree valuations of $\mathrm{Uvals}(\Sigma, Z, \underline{A})$ is denoted by $\mathrm{Rec}(\Sigma, Z, \underline{A})$. Furthermore, we abbreviate $\bigcup_{Z \text{ finite set}} \mathrm{Rec}(\Sigma, Z, \underline{A})$ by $\mathrm{Rec}(\Sigma, \mathrm{fin}, \underline{A})$.

Now we relate our wta model to the wta model of [17]. In fact, we show that our wta with $Z = \emptyset$ is semantically equivalent with the wta of [17] if the underlying M-monoid is distributive.

Observation 3.9. If $\underline{A}$ is distributive, then $(S(M), t) = \bigoplus_{r \in R_M(t)} F_{\pi_2(r(\varepsilon))}(c_M(r))$ for every $t \in T_\Sigma(Z)$.

Proof. We observe that $(S(M), t) = \bigoplus_{q \in Q} F_q(\bigoplus_{r \in R_M(t,q)} c_M(r))$ by Definition 3.7. Due to Observation 3.3 (with $\omega = F_q$) the latter equals $\bigoplus_{q \in Q} \bigoplus_{r \in R_M(t,q)} F_q(c_M(r))$. Clearly, this is the desired result. □

Next let us give three examples of wta over M-monoids in order to illustrate the power of this concept and to show several M-monoids and DM-monoids.

Example 3.10. Let Σ be a ranked alphabet. The simple function $\mathrm{ht} : T_\Sigma \to \mathbb{N}$ can be computed by wta over M-monoids as follows. Let $M = (Q, \Sigma, \emptyset, \underline{A}, F, \mu, \nu)$ be a wta where $Q = \{q\}$, $\underline{A} = (\mathbb{N}, -, -, \Omega)$ with $\{\overline{\sigma} \mid \sigma \in \Sigma\} \cup \{\mathrm{id}\} \subseteq \Omega$ and for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $n_1, \ldots, n_k \in \mathbb{N}$, we define $\overline{\sigma}(n_1, \ldots, n_k) = 1 + \max\{n_1, \ldots, n_k\}$; in fact, the addition and the 0 of $\underline{A}$ are irrelevant. Moreover, $F_q = \mathrm{id}$ and for every $k \geq 0$ and $\sigma \in \Sigma^{(k)}$, let $\mu_k(\sigma)_{q\ldots q,q} = \overline{\sigma}$. We observe that M is deterministic. Clearly $S(M) = \mathrm{ht}$.

Example 3.11. Next let us construct a wta M which takes any tree $t \in T_\Sigma(Z)$ with $\Sigma = \{\sigma^{(2)}, \alpha^{(0)}\}$ and $Z = \{z\}$ and, provided that t contains exactly one occurrence of z and this labels the leaf on the zig-zag-path through t, $S(M)$ maps t to a unary language function f, such that $f(L) = \{w\}[z \leftarrow L]$ where w has the form of the zig-zag-path through t, and the deleted subtrees are represented by certain α-strings. For instance: for arbitrary trees $t_1, t_2, t_3 \in T_\Sigma$,

$$\big(S(M), \sigma(\sigma(t_1, \sigma(z, t_2)), t_3)\big) = \lambda x.1\,2\,\alpha^{n_1}\,1\,x\,\alpha^{n_2}\,\alpha^{n_3}$$

where n_i is the size of t_i for every $i \in \{1,2,3\}$. Then $\lambda x.1\,2\,\alpha^{n_1}\,1\,x\,\alpha^{n_2}\,\alpha^{n_3}$ is a unary language function of type $\mathcal{P}(\Delta^*) \to \mathcal{P}(\Delta^*)$ which takes any language $L \subseteq \Delta^*$ as argument and delivers the language $\{1\,2\,\alpha^{n_1}\,1\,w\,\alpha^{n_2}\,\alpha^{n_3} \mid w \in L\}$ as result. For this, we construct the wta $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ where:

- $Q = \{e, o, \star\}$ where e and o abbreviate even and odd, respectively, for counting modulo 2, and $\star$ is needed to accumulate the α-strings.

- $\underline{A} = (\mathcal{P}(\Delta^*), \cup, \emptyset, \Omega)$ where $\Delta = \{1, 2, \alpha\}$ and

$$\Omega = \{\mathrm{one}^{(2)}, \mathrm{two}^{(2)}, \mathrm{alpha}_2^{(2)}, \mathrm{alpha}_0^{(0)}, \mathrm{null}^{(0)}\} \cup \{\mathrm{id}\} \cup \{\emptyset^k \mid k \geq 0\}$$

  where for every $L_1, L_2 \in \mathcal{P}(\Delta^*)$ we define

$$\begin{aligned} \mathrm{one}(L_1, L_2) &= \{1\} \cdot L_1 \cdot L_2 & \qquad \mathrm{alpha}_0() &= \{\alpha\} \\ \mathrm{two}(L_1, L_2) &= \{2\} \cdot L_1 \cdot L_2 & \mathrm{null}() &= \{\varepsilon\} \\ \mathrm{alpha}_2(L_1, L_2) &= \{\alpha\} \cdot L_1 \cdot L_2. \end{aligned}$$

  Note that $\underline{A}$ is in fact a DM-monoid.

- $F_e = \mathrm{id}$ and $F_o = F_\star = \emptyset$.

- The transition mappings are given by

$$\begin{aligned} \mu_2(\sigma)_{o\star,e} &= \mathrm{one} & \qquad \mu_0(\alpha)_{\varepsilon,e} &= \mathrm{null} & \qquad \mu_2(\sigma)_{\star\star,\star} &= \mathrm{alpha}_2 \\ \mu_2(\sigma)_{\star e,o} &= \mathrm{two} & \mu_0(\alpha)_{\varepsilon,o} &= \mathrm{null} & \mu_0(\alpha)_{\varepsilon,\star} &= \mathrm{alpha}_0 \end{aligned}$$

  and every other entry in $\mu_2(\sigma)$ is mapped to $\emptyset$.

- $\nu(z)_e = \nu(z)_o = \mathrm{id}$ and $\nu(z)_\star = \emptyset$.

Now consider again the tree $t = \sigma(\sigma(t_1, \sigma(z, t_2)), t_3)$ and the run

$$r = \langle\sigma, e\rangle(\langle\sigma, o\rangle(t'_1, \langle\sigma, e\rangle(\langle z, o\rangle, t'_2)), t'_3)$$

in $R_M(t, e)$ where for every $i \in \{1, 2, 3\}$ the tree t'_i is obtained from t_i by adding the state $\star$ to every node symbol. It should be clear that $c_M(t'_i) = \{\alpha^{n_i}\}$ where n_i is the size of t_i. Then

$$c_M(r) = \mathrm{one}(\mathrm{two}(\{\alpha^{n_1}\}, \mathrm{one}(\mathrm{id}, \{\alpha^{n_2}\})), \{\alpha^{n_3}\})$$

Then, e.g., the subexpression $\mathrm{one}(\mathrm{id}, \{\alpha^{n_2}\})$ evaluates to the unary language function g such that for every language L we have:

$$g(L) = \mathrm{one}(\mathrm{id}(L), \{\alpha^{n_2}\}) = \mathrm{one}(L, \{\alpha^{n_2}\}) = \{1\, w\, \alpha^{n_2} \mid w \in L\}.$$

Thus $c_M(r)$ evaluates to the unary language function $\lambda x.1\, 2\, \alpha^{n_1}\, 1\, x\, \alpha^{n_2}\, \alpha^{n_3}$. Then

$$\begin{aligned} (S(M), t) &= \bigcup_{q \in Q} F_q((S(M)_q, t)) \\ &= F_e((S(M)_e, t)) \cup F_o((S(M)_o, t)) \cup F_\star((S(M)_\star, t)) \\ &= \mathrm{id}((S(M)_e, t)) \cup \emptyset^1((S(M)_o, t)) \cup \emptyset^1((S(M)_\star, t)) \\ &= (S(M)_e, t) \cup \emptyset^1 \cup \emptyset^1 = (S(M)_e, t). \end{aligned}$$

Also it is obvious that for every other e-run r' on t, $c_M(r') = \emptyset$. Thus we obtain that $(S(M), t) = \lambda x.1\, 2\, \alpha^{n_1}\, 1\, x\, \alpha^{n_2}\, \alpha^{n_3}$.

Example 3.12. In [17] it has been shown that wta (without variables) over particular DM-monoids exactly characterize the bottom-up tree series transducers [9, 11, 16, 18]. Roughly speaking, if a bottom-up tree series transducer is specified over some ranked input (output) alphabet Σ (Δ, respectively) and some semiring B, then the DM-monoid

$\underline{A}$ for the wta has the form $(B\langle\langle T_\Delta\rangle\rangle, \oplus, \widetilde{\mathbf{0}}, \Omega)$ where $\Omega = \Gamma \cup \{\mathrm{id}\} \cup \{\mathbf{0}^k \mid k \geq 0\}$ and $\Gamma^{(k)} = \{\overline{\varphi}_k \mid \varphi \in A\langle\langle T_\Delta(X_k)\rangle\rangle\}$. The k-ary operation $\overline{\varphi}_k : B\langle\langle T_\Delta\rangle\rangle^k \to B\langle\langle T_\Delta\rangle\rangle$ is defined by $\overline{\varphi}_k(\psi_1, \ldots, \psi_k) = \varphi \leftarrow (\psi_1, \ldots, \psi_k)$ where $\leftarrow$ is the IO-substitution of tree series [3, 9] (using the elements of $X_k = \{x_1, \ldots, x_k\}$ as substitution variables). For more details we refer to [17].

Finally, let us present an inductive way to compute the run semantics of a wta. For this, we need that the underlying M-monoid $\underline{A}$ is distributive (see Proposition 1 of [17]). In the sequel we will use the notation $M(t)_q$, which supports the inductive definition, as an abbreviation of $(S(M)_q, t)$ for every $t \in T_\Sigma(Z)$ and $q \in Q$.

Lemma 3.13. If $\underline{A}$ is distributive, then

(i) $M(z)_q = \nu(z)_q$ and

(ii) $M(\sigma(t_1, \ldots, t_k))_q = \bigoplus_{q_1,\ldots,q_k \in Q} \mu_k(\sigma)_{q_1\ldots q_k,q}(M(t_1)_{q_1}, \ldots, M(t_k)_{q_k})$

for every $q \in Q$, $z \in Z$, $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $t_1, \ldots, t_k \in T_\Sigma(Z)$.

Proof. Clearly, $M(z)_q = \bigoplus_{r \in R_M(z,q)} c_M(r) = c_M(\langle z, q\rangle) = \nu(z)_q$, which proves (i). For (ii), we calculate as follows where $t = \sigma(t_1, \ldots, t_k)$:

$$M(\sigma(t_1, \ldots, t_k))_q = \bigoplus_{r \in R_M(t,q)} c_M(r)$$

$=$ (since $R_M(t,q) = \{\langle\sigma, q\rangle(r_1, \ldots, r_k) \mid r_1 \in R_M(t_1), \ldots, r_k \in R_M(t_k)\}$)

$$\cdots \bigoplus_{r_i \in R_M(t_i)} \cdots \ c_M(\langle\sigma, q\rangle(r_1, \ldots, r_k))$$

$=$ (by definition of $R_M(t_i)$ and c_M)

$$\bigoplus_{q_1,\ldots,q_k \in Q} \cdots \bigoplus_{r_i \in R_M(t_i,q_i)} \cdots \ \mu_k(\sigma)_{q_1\ldots q_k,q}(\ldots, c_M(r_i), \ldots)$$

$=$ (by distributivity; see Observation 3.3)

$$\bigoplus_{q_1,\ldots,q_k \in Q} \mu_k(\sigma)_{q_1\ldots q_k,q}\Big(\ldots, \bigoplus_{r_i \in R_M(t_i,q_i)} c_M(r_i), \ldots\Big)$$

$=$ (by definition of $S(M)_q$)

$$\bigoplus_{q_1,\ldots,q_k \in Q} \mu_k(\sigma)_{q_1\ldots q_k,q}(M(t_1)_{q_1}, \ldots, M(t_k)_{q_k}).$$

□

We note that Lemma 3.13 presents an inductive definition of $M(t)_q$. We will use this inductive definition in several of the following proofs.

4 Normal forms for wta

In this section we will present certain normal forms of wta with respect to the variables or the final distribution function. Moreover, we show constructions that normalize wta provided that the underlying M-monoid has certain properties. To this end, let $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ be an arbitrary wta with $\underline{A} = (A, \oplus, \mathbf{0}, \Omega)$ for the rest of this section.

First we consider a normal form which concerns the variables. More specifically, for a variable x of a subset $X \subseteq Z$, there shall exist a state q in the wta such that there exists an x-transition with weight id into q and for all other states there shall not be an x-transition. Clearly, such a state q is unique, whenever it exists. Consequently we call such a state q *the x-initial state*. If, moreover, there shall not be a z-transition into q for any $z \in X$ with $z \neq x$, then we call it *the X-private x-initial state*. A Z-private x-initial state is just called *private x-initial state*.

The concept of X-private x-initial state with $X \neq Z$ will only be considered in Section 6.2. Another important concept will be *variable states*. A state $q \in Q$ is a variable state, whenever it is not reachable in μ; i.e., there does not exist a transition in μ leading into q. Hence there may only be z-transitions (of ν) that lead to q. Then we call a state an X-private x-initial variable state if it is both, a variable state and X-private x-initial. A wta that has a private x-initial variable state is called initial x-state normalized in Definition 4.10 of [7].

Definition 4.1. A state $q \in Q$ is a *variable state*, if $\mu_k(\sigma)_{q_1 \ldots q_k, q} = \mathbf{0}$ for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q_1, \ldots, q_k \in Q$.

Clearly a variable state cannot accept any tree that contains any symbol of Σ. This is formalized in the next observation.

Observation 4.2. Let $q \in Q$ be a variable state. Then $\mathrm{supp}(S(M)_q) \subseteq Z$. Moreover, if q is the private x-initial variable state for some $x \in Z$, then $S(M)_q = \mathrm{id}.x$ and $x \notin \mathrm{supp}(S(M)_p)$ for every $p \in Q$ with $p \neq q$.

For the results in the sequel we need the following properties of M-monoids.

Definition 4.3. Let $\Phi \subseteq \mathrm{Ops}(A)$. We say that Φ is

- *sum closed*, if $\omega_1 \oplus \omega_2 \in \Phi^{(k)}$ for every $k \geq 0$ and $\omega_1, \omega_2 \in \Phi^{(k)}$.
- $(1, \star)$-*composition closed*, if $\omega(\omega') \in \Phi^{(k)}$ for every $k \geq 0$, $\omega \in \Phi^{(1)}$, and $\omega' \in \Phi^{(k)}$.
- $(\star, 1)$-*composition closed*, if $\omega(\omega_1, \ldots, \omega_k) \in \Phi^{(k)}$ for every $k \geq 0$, $\omega \in \Phi^{(k)}$, and $\omega_1, \ldots, \omega_k \in \Phi^{(1)}$.
- *unary-composition closed*, if Φ is $(\star, 1)$- and $(1, \star)$-composition closed.

We say that $\underline{A}$ is sum closed (respectively, $(1, \star)$-composition closed, $(\star, 1)$-composition closed, and unary-composition closed), if Ω is so.

Let us show that for every distributive $\underline{A}$ we can construct a DM-monoid $(A, \oplus, \mathbf{0}, \Omega')$ with $\Omega \subseteq \Omega'$ that is sum closed and unary-composition closed.

Lemma 4.4. Let $\underline{A}$ be distributive, and let $\underline{B} = (A, \oplus, \mathbf{0}, \Omega')$ be the M-monoid such that Ω' is the smallest sum closed and unary-composition closed subset of $\mathrm{Ops}(A)$ that contains Ω. Then $\underline{B}$ is a sum closed and unary-composition closed DM-monoid.

Proof. By definition, $\underline{B}$ is sum closed and unary-composition closed. It remains to prove that distributivity is preserved. For this it suffices to show that the set of distributive operations on A is sum closed and unary-composition closed. This is straightforward and hence omitted. ☐

Next we show that, under appropriate conditions on $\underline{A}$, there exists an equivalent wta M' with a private z-initial variable state ($z \in Z$).

Lemma 4.5. Let $\underline{A}$ be a $(\star, 1)$-composition closed and sum closed DM-monoid, and let $z \in Z$. There exists a wta M' with a private z-initial variable state such that $S(M') = S(M)$.

Proof. Without loss of generality, let us suppose that $z \notin Q$. Define the wta $M' = (Q', \Sigma, Z, \underline{A}, F', \mu', \nu')$ such that:

- $Q' = Q \cup \{z\}$;
- $F'_z = \bigoplus_{q \in Q} F_q(\nu(z)_q)$ and $F'_q = F_q$ for every $q \in Q$;
- $\nu'(z)_z = \mathrm{id}$ and $\nu'(z)_q = \mathbf{0}$ for every $q \in Q$;
- $\nu'(x)_z = \mathbf{0}$ and $\nu'(x)_q = \nu(x)_q$ for every $x \in Z$ with $x \neq z$ and $q \in Q$;
- for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, $q_1, \ldots, q_k \in Q'$, and $q \in Q$ let

$$\mu'_k(\sigma)_{q_1 \ldots q_k, q} = \bigoplus_{\substack{p_1, \ldots, p_k \in Q, \\ (\forall i \in [k]) \colon p_i = q_i \text{ if } q_i \neq z}} \mu_k(\sigma)_{p_1 \ldots p_k, q}(f_{p_1, q_1}, \ldots, f_{p_k, q_k})$$

  where for every $p \in Q$ and $q \in Q'$

$$f_{p,q} = \begin{cases} \nu(z)_p & \text{if } q = z, \\ \mathrm{id} & \text{otherwise} \end{cases}$$

- for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q_1, \ldots, q_k \in Q'$ let $\mu'_k(\sigma)_{q_1 \cdots q_k, z} = \mathbf{0}$.

We note that M' is well defined because of the given properties of $\underline{A}$. Obviously, z is a variable state in M', and z is the private z-initial state. It remains to show that the wta M and M' are equivalent; i.e., $S(M') = S(M)$. Since $\underline{A}$ is a DM-monoid, we use the inductive semantics given in Lemma 3.13 for the proof.

By Observation 4.2 we have that $S(M')_z = \mathrm{id}.z$ and $z \notin \mathrm{supp}(S(M')_q)$ for every $q \in Q$, and thus (i) $M'(t)_z = \mathbf{0}$ for every $t \in T_\Sigma(Z)$ with $t \neq z$, (ii) $M'(z)_z = \mathrm{id}$, and (iii) $M'(z)_q = \mathbf{0}$ for every $q \in Q$. Using these facts we now prove that $M'(t)_q = M(t)_q$ for every $t \in T_\Sigma(Z)$ with $t \neq z$ and $q \in Q$. We achieve this by induction on t.

For $t = x$ with $x \in Z$ and $x \neq z$, we have $M'(x)_q = \nu'(x)_q = \nu(x)_q = M(x)_q$ by Lemma 3.13. In the induction step suppose that $t = \sigma(t_1, \ldots, t_k)$ for some $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and trees $t_1, \ldots, t_k \in T_\Sigma(Z)$. By the induction hypothesis we have $M'(t_i)_q = M(t_i)_q$ for every $q \in Q$ and $i \in [k]$ such that $t_i \neq z$. Then

$$\begin{aligned}
& M'(\sigma(t_1, \ldots, t_k))_q \\
= \quad & \text{(by Lemma 3.13)} \\
& \bigoplus_{q_1, \ldots, q_k \in Q'} \mu'_k(\sigma)_{q_1 \ldots q_k, q}(M'(t_1)_{q_1}, \ldots, M'(t_k)_{q_k}) \\
= \quad & \text{(by definition of } \mu'\text{)} \\
& \bigoplus_{q_1, \ldots, q_k \in Q'} \Big(\bigoplus_{\substack{p_1, \ldots, p_k \in Q, \\ (\forall i \in [k]) \colon p_i = q_i \text{ if } q_i \neq z}} \mu_k(\sigma)_{p_1 \ldots p_k, q}(f_{p_1, q_1}, \ldots, f_{p_k, q_k})\big(M'(t_1)_{q_1}, \ldots, M'(t_k)_{q_k}\big) \Big) \\
= \quad & \text{(by Observation 3.4)} \\
& \bigoplus_{q_1, \ldots, q_k \in Q'} \Big(\bigoplus_{\substack{p_1, \ldots, p_k \in Q, \\ (\forall i \in [k]) \colon p_i = q_i \text{ if } q_i \neq z}} \mu_k(\sigma)_{p_1 \ldots p_k, q}(f_{p_1, q_1}(M'(t_1)_{q_1}), \ldots, f_{p_k, q_k}(M'(t_k)_{q_k})) \Big)
\end{aligned}$$

$$
\begin{aligned}
= \quad & \text{(because } M'(t_i)_z = \mathbf{0} \text{ and } M'(z)_{q_i} = \mathbf{0} \text{ for } t_i \neq z \text{ and } q_i \neq z \text{ by (i) and (iii))} \\
& \bigoplus_{\substack{q_1,\dots,q_k \in Q', \\ (\forall i \in [k])\colon q_i = z \text{ iff } t_i = z}} \Bigl(\bigoplus_{\substack{p_1,\dots,p_k \in Q, \\ (\forall i \in [k])\colon p_i = q_i \text{ if } q_i \neq z}} \mu_k(\sigma)_{p_1\dots p_k,q}(f_{p_1,q_1}(M'(t_1)_{q_1}),\dots,f_{p_k,q_k}(M'(t_k)_{q_k}))\Bigr) \\
= \quad & \text{(by evaluation of } f_{p_i,q_i}(M'(t_i)_{q_i}) \text{ using (ii))} \\
& \bigoplus_{\substack{q_1,\dots,q_k \in Q', \\ (\forall i \in [k])\colon q_i = z \text{ iff } t_i = z}} \Bigl(\bigoplus_{\substack{p_1,\dots,p_k \in Q, \\ (\forall i \in [k])\colon p_i = q_i \text{ if } q_i \neq z}} \mu_k(\sigma)_{p_1\dots p_k,q}(g(p_1,q_1,t_1),\dots,g(p_k,q_k,t_k))\Bigr)
\end{aligned}
$$

where for every $p \in Q$, $q' \in Q'$, and $t \in T_\Sigma(Z)$, we define $g(p,q',t) = \nu(z)_p$ if $t = z$, and $M'(t)_p$ otherwise. Thus by induction hypothesis and Lemma 3.13 we have that $g(p,q',t_i) = M(t_i)_p$. Hence we continue with

$$
\begin{aligned}
& \bigoplus_{\substack{q_1,\dots,q_k \in Q', \\ (\forall i \in [k])\colon q_i = z \text{ iff } t_i = z}} \Bigl(\bigoplus_{\substack{p_1,\dots,p_k \in Q, \\ (\forall i \in [k])\colon p_i = q_i \text{ if } q_i \neq z}} \mu_k(\sigma)_{p_1\dots p_k,q}(M(t_1)_{p_1},\dots,M(t_k)_{p_k})\Bigr) \\
= & \bigoplus_{p_1,\dots,p_k \in Q} \mu_k(\sigma)_{p_1\dots p_k,q}(M(t_1)_{p_1},\dots,M(t_k)_{p_k}) \\
= & \quad \text{(by Lemma 3.13)} \\
& M(\sigma(t_1,\dots,t_k))_q.
\end{aligned}
$$

Note that we used the absorption property freely. Now we can finish the proof. For every $t \in T_\Sigma(Z)$ with $t \neq z$ we immediately obtain $(S(M'),t) = (S(M),t)$. For $t = z$ we can compute as follows:

$$
(S(M'),z) = \bigoplus_{q \in Q'} F'_q(M'(z)_q) = F'_z(\mathrm{id}) = F'_z = \bigoplus_{q \in Q} F_q(\nu(z)_q) = (S(M),z),
$$

where we used (ii) and (iii) in the second step and the definition of F'_z in the fourth step. Hence we proved that $S(M') = S(M)$. □

So far, we have moved weights from the initial transitions at variables towards inner transitions. In essence we moved the weight upwards away from the leaf. Now we turn to the final distribution and will move weights of it down to the last transition (at the root) of the input tree.

Definition 4.6. A state $p \in Q$ is a *terminating state*, if $F_p = \mathrm{id}$, $F_q = 0$ for every $q \in Q$ with $q \neq p$, and, for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q, q_1, \dots, q_k \in Q$ we have that $\mu_k(\sigma)_{q_1\dots q_k,q} = \mathbf{0}$ whenever there exists $1 \leq i \leq k$ such that $q_i = p$. Clearly, such a state is unique if it exists.

First we observe that the semantics of a wta with a terminating state can be simplified because the final distribution entry is either $\mathbf{0}$ (non-terminating state) or id (the terminating state). This simplification is stated in the next observation.

Observation 4.7. If $p \in Q$ is a terminating state, then $S(M) = S(M)_p$.

Next we show that there exists an equivalent wta M' with a terminating state. For this we need that $\underline{A}$ is a $(1,\star)$-composition closed and sum closed DM-monoid. The construction pushes the final weight down to the last transition, and nondeterminism is used to guess whether the next transition is the last transition or not.

Lemma 4.8 (cf. Lemma 4.8 of [7] and Lemma 22 of [2]). Let $\underline{A}$ be a $(1, \star)$-composition closed and sum closed DM-monoid. There exists a wta M' with a terminating state such that $S(M') = S(M)$.

Proof. Let $p \notin Q$ be a new state. The wta $M' = (Q', \Sigma, Z, \underline{A}, F', \mu', \nu')$ is constructed as follows.

- $Q' = Q \cup \{p\}$;
- $F'_q = \mathbf{0}$ for every $q \in Q$ and $F'_p = \mathrm{id}$;
- $\nu'(z)_q = \nu(z)_q$ and $\nu'(z)_p = \bigoplus_{q' \in Q} F_{q'}(\nu(z)_{q'})$ for every $z \in Z$ and $q \in Q$; and
- $\mu'_k(\sigma)_{q_1 \ldots q_k, q} = \mu_k(\sigma)_{q_1 \ldots q_k, q}$ and $\mu'_k(\sigma)_{q_1 \ldots q_k, p} = \bigoplus_{q' \in Q} F_{q'}(\mu_k(\sigma)_{q_1 \ldots q_k, q'})$ for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q, q_1, \ldots, q_k \in Q$.
- All remaining entries in μ' are $\mathbf{0}$.

It is immediately clear that p is a terminating state. We leave it as an exercise to the reader to prove the statement $M'(t)_q = M(t)_q$ for every $t \in T_\Sigma(Z)$ and $q \in Q$ by a straightforward induction on t (see Lemma 3.13). With this statement, it is easy to see that $M'(t)_p = \bigoplus_{q' \in Q} F_{q'}(M(t)_{q'})$, which yields the desired $S(M') = S(M)$. $\square$

Finally, let us consider the combination of a private z-initial variable state and a terminating state. Recall that z-initial variable states required us to move weights upward (away from the variable leaves) and the terminating state requires us to move weights downward. In case the input tree is just a variable, there exists no transition such that we can perform those two moves consistently. However, for every uniform tree valuation $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ such that $(\psi, z) = \mathbf{0}$ this approach is possible.

Definition 4.9. Let $z \in Z$. A uniform tree valuation $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ is *z-proper* if $(\psi, z) = \mathbf{0}$.

Lemma 4.10. Let $\underline{A}$ be a unary-composition closed and sum closed DM-monoid. Moreover, let $z \in Z$ and $\psi \in \mathrm{Rec}(\Sigma, Z, \underline{A})$ be z-proper. Then there exists a wta that

- recognizes ψ;
- has a private z-initial variable state; and
- has a terminating state.

Proof. Let $M' = (Q', \Sigma, Z, \underline{A}, F', \mu', \nu')$ be a wta with a private z-initial variable state q such that $S(M') = \psi$. Such a wta exists by Lemma 4.5. By Observation 4.2, $(S(M'), z) = F'_q(\nu'(z)_q) = F'_q$. However, $(\psi, z) = \mathbf{0}$ and thus $F'_q = \mathbf{0}$. Let M'' be the wta with terminating state that results from the application of the construction of Lemma 4.8 to M'. Because $F'_q = \mathbf{0}$ we observe that q is still a private z-initial variable state. Moreover, $S(M'') = S(M') = \psi$. $\square$

5 Rational operations and rational expressions

In this section we introduce the rational operations (and correspondingly the rational expressions) on uniform tree valuations over M-monoids.

Before we are ready to define the rational operations, we need one more composition operation. More specifically, we need a composition that composes a function ω with several other functions $\omega_1, \ldots, \omega_k$ such that the result of the operation ω_i is forwarded to a certain parameter of ω. The selection of the exact positions is driven by two subsets of $\mathbb{N}^*$. The first gives references to all parameters and the second subset singles out the parameters at which the composition should take place.

Definition 5.1. Let $k, n \geq 0$ and let $W, V \subseteq \mathbb{N}^*$ such that $V \subseteq W$, $\mathrm{card}(W) = n$, and $\mathrm{card}(V) = k$. Suppose that $W = \{w_1, \ldots, w_n\}$ and $V = \{v_1, \ldots, v_k\}$ such that $w_1 <_{\mathrm{lex}} \cdots <_{\mathrm{lex}} w_n$ and $v_1 <_{\mathrm{lex}} \cdots <_{\mathrm{lex}} v_k$. We define the mapping

$$\circ_{W,V} \colon \mathrm{Ops}^n(A) \times \mathrm{Ops}(A)^k \to \mathrm{Ops}(A)$$

for every $\omega \in \mathrm{Ops}^n(A)$ and $\omega_1, \ldots, \omega_k \in \mathrm{Ops}(A)$ by $\omega \circ_{W,V} (\omega_1, \ldots, \omega_k) = \omega(\omega'_1, \ldots, \omega'_n)$ where for every $1 \leq i \leq n$

$$\omega'_i = \begin{cases} \omega_j & \text{if } w_i = v_j \text{ for some } 1 \leq j \leq k, \\ \mathrm{id} & \text{otherwise.} \end{cases}$$

In the sequel, we will mostly use $\circ_{\mathrm{pos}_Z(t), \mathrm{pos}_z(t)}$ for some $t \in T_\Sigma(Z)$ and $z \in Z$. Thus we abbreviate it to simply $\circ_{t,z}$. Before we define the rational operations, let us prove the associativity of $\circ_{W,V}$.

Proposition 5.2. Let $m \geq 0$, $W_i \subseteq W$, and $V_i \subseteq V$ such that $V_i \subseteq W_i$, $n_i = \mathrm{card}(W_i)$, and $k_i = \mathrm{card}(V_i)$ for every $1 \leq i \leq m$. Finally, let $\omega \in \mathrm{Ops}^m(A)$, $\omega_i \in \mathrm{Ops}^{n_i}(A)$ for every $1 \leq i \leq m$, and $\omega_{i,j} \in \mathrm{Ops}(A)$ for every $1 \leq i \leq m$ and $1 \leq j \leq k_i$. Provided that $W_1 <_{\mathrm{lex}} \cdots <_{\mathrm{lex}} W_k$,

$$\begin{aligned} &\omega\big(\omega_1 \circ_{W_1,V_1} (\omega_{1,1}, \ldots, \omega_{1,k_1}), \ldots, \omega_m \circ_{W_m,V_m} (\omega_{m,1}, \ldots, \omega_{m,k_m})\big) \\ &= \omega(\omega_1, \ldots, \omega_m) \circ_{W,V} (\omega_{1,1}, \ldots, \omega_{1,k_1}, \ldots, \omega_{m,1}, \ldots, \omega_{m,k_m}) \end{aligned}$$

where $W = \bigcup_{i=1}^m W_i$ and $V = \bigcup_{i=1}^m V_i$.

Proof. Suppose that $W_i = \{w_{i,1}, \ldots, w_{i,n_i}\}$ and $V_i = \{v_{i,1}, \ldots, v_{i,k_i}\}$ such that we have $w_{i,1} <_{\mathrm{lex}} \cdots <_{\mathrm{lex}} w_{i,n_i}$ and $v_{i,1} <_{\mathrm{lex}} \cdots <_{\mathrm{lex}} v_{i,k_i}$ for every $1 \leq i \leq k$. For every $1 \leq i \leq m$ and $1 \leq j \leq n_i$ let

$$\omega'_{i,j} = \begin{cases} \omega_{i,j'} & \text{if } w_{i,j} = v_{i,j'} \text{ for some } 1 \leq j' \leq k_i, \\ \mathrm{id} & \text{otherwise.} \end{cases}$$

Then we obtain

$$\begin{aligned} &\omega\big(\omega_1 \circ_{W_1,V_1} (\omega_{1,1}, \ldots, \omega_{1,k_1}), \ldots, \omega_m \circ_{W_m,V_m} (\omega_{m,1}, \ldots, \omega_{m,k_m})\big) \\ &= \omega\big(\omega_1(\omega'_{1,1}, \ldots, \omega'_{1,n_1}), \ldots, \omega_m(\omega'_{m,1}, \ldots, \omega'_{m,n_m})\big) \\ &= \big(\omega(\omega_1, \ldots, \omega_m)\big)(\omega'_{1,1}, \ldots, \omega'_{1,n_1}, \ldots, \omega'_{m,1}, \ldots, \omega'_{m,n_m}) \\ &= \omega(\omega_1, \ldots, \omega_m) \circ_{W,V} (\omega_{1,1}, \ldots, \omega_{1,k_1}, \ldots, \omega_{m,1}, \ldots, \omega_{m,k_m}) \end{aligned}$$

by definition of $\circ_{W_i,V_i}$, Observation 3.4, and the definition of $\circ_{W,V}$, respectively. $\square$

Next we define four kinds of operations, called *(complex) rational operations*, on $\mathrm{Uvals}(\Sigma, Z, \underline{A})$. In fact, we will not prove that the result of the operations applied to uniform tree valuations is again a uniform tree valuation because this is obvious.

Definition 5.3. We define the following *rational operations* on $\mathrm{Uvals}(\Sigma, Z, \underline{A})$.

1. The *sum* $\oplus^{\mathrm{u}}$ is a rational operation (for the definition, cf. Section 3.1).
2. For every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $\omega \in \Omega^{(k)}$, the *top-concatenation (with σ)* $\mathrm{top}_{\sigma,\omega}$ is rational. For all uniform tree valuations $\psi_1, \dots, \psi_k \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$, the uniform tree valuation $\mathrm{top}_{\sigma,\omega}(\psi_1, \dots, \psi_k)$ is defined by
$$\mathrm{top}_{\sigma,\omega}(\psi_1, \dots, \psi_k) = \bigoplus^{\mathrm{u}}_{t_1,\dots,t_k \in T_\Sigma(Z)} \omega((\psi_1, t_1), \dots, (\psi_k, t_k)).\sigma(t_1, \dots, t_k).$$
3. For every $z \in Z$ the *z-concatenation* $\cdot_z$ is rational. For every $\psi, \psi' \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ the z-concatenation of ψ and ψ' is the uniform tree valuation which is defined by
$$\psi \cdot_z \psi' = \bigoplus^{\mathrm{u}}_{\substack{s \in T_\Sigma(Z), l = |s|_z \\ t_1,\dots,t_l \in T_\Sigma(Z)}} \Big((\psi, s) \circ_{s,z} ((\psi', t_1), \dots, (\psi', t_l))\Big).s[z \leftarrow (t_1, \dots, t_l)] \ .$$
4. For every $z \in Z$ the z-Kleene-*star* is rational. For the definition of it, we need the iteration first. For every $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ and $n \geq 0$ we define the uniform tree valuation ψ_z^n inductively over n as follows:
   (i) $\psi_z^0 = \widetilde{\mathbf{0}}$; and
   (ii) $\psi_z^{n+1} = (\psi \cdot_z \psi_z^n) \oplus^{\mathrm{u}} \mathrm{id}.z$.

   The z-Kleene star of a z-proper $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ is the uniform tree valuation, denoted by ψ_z^*, that is defined by $(\psi_z^*, t) = (\psi_z^{\mathrm{ht}(t)+1}, t)$ for every $t \in T_\Sigma(Z)$. For every non-z-proper $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ we define $\psi_z^* = \widetilde{\mathbf{0}}$.

We note that the top-concatenation operation can be considered basic (if $k = 0$) or complex (if $k \geq 1$). In the basic case top-concatenation yields the series $\omega.\alpha$ where $\omega \in \Omega^{(0)}$ and $\alpha \in \Sigma^{(0)}$. In essence, our top-concatenation is different from the one of [7] in the fact that we not only concatenate a symbol from Σ and accumulate the weights, but apply a k-ary operation as well. In this sense, it can be seen as a combination of the top-concatenation of [7] immediately followed by a scalar multiplication with the corresponding transition weight. In [7] these operations could be applied independently, because the weights could be permuted due to the required commutativity of the underlying semiring.

The following lemma justifies the definition of the z-Kleene-star in Definition 5.3.

Lemma 5.4 (cf. [1] and Lemma 3.10 of [7])**.** Let $z \in Z$, $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ be z-proper, and $t \in T_\Sigma(Z)$. If $n \geq \mathrm{ht}(t) + 1$, then $(\psi_z^{n+1}, t) = (\psi_z^n, t)$.

Proof. We prove the statement by induction on $\mathrm{ht}(t)$.

Induction base: Let $\mathrm{ht}(t) = 0$. First let us consider the case $t = z$. We now show that, for every $n \geq 1$, $(\psi_z^n, z) = \mathrm{id}$. In fact,

$$\begin{aligned} &(\psi_z^n, z) = (\psi \cdot_z \psi_z^{n-1}, z) \oplus (\mathrm{id}.z, z) \\ &= \big((\psi, z) \circ_{\{\varepsilon\},\{\varepsilon\}} (\psi_z^{n-1}, z)\big) \oplus (\mathrm{id}.z, z) = \mathbf{0}^1((\psi_z^{n-1}, z)) \oplus (\mathrm{id}.z, z) \end{aligned}$$

$$= \mathbf{0} \oplus \mathrm{id} \qquad\qquad\qquad\qquad = \mathrm{id}\,.$$

Secondly, we consider the case that $t \neq z$. We show that, for every $n \geq 1$, we have $(\psi_z^n, t) = (\psi, t)$. In fact, $(\psi_z^n, t) = (\psi \cdot_z \psi_z^{n-1}, t) \oplus (\mathrm{id}.z, t)$, which equals

$$\big((\psi, t) \circ_{\mathrm{pos}_Z(t),\emptyset} ()\big) \oplus \big((\psi, z) \circ_{\{\varepsilon\},\{\varepsilon\}} (\psi_z^{n-1}, t)\big) \oplus (\mathrm{id}.z, t),$$

which is equal to $(\psi, t) \oplus \mathbf{0}^1((\psi_z^{n-1}, t)) \oplus \mathbf{0} = (\psi, t) \oplus \mathbf{0} \oplus \mathbf{0} = (\psi, t)$.

Induction step: Let $\mathrm{ht}(t) > 0$ and $n \geq \mathrm{ht}(t)+1$. By the induction hypothesis the statement holds for every tree t' such that $\mathrm{ht}(t') < \mathrm{ht}(t)$. Then $(\psi_z^{n+1}, t) = (\psi \cdot_z \psi_z^n, t) \oplus (\mathrm{id}.z, t)$. This is equal to $(\psi \cdot_z \psi_z^n, t)$ because $t \neq z$. By definition of $\cdot_z$ this equals

$$\bigoplus_E (\psi, s) \circ_{s,z} ((\psi_z^n, t_1), \ldots, (\psi_z^n, t_l))$$

where E abbreviates $\{(s, t_1, \ldots, t_l) \in T_\Sigma(Z)^{l+1} \mid s \neq z, l = |s|_z, t = s[z \leftarrow (t_1, \ldots, t_l)]\}$. Note that we can take $s \neq z$ because for $s = z$ we have $(\psi, s) = \mathbf{0}$. Now, since $s \neq z$, we have that $\mathrm{ht}(t_i) < \mathrm{ht}(t)$ and so $n - 1 \geq \mathrm{ht}(t_i) + 1$. Hence $(\psi_z^n, t_i) = (\psi_z^{n-1}, t_i)$ by the induction hypothesis. Obviously, substituting this in the above expression we obtain (ψ_z^n, t) by the same computation as above (for n instead of $n+1$). □

This yields the following important property concerning z-Kleene-stars.

Lemma 5.5. Let $z \in Z$, and let $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ be z-proper. Then $\psi_z^* = (\psi \cdot_z \psi_z^*) \oplus^{\mathrm{u}} \mathrm{id}.z$.

Proof. Let $t \in T_\Sigma(Z)$. Then $(\psi_z^*, t) = (\psi_z^{\mathrm{ht}(t)+1}, t) = (\psi_z^{\mathrm{ht}(t)+2}, t)$ by Lemma 5.4. By definition, this is equal to $((\psi \cdot_z \psi_z^{\mathrm{ht}(t)+1}) \oplus^{\mathrm{u}} \mathrm{id}.z, t) = (\psi \cdot_z \psi_z^{\mathrm{ht}(t)+1}, t) \oplus (\mathrm{id}.z, t)$. Using the definition of $\cdot_z$, we obtain

$$(\psi \cdot_z \psi_z^{\mathrm{ht}(t)+1}, t) = \bigoplus_E (\psi, s) \circ_{s,z} ((\psi_z^{\mathrm{ht}(t)+1}, t_1), \ldots, (\psi_z^{\mathrm{ht}(t)+1}, t_l))$$

where E abbreviates $\{(s, t_1, \ldots, t_l) \in T_\Sigma(Z)^{l+1} \mid l = |s|_z, t = s[z \leftarrow (t_1, \ldots, t_l)]\}$. Since $\mathrm{ht}(t) + 1 \geq \mathrm{ht}(t_i) + 1$, it follows that $(\psi_z^{\mathrm{ht}(t)+1}, t_i) = (\psi_z^{\mathrm{ht}(t_i)+1}, t_i) = (\psi_z^*, t_i)$. Thus we obtain

$$(\psi \cdot_z \psi_z^{\mathrm{ht}(t)+1}, t) = \bigoplus_E (\psi, s) \circ_{s,z} ((\psi_z^*, t_1), \ldots, (\psi_z^*, t_l)) = (\psi \cdot_z \psi_z^*, t).$$

Consequently, $(\psi_z^*, t) = (\psi \cdot_z \psi_z^{\mathrm{ht}(t)+1}, t) \oplus (\mathrm{id}.z, t) = (\psi \cdot_z \psi_z^*, t) \oplus (\mathrm{id}.z, t)$, which proves the statement. □

After we have defined the rational operations and proved some essential properties of the z-Kleene-star, we are now ready to define rational expressions and their semantics.

Definition 5.6. The set $\mathrm{RatExp}(\Sigma, Z, \underline{A})$ of *rational expressions* (over Σ, Z, and $\underline{A}$) is defined inductively as the smallest set R satisfying Conditions (i)–(v). For every rational expression $\eta \in \mathrm{RatExp}(\Sigma, Z, \underline{A})$ we define its semantics $[\![\eta]\!] \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ simultaneously.

(i) For every $z \in Z$ and $\omega \in \Omega^{(1)}$ we have $\omega.z \in R$ and $[\![\omega.z]\!] = \omega.z$.

(ii) For every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, $\omega \in \Omega^{(k)}$, and rational expressions $\eta_1, \ldots, \eta_k \in R$ we have $\mathrm{top}_{\sigma,\omega}(\eta_1, \ldots, \eta_k) \in R$ and $[\![\mathrm{top}_{\sigma,\omega}(\eta_1, \ldots, \eta_k)]\!] = \mathrm{top}_{\sigma,\omega}([\![\eta_1]\!], \ldots, [\![\eta_k]\!])$.

(iii) For every $\eta_1, \eta_2 \in R$ we have $\eta_1 + \eta_2 \in R$ and $[\![\eta_1 + \eta_2]\!] = [\![\eta_1]\!] \oplus^{\mathrm{u}} [\![\eta_2]\!]$.

(iv) For every $\eta_1, \eta_2 \in R$ and $z \in Z$ we have $\eta_1 \cdot_z \eta_2 \in R$ and $[\![\eta_1 \cdot_z \eta_2]\!] = [\![\eta_1]\!] \cdot_z [\![\eta_2]\!]$.

(v) For every $\eta \in R$ and $z \in Z$ we have $\eta_z^* \in R$ and $[\![\eta_z^*]\!] = [\![\eta]\!]_z^*$.

Definition 5.7. We call $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ *rational*, if there exists a rational expression $\eta \in \mathrm{RatExp}(\Sigma, Z, \underline{A})$ such that $[\![\eta]\!] = \psi$. The set of all rational uniform tree valuations of $\mathrm{Uvals}(\Sigma, Z, \underline{A})$ is denoted by $\mathrm{Rat}(\Sigma, Z, \underline{A})$. Moreover, we abbreviate $\bigcup_{Z \text{ finite set}} \mathrm{Rat}(\Sigma, Z, \underline{A})$ by $\mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A})$.

Let us conclude this section by a small example and an easy property.

Example 5.8. Let $\Sigma = \{\sigma^{(2)}, \alpha^{(0)}\}$ and $\underline{A} = (\mathbb{N}, -, -, \Omega)$ be the M-monoid with $\Omega = \{\overline{\sigma}^{(2)}, \overline{\alpha}^{(0)}, \mathrm{id}\}$ and for every $n_1, n_2 \in \mathbb{N}$, we define $\overline{\sigma}(n_1, n_2) = 1 + \max\{n_1, n_2\}$ and $\overline{\alpha}() = 0$. Moreover, let

$$\eta = \mathrm{top}_{\sigma,\overline{\sigma}}(\mathrm{id}.z, \mathrm{id}.z) + \mathrm{top}_{\alpha,\overline{\alpha}}() \quad \text{and} \quad \eta' = \eta_z^*.$$

Clearly, $\eta' \in \mathrm{RatExp}(\Sigma, \{z\}, \underline{A})$. For the sake of illustration, let us prove that $[\![\eta']\!]|_{T_\Sigma} = \mathrm{ht}$. For $t = \alpha$, we have $([\![\eta']\!], \alpha) = ([\![\eta]\!]_z^*, \alpha) = ([\![\eta]\!]_z^1, \alpha) = \overline{\alpha} = 0 = \mathrm{ht}(\alpha)$. Next, let $t = \sigma(t_1, t_2)$ for some $t_1, t_2 \in T_\Sigma$. By induction hypothesis, $([\![\eta']\!], t_1) = \mathrm{ht}(t_1)$ and $([\![\eta']\!], t_2) = \mathrm{ht}(t_2)$. Then

$$\begin{aligned}
&\big([\![\eta']\!], \sigma(t_1, t_2)\big) \;=\; \big([\![\eta]\!]_z^*, \sigma(t_1, t_2)\big) \\
=&\qquad \text{(by Lemma 5.5)} \\
&\big(\min{}^{\mathrm{u}}(\min{}^{\mathrm{u}}(\overline{\sigma}.\sigma(z,z), \overline{\alpha}.\alpha) \cdot_z [\![\eta]\!]_z^*, \mathrm{id}.z), \sigma(t_1, t_2)\big) \\
=&\; \overline{\sigma}(([\![\eta]\!]_z^*, t_1), ([\![\eta]\!]_z^*, t_2)) \;=\; \overline{\sigma}(([\![\eta']\!], t_1), ([\![\eta']\!], t_2)) \\
=&\qquad \text{(by induction hypothesis)} \\
&\overline{\sigma}(\mathrm{ht}(t_1), \mathrm{ht}(t_2)) \;=\; 1 + \max\{\mathrm{ht}(t_1), \mathrm{ht}(t_2)\} \\
=&\; \mathrm{ht}(\sigma(t_1, t_2)).
\end{aligned}$$

Finally, we present an easy observation that will prove useful in the proof of Theorem 6.8. For a unary operation $\omega \in \Omega^{(1)}$ and a uniform tree valuation $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ let us define the uniform tree valuation $\omega \diamond \psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ by letting $(\omega \diamond \psi, t) = \omega((\psi, t))$ for every $t \in T_\Sigma(Z)$.

Observation 5.9. Let $Z \neq \emptyset$, $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ be rational, and $\omega \in \Omega^{(1)}$. Then $\omega \diamond \psi$ is rational.

Proof. Let $\eta \in \mathrm{RatExp}(\Sigma, Z, \underline{A})$ be a rational expression such that $[\![\eta]\!] = \psi$. We set $\eta' = (\omega.z) \cdot_z \eta$ for some arbitrary $z \in Z$ and prove that $[\![\eta']\!] = \omega \diamond \psi$. For $t \in T_\Sigma(Z)$, we have $([\![\eta']\!], t) = \big(\omega.z \cdot_z \psi, t\big) = \omega \circ_{\{\varepsilon\},\{\varepsilon\}} ((\psi, t)) = \omega((\psi, t)) = (\omega \diamond \psi, t)$. □

This means that the operation $\diamond$ can be added to the rational operations. It corresponds to scalar multiplication in [7] where ω is the scalar.

6 From wta to rational expressions

6.1 Decomposition of runs

Here we will decompose runs of a wta at certain nodes for which a particular property U holds. The used notion of decomposition (and the used notation) was introduced in [7], and we extend this here to wta which can handle variables. Henceforth, let $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ be a wta. We decompose a run r of M by cutting off the prefix r' of r, whose leaves are nodes that have the property U and no inner node (except potentially the root) of r' fulfils U. Let us first define the used properties, called node properties, formally.

Definition 6.1. A *node property (of M)* is a mapping $U \colon R_M \to \mathcal{P}(\mathbb{N}^*)$ such that $U(r) \subseteq \mathrm{pos}(r)$ for every $r \in R_M$.

In essence a node property just selects some nodes of a run. Now we define the decomposition of a run r of M into subruns according to a given node property U. It is convenient to consider a special (uniquely determined) decomposition, such that the cuts are (uniquely) determined by U. We achieve this by performing the decomposition at the topmost positions of r (disregarding the root) for which the node property U holds, i.e., which are elements of $U(r)$. Those nodes w at which the cuts occur are replaced by the variable z.

Definition 6.2. Let U be a node property, $z \in Z$, and $r \in R_M$. We define the (U, z)-*decomposition of* r, denoted by $\mathrm{dec}_{U,z}(r)$, by

$$\mathrm{dec}_{U,z}(r) = (r', (w_1, r|_{w_1}), \ldots, (w_m, r|_{w_m}))$$

with

$$r' = r[w_i \leftarrow \langle z, \pi_2(r(w_i))\rangle \mid i \in [m]\,]$$

where (i) $\{w_1, \ldots, w_m\}$ is the set of all positions $w \in U(r) \setminus \{\varepsilon\}$ such that $v \notin U(r)$ for every prefix $v \in \mathrm{pos}(r) \setminus \{\varepsilon\}$ of w and (ii) $w_1 <_{\mathrm{lex}} \cdots <_{\mathrm{lex}} w_m$.

Intuitively speaking, given a run r, the decomposition yields the prefix run, in which the label of each cut-point w is now $\langle z, q\rangle$ where q was the state associated to node w in r, and the subruns that were cut off from r together with their positions in r. It is important to note that given U and z the decomposition is uniquely determined and the decomposition also uniquely determines the run r (see Observation 6.4(3)). In other words the mapping $\mathrm{dec}_{U,z}$ is injective.

We will need the fact that the weight mapping c_M distributes over the replacement of runs.

Lemma 6.3. For every $r' \in R_M$, $V = \{w_1, \ldots, w_m\} \subseteq W = \mathrm{pos}_{\langle Z,Q\rangle}(r')$ with $w_1 <_{\mathrm{lex}} \cdots <_{\mathrm{lex}} w_m$, and for all $r_1, \ldots, r_m \in R_M$ such that $\pi_2(r_i(\varepsilon)) = \pi_2(r'(w_i))$ for every $i \in [m]$, if $\nu(z)_q = \mathrm{id}$ for every $\langle z, q\rangle \in \{r'(w_i) \mid i \in [m]\}$, then

$$c_M(r'[w_1 \leftarrow r_1, \ldots, w_m \leftarrow r_m]) = c_M(r') \circ_{W,V} (c_M(r_1), \ldots, c_M(r_m)).$$

Proof. The proof is done by induction on r'. Since this is straightforward, we will only discuss its main steps and leave the details to the reader. Assume that r' is of the form $\langle \sigma, q\rangle(r'_1, \ldots, r'_k)$. Then the replacement $[w_1 \leftarrow r_1, \ldots, w_m \leftarrow r_m]$ is split and distributed to the subruns $r'_1, \ldots, r'_k$ appropriately. After having used the definition of c_M on the root $\langle \sigma, q\rangle$ of the run $r'[w_1 \leftarrow r_1, \ldots, w_m \leftarrow r_m]$, we can apply the induction hypothesis. Finally we can apply Proposition 5.2, use the definition of c_M on the run r', and obtain the desired equality. $\square$

The next statement collects some further trivial observations about (U, z)-decompositions, where the fourth part follows from Lemma 6.3.

Observation 6.4. Let $X \subseteq Z$, $z \in X$, $P \subseteq Q$, and U be a node property. Moreover, let $r \in R_M^{X,P}$ and $\mathrm{dec}_{U,z}(r) = (r', (w_1, r_1), \ldots, (w_m, r_m))$. Finally, let $W = \mathrm{pos}_{\langle Z,Q\rangle}(r')$ and $V = \{w_1, \ldots, w_m\}$. Then

1. $r' \in R_M^{X,P}$ and $r'(\varepsilon) = r(\varepsilon)$;
2. $r_i \in R_M^{X,P}$ and $\pi_2(r_i(\varepsilon)) = \pi_2(r'(w_i))$ for every $i \in [m]$;
3. $r = r'[w_1 \leftarrow r_1, \ldots, w_m \leftarrow r_m]$; and
4. if $\nu(z)_q = \mathrm{id}$ for every $\langle z, q\rangle \in \{r'(w_i) \mid i \in [m]\}$ then

$$c_M(r) = c_M(r') \circ_{W,V} (c_M(r_1), \ldots, c_M(r_m)).$$

We note that in Observation 4.6(4) of [7] the commutativity of the semiring is needed.

6.2 The analysis of the wta

In this section we will show that, for every wta over a DM-monoid, a (semantically) equivalent rational expression can effectively be constructed. We will use the concept of an X-private z-initial state in the wta, which is established in Section 4.

Definition 6.5. For every $P \subseteq Q$, $X \subseteq Z$, and $q \in Q$ we define the mapping $S(M)_q^{X,P} \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ such that for every $t \in T_\Sigma(Z)$

$$(S(M)_q^{X,P}, t) = \begin{cases} \bigoplus_{r \in R_M^{X,P}(t,q)} c_M(r) & \text{if } t \in T_\Sigma(Z) \setminus X, \\ \mathbf{0} & \text{if } t \in X. \end{cases}$$

Note that, by definition, $S(M)_q^{X,P}$ is x-proper for every $x \in X$. The mappings $S(M)_q^{X,P}$ (which are due to [8]) are central in the following development. We now show a recursion equation which specifies $S(M)_q^{X,P\cup\{p\}}$ only in terms of $S(M)_q^{X,P}$, $S(M)_p^{X,P}$, z-concatenation, and z-Kleene-star. If we apply the obtained recursion equation exhaustively, then this allows us to compute $S(M)_q^{X,Q}$ using the rational operations (z-concatenation and z-Kleene-star for various $z \in Z$) and $S(M)_q^{X,\emptyset}$ for every $q \in Q$.

Lemma 6.6 (cf. [8])**.** Let $\underline{A}$ be distributive. Moreover, let $X \subseteq Z$ and $z \in X$ be such that $p \in Q$ is an X-private z-initial state. For every $P \subseteq Q$ with $p \notin P$ and $q \in Q$,

$$S(M)_q^{X,P\cup\{p\}} = S(M)_q^{X,P} \cdot_z \left(S(M)_p^{X,P}\right)_z^*.$$

Proof. It is easily shown that both the left and the right hand side are x-proper for every $x \in X$. We let $U \colon R_M \to \mathcal{P}(\mathbb{N}^*)$ be the node property such that

$$U(r) = \{w \in \mathrm{pos}(r) \mid \pi_2(r(w)) = p\}$$

for every $r \in R_M$, i.e., all nodes of the run r that are labelled with the state p. Let $P' = P \cup \{p\}$. It remains to prove the statement for every $t \in T_\Sigma(Z) \setminus X$. This is achieved by induction on t, as follows:

$$\begin{aligned} (S(M)_q^{X,P'}, t) &= \bigoplus_{r \in R_M^{X,P'}(t,q)} c_M(r) \\ &= \quad \text{(by Observation 6.4 which is applicable because } r'(w_i) = \langle z, p\rangle \text{ and } \nu(z)_p = \mathrm{id}) \\ &\quad \bigoplus_{\substack{r \in R_M^{X,P'}(t,q) \\ \mathrm{dec}_{U,z}(r) = (r', (w_1, r_1), \ldots, (w_m, r_m))}} c_M(r') \circ_{\mathrm{pos}_{\langle Z,Q\rangle}(r'), \mathrm{pos}_{\langle z,p\rangle}(r')} (c_M(r_1), \ldots, c_M(r_m)) \end{aligned}$$

$=$ (because (i) the subruns uniquely determine a run; and
(ii) $c_M(r') = \mathbf{0}$ if r' contains a node $\langle z, q' \rangle$ for some $q' \in Q \setminus \{p\}$ as $\nu(z)_{q'} = \mathbf{0}$)

$$\bigoplus_{\substack{t' \in T_\Sigma(Z) \setminus X, r' \in R_M^{X,P}(t',q), m = |t'|_z, \\ (\forall i \in [m]) \colon r_i \in R_M^{X,P'}(t_i,p), \\ E}} c_M(r') \circ_{t',z} (c_M(r_1), \dots, c_M(r_m))$$

where $E = \{(t_1, \dots, t_m) \in T_\Sigma(Z)^m \mid t = t'[z \leftarrow (t_1, \dots, t_m)]\}$

$=$ (by distributivity, cf. Observations 3.2 and 3.3)

$$\bigoplus_{\substack{t' \in T_\Sigma(Z) \setminus X, \\ m = |t'|_z, E}} \Big(\bigoplus_{r' \in R_M^{X,P}(t',q)} c_M(r') \Big) \circ_{t',z} \Big(\bigoplus_{r_1 \in R_M^{X,P'}(t_1,p)} c_M(r_1), \dots, \bigoplus_{r_m \in R_M^{X,P'}(t_m,p)} c_M(r_m) \Big)$$

$=$ (by definition of $S(M)_q^{X,P}$ because $t' \notin X$, and by definition of $S(M)_p^{X,P'}$ because $\nu(z)_p = \mathrm{id}$ and $\nu(x)_p = \mathbf{0}$ for every $x \in X$ with $x \neq z$)

$$\bigoplus_{\substack{t' \in T_\Sigma(Z), \\ m = |t'|_z, E}} (S(M)_q^{X,P}, t') \circ_{t',z} \Big(\big(S(M)_p^{X,P'} \oplus^{\mathrm{u}} \mathrm{id}.z, t_1 \big), \dots, \big(S(M)_p^{X,P'} \oplus^{\mathrm{u}} \mathrm{id}.z, t_m \big) \Big)$$

$=$ (by induction hypothesis because t_i is a proper subtree of t)

$$\bigoplus_{\substack{t' \in T_\Sigma(Z), \\ m = |t'|_z, E}} (S(M)_q^{X,P}, t') \circ_{t',z} \Big(\big(\big(S(M)_p^{X,P} \cdot_z \big(S(M)_p^{X,P} \big)_z^* \big) \oplus^{\mathrm{u}} \mathrm{id}.z, t_1 \big), \dots,$$
$$\big(\big(S(M)_p^{X,P} \cdot_z \big(S(M)_p^{X,P} \big)_z^* \big) \oplus^{\mathrm{u}} \mathrm{id}.z, t_m \big) \Big)$$

$=$ (by definition of $\cdot_z$)

$$\Big(S(M)_q^{X,P} \cdot_z \big(\big(S(M)_p^{X,P} \cdot_z \big(S(M)_p^{X,P} \big)_z^* \big) \oplus^{\mathrm{u}} \mathrm{id}.z \big), t \Big)$$

$=$ (by Lemma 5.5)

$$\Big(S(M)_q^{X,P} \cdot_z \big(S(M)_p^{X,P} \big)_z^*, t \Big).$$ □

Using the previous lemma, we can analyse the uniform tree valuations which are recognized by a wta.

Lemma 6.7. Let $\underline{A}$ be distributive, and let $X \subseteq Z$ be such that for every $q \in Q$ there exists an $x \in X$ such that q is the X-private x-initial state. For every $P \subseteq Q$ and $q \in Q$ there exists a rational expression $\eta \in \mathrm{RatExp}(\Sigma, Z, \underline{A})$ such that $[\![\eta]\!] = S(M)_q^{X,P}$.

Proof. We prove the statement by induction on the size of P.

Induction base: Let $P = \emptyset$. We distinguish two cases. On the one hand, let $t = \sigma(t_1, \dots, t_k)$ for some $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $t_1, \dots, t_k \in T_\Sigma(Z)$. Then $(S(M)_q^{X,\emptyset}, t) = \bigoplus_{r \in R_M^{X,\emptyset}(t,q)} c_M(r)$. This equals $\mathbf{0}$ if $\{t_1, \dots, t_k\} \not\subseteq X$, because then $R_M^{X,\emptyset}(t,q) = \emptyset$. For $t_1 = x_1, \dots, t_k = x_k \in X$, the last sum equals

$$\bigoplus_{q_1, \dots, q_k \in Q} c_M(\langle \sigma, q \rangle (\langle x_1, q_1 \rangle, \dots, \langle x_k, q_k \rangle)) = \bigoplus_{q_1, \dots, q_k \in Q} \mu_k(\sigma)_{q_1 \dots q_k, q}(\nu(x_1)_{q_1}, \dots, \nu(x_k)_{q_k})$$

by definition of c_M.

On the other hand, suppose that $t = z$ for some $z \in Z$. We immediately have

$$(S(M)_q^{X,\emptyset}, z) = \begin{cases} \mathbf{0} & \text{if } z \in X \\ \nu(z)_q & \text{otherwise.} \end{cases}$$

Thus we obtain

$$S(M)_q^{X,\emptyset} = \Big(\bigoplus^{\mathrm{u}}_{\substack{k\geq 0,\sigma\in\Sigma^{(k)}\\ x_1,\dots,x_k\in X\\ q_1,\dots,q_k\in Q}} \mathrm{top}_{\sigma,\mu_k(\sigma)_{q_1\cdots q_k,q}}(\nu(x_1)_{q_1}.x_1,\dots,\nu(x_k)_{q_k}.x_k)\Big).$$

We let

$$\eta = \Big(\sum_{\substack{k\geq 0,\sigma\in\Sigma^{(k)}\\ x_1,\dots,x_k\in X\\ q_1,\dots,q_k\in Q}} \mathrm{top}_{\sigma,\mu_k(\sigma)_{q_1\cdots q_k,q}}(\nu(x_1)_{q_1}.x_1,\dots,\nu(x_k)_{q_k}.x_k)\Big).$$

Note that $(S(M)_q^{X,\emptyset}, z) = \mathbf{0}$ for $z \in Z\setminus X$. This follows from the fact that q is an X-private x-initial state, which yields $\nu(z)_q = \mathbf{0}$ for every $z \in Z \setminus X$.

Clearly, $\eta \in \mathrm{RatExp}(\Sigma, Z, \underline{A})$ and $[\![\eta]\!] = S(M)_q^{X,\emptyset}$. This completes the induction base.

Induction step: Let $p \notin P$ and $x \in X$ be such that p is the X-private x-initial state. By Lemma 6.6, $S(M)_q^{X,P\cup\{p\}} = S(M)_q^{X,P} \cdot_x \big(S(M)_p^{X,P}\big)^*_x$. By induction hypothesis there exist $\eta_1, \eta_2 \in \mathrm{RatExp}(\Sigma, Z, \underline{A})$ such that $[\![\eta_1]\!] = S(M)_q^{X,P}$ and $[\![\eta_2]\!] = S(M)_p^{X,P}$. Since $S(M)_p^{X,P}$ is x-proper, $[\![\eta_1 \cdot_x (\eta_2)^*_x]\!] = S(M)_q^{X,P\cup\{p\}}$. □

Finally, let us present the relationship between wta and rational expressions in a theorem. We use additional variables in the rational expressions. These variables correspond to the states of the automaton.

Theorem 6.8. Let $\underline{A}$ be distributive and $Z \cap Q = \emptyset$. There exists a rational expression $\eta \in \mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ such that $S(M) = [\![\eta]\!]|_{T_\Sigma(Z)}$. Hence, if Z is finite, then we have $\mathrm{Rec}(\Sigma, Z, \underline{A}) \subseteq \mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A})|_{T_\Sigma(Z)}$.

Proof. We first extend M to a wta $M' = (Q, \Sigma, Z \cup Q, \underline{A}, F, \mu, \nu')$ as follows:

- $\nu'(q)_q = \mathrm{id}$ and $\nu'(q)_p = \mathbf{0}$ for every $p, q \in Q$ with $p \neq q$; and
- $\nu'(z)_q = \nu(z)_q$ for every $z \in Z$ and $q \in Q$.

Clearly, every $q \in Q$ is the Q-private q-initial state in M'. Moreover, we have that $(S(M'), t) = (S(M), t)$ for every $t \in T_\Sigma(Z)$ and thus $S(M')|_{T_\Sigma(Z)} = S(M)$. It remains to prove that there exists a rational expression $\eta \in \mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ such that $[\![\eta]\!] = S(M')$.

By Lemma 6.7 for every $q \in Q$ there exists a rational expression $\eta_q \in \mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ such that $[\![\eta_q]\!] = S(M')_q^{Q,Q}$. We obtain for every $t \in T_\Sigma(Z \cup Q)$

$$(S(M'), t) = \bigoplus_{q\in Q} F_q((S(M')_q, t))$$

$=$ (by the proof of Observation 5.9)

$$\Big(\bigoplus_{q\in Q}\big(F_q.q \cdot_q S(M')_q\big), t\Big)$$

$=$ (by definition of $S(M')_q$ and $S(M')_q^{Q,Q}$ because q is the Q-private q-initial state)

$$\Big(\bigoplus_{q\in Q}\big(F_q.q \cdot_q (S(M')_q^{Q,Q} \oplus^{\mathrm{u}} \mathrm{id}.q\big), t\Big).$$

We thus set $\eta = \sum_{q\in Q}(F_q.q \cdot_q (\eta_q + \mathrm{id}.q))$. Obviously $\eta \in \mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ and $[\![\eta]\!] = S(M')$. □

7 From rational expressions to wta

7.1 Recognizable uniform tree valuations are closed under relabeling

In this subsection we show that recognizable uniform tree valuations are closed under (total and deterministic) relabelings. Let us first introduce the required notions. Let Σ, Δ be ranked alphabets and Z, Y finite sets of variables which are disjoint to Σ and Δ. A *relabeling* is a mapping $f\colon \Sigma \to \Delta$ such that $f(\sigma) \in \Delta^{(k)}$ for every $k \geq 0$ and $\sigma \in \Sigma^{(k)}$. Let $g\colon Z \to Y$ be another mapping. Then (f, g) *induces* a mapping $h\colon T_\Sigma(Z) \to T_\Delta(Y)$, also called a relabeling, which is defined as follows. We have $h(z) = g(z)$ for every $z \in Z$, and $h(\sigma(t_1, \ldots, t_k)) = f(\sigma)(h(t_1), \ldots, h(t_k))$ for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$ and $t_1, \ldots, t_k \in T_\Sigma(Z)$. Now let $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ be a uniform tree valuation. By $h(\psi)$ we denote the uniform tree valuation $h(\psi) \in \mathrm{Uvals}(\Delta, Y, \underline{A})$ which is given by

$$(h(\psi), u) = \bigoplus_{t \in h^{-1}(u)} (\psi, t)$$

for every $u \in T_\Delta(Y)$.

Lemma 7.1. Let $\underline{A}$ be a sum closed DM-monoid. Let $f\colon \Sigma \to \Delta$ be a relabeling, $g\colon Z \to Y$ a mapping, and $h : T_\Sigma(Z) \to T_\Delta(Y)$ the relabeling induced by (f, g). For every $\psi \in \mathrm{Rec}(\Sigma, Z, \underline{A})$ we have $h(\psi) \in \mathrm{Rec}(\Delta, Y, \underline{A})$.

Proof. Let $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ be a wta such that $S(M) = \psi$. We construct the wta $M' = (Q, \Delta, Y, \underline{A}, F, \mu', \nu')$ where

- $\nu'(y)_q = \bigoplus_{z \in g^{-1}(y)} \nu(z)_q$ for every $y \in Y$ and $q \in Q$; and
- $\mu'_k(\delta)_{q_1 \ldots q_k, q} = \bigoplus_{\sigma \in f^{-1}(\delta)} \mu_k(\sigma)_{q_1 \ldots q_k, q}$ for every $k \geq 0$, $\delta \in \Delta^{(k)}$, and $q, q_1, \ldots, q_k \in Q$.

Using the inductive semantics, it can be shown straightforwardly that $S(M')_q = h(S(M)_q)$ for every $q \in Q$. From this we can derive the result as follows for every $u \in T_\Delta(Y)$:

$$\begin{aligned}
(S(M'), u) &= \bigoplus_{q \in Q} F_q((S(M')_q, u)) && = \bigoplus_{q \in Q} F_q((h(S(M)_q), u)) \\
&= \bigoplus_{q \in Q} F_q\Big(\bigoplus_{t \in h^{-1}(u)} (S(M)_q, t)\Big) &&= \bigoplus_{t \in h^{-1}(u)} \Big(\bigoplus_{q \in Q} F_q((S(M)_q, t))\Big) \\
&= (h(S(M)), u)
\end{aligned}$$

where the one-before-last step is by distributivity (Observation 3.3). □

7.2 The synthesis of the wta

Now we show that every rational uniform tree valuation is indeed also recognizable provided that the underlying M-monoid $\underline{A} = (A, \oplus, \mathbf{0}, \Omega)$ is distributive and expressive enough. We begin with showing that the basic rational uniform tree valuation $\omega.z$, where $\omega \in \Omega^{(1)}$ and $z \in Z$, is recognizable. Then we show that recognizable uniform tree valuations are closed under rational operations. This proves the desired inclusion, because the set of rational uniform tree valuations is the smallest set containing $\omega.z$ and being closed under the rational operations.

Lemma 7.2. For every $z \in Z$ and $\omega \in \Omega^{(1)}$ we have $\omega.z \in \mathrm{Rec}(\Sigma, Z, \underline{A})$.

Proof. Construct the wta $M = (\{q\}, \Sigma, Z, \underline{A}, F, \mu, \nu)$ such that $F_q = \mathrm{id}$ and $\nu(z)_q = \omega$. All remaining entries in ν and μ are $\mathbf{0}$. It should be clear that $S(M) = \omega.z$. $\square$

Next we consider the sum operation. We present the construction and its correctness proof separately, because they will also be useful when considering top-concatenation.

Definition 7.3. Let $M' = (Q', \Sigma, Z, \underline{A}, F', \mu', \nu')$ and $M'' = (Q'', \Sigma, Z, \underline{A}, F'', \mu'', \nu'')$ be wta such that $Q' \cap Q'' = \emptyset$. We define the wta $M' \oplus M'' = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ as follows.

- $Q = Q' \cup Q''$;
- $F_p = F'_p$ for every $p \in Q'$ and $F_q = F''_q$ for every $q \in Q''$;
- for every $k \geq 0$ and $\sigma \in \Sigma^{(k)}$ let
$$\mu_k(\sigma)_{p_1 \ldots p_k, p} = \mu'_k(\sigma)_{p_1 \ldots p_k, p} \quad \text{and} \quad \mu_k(\sigma)_{q_1 \ldots q_k, q} = \mu''_k(\sigma)_{q_1 \ldots q_k, q}$$
for every $p, p_1, \ldots, p_k \in Q'$ and $q, q_1, \ldots, q_k \in Q''$; and
- for every $z \in Z$ let $\nu(z)_p = \nu'(z)_p$ for every $p \in Q'$ and $\nu(z)_q = \nu''(z)_q$ for every $q \in Q''$.

Note that μ is $\mathbf{0}$ at all unmentioned entries.

Observation 7.4. $S(M' \oplus M'') = S(M') \oplus^{\mathrm{u}} S(M'')$, and in particular, $S(M' \oplus M'')_p = S(M')_p$ for every $p \in Q'$ and $S(M' \oplus M'')_q = S(M'')_q$ for every $q \in Q''$.

Proof. Let $M = M' \oplus M''$. The proof is straightforward using the following argument: if a run $r \in R_M(t, q')$ with $q' \in Q'$ uses states from Q'', then it has weight $c_M(r) = \mathbf{0}$. Thus the set $R_M(t)$ can be partitioned into $R_{M'}(t)$ and $R_{M''}(t)$. $\square$

Lemma 7.5. The set $\mathrm{Rec}(\Sigma, Z, \underline{A})$ is closed under sum.

Next let us consider the top-concatenation $\mathrm{top}_{\sigma,\omega}$ for some $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $\omega \in \Omega^{(k)}$.

Lemma 7.6. Let $\underline{A}$ be a $(1,\star)$-composition closed and sum closed DM-monoid, and let $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $\omega \in \Omega^{(k)}$. The set $\mathrm{Rec}(\Sigma, Z, \underline{A})$ is closed under $\mathrm{top}_{\sigma,\omega}$.

Proof. For every $i \in [k]$ let $\psi_i \in \mathrm{Rec}(\Sigma, Z, \underline{A})$ and $M_i = (Q_i, \Sigma, Z, \underline{A}, F_i, \mu_i, \nu_i)$ be a wta such that $S(M_i) = \psi_i$. Since $\underline{A}$ is $(1,\star)$-composition closed and sum closed, we can assume by Lemma 4.8 that $M_1, \ldots, M_k$ are wta with terminating states, say, $p_1, \ldots, p_k$, respectively. Without loss of generality, we can assume that $Q_1, \ldots, Q_k$ are pairwise disjoint. Moreover, let $\star \notin \bigcup_{i \in [k]} Q_i$. Finally, let $M' = M_1 \oplus \cdots \oplus M_k$. Suppose that $M' = (Q', \Sigma, Z, \underline{A}, F', \mu', \nu')$. We construct the wta $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ as follows:

- $Q = Q' \cup \{\star\}$;
- $F_\star = \mathrm{id}$ and $F_q = \mathbf{0}$ for every $q \in Q'$;
- $\nu(z)_q = \nu'(z)_q$ for every $z \in Z$ and $q \in Q'$;
- $\mu_n(\delta)_{q_1 \ldots q_n, q} = \mu'_n(\delta)_{q_1 \ldots q_n, q}$ for every $n \geq 0$, $\delta \in \Sigma^{(n)}$, and $q, q_1, \ldots, q_n \in Q'$;
- $\mu_k(\sigma)_{p_1 \ldots p_k, \star} = \omega$ (let us note that $F'_{p_i} = \mathrm{id}$).

- All remaining entries in ν and μ are $\mathbf{0}$.

We claim that $S(M) = \mathrm{top}_{\sigma,\omega}(\psi_1, \ldots, \psi_k)$. Since M is a wta with terminating state $\star$ we immediately conclude that $S(M) = S(M)_\star$ by Observation 4.7. Now we prove $(S(M)_\star, t) = (\mathrm{top}_{\sigma,\omega}(\psi_1, \ldots, \psi_k), t)$ for every $t \in T_\Sigma(Z)$ by case analysis.

Suppose that $t = z$ for some $z \in Z$. Then $(S(M)_\star, z) = \nu(z)_\star = \mathbf{0} = (\mathrm{top}_{\sigma,\omega}(\psi_1, \ldots, \psi_k), z)$.

Now suppose that $t = \delta(t_1, \ldots, t_n)$ for some $n \geq 0$, $\delta \in \Sigma^{(n)}$, and $t_1, \ldots, t_n \in T_\Sigma(Z)$. If $\delta \neq \sigma$, then we have $(S(M)_\star, t) = \mathbf{0} = (\mathrm{top}_{\sigma,\omega}(\psi_1, \ldots, \psi_k), t)$. Now let $\delta = \sigma$. Then, by the fact that the state sets $Q_1, \ldots, Q_k$ are pairwise disjoint, we have that

$$(S(M)_\star, t) = \bigoplus_{\substack{q_1 \in Q_1, \ldots, q_k \in Q_k \\ (\forall i \in [k]) \colon r_i \in R_{M_i}(t_i, q_i)}} \mu_k(\sigma)_{q_1 \ldots q_k, \star}(c_{M_1}(r_1), \ldots, c_{M_k}(r_k)).$$

By construction of $\mu_k(\sigma)$, this is equal to $\bigoplus_{(\forall i \in [k]) \colon r_i \in R_{M_i}(t_i, p_i)} \omega(c_{M_1}(r_1), \ldots, c_{M_k}(r_k))$. By distributivity and the definition of $(S(M_i)_{p_i}, t_i)$ this is equal to $\omega((S(M_1)_{p_1}, t_1), \ldots, (S(M_k)_{p_k}, t_k))$. Since p_i is the terminating state of M_i, we have that $(S(M_i)_{p_i}, t_i) = (S(M_i), t_i) = (\psi_i, t_i)$, and hence we obtain that

$$\omega((S(M_1)_{p_1}, t_1), \ldots, (S(M_k)_{p_k}, t_k)) = (\mathrm{top}_{\sigma,\omega}(\psi_1, \ldots, \psi_k), t). \qquad \square$$

Let us proceed with concatenation.

Lemma 7.7. Let $\underline{A}$ be a $(1, \star)$-composition closed and sum closed DM-monoid and $z \in Z$. The set $\mathrm{Rec}(\Sigma, Z, \underline{A})$ is closed under z-concatenation.

Proof. Let $\psi', \psi'' \in \mathrm{Rec}(\Sigma, Z, \underline{A})$. Moreover, let $M' = (Q', \Sigma, Z, \underline{A}, F', \mu', \nu')$ be a wta with a terminating state $q' \in Q'$ such that $S(M') = S(M')_{q'} = \psi'$ (see Lemma 4.8 and Observation 4.7). Moreover, let Δ be a ranked alphabet such that $\Sigma \cap \Delta = \emptyset$ and there exists a bijective relabeling $f \colon \Sigma \to \Delta$. Let Y be a set such that $Z \cap Y = \emptyset$ and there exists a bijective mapping $g \colon Z \to Y$. Let θ be the (f, g)-induced relabeling. By Theorem 7.1 we have that $\theta(\psi'') \in \mathrm{Rec}(\Delta, Y, \underline{A})$. Finally, let $M'' = (Q'', \Delta, Y, \underline{A}, F'', \mu'', \nu'')$ be a wta with $Q' \cap Q'' = \emptyset$ and a terminating state $q'' \in Q''$ such that $S(M'') = S(M'')_{q''} = \theta(\psi'')$ (see again Lemma 4.8 and Observation 4.7). We construct the wta $M = (Q, \Sigma \cup \Delta, Z \cup Y, \underline{A}, F, \mu, \nu)$ as follows:

- $Q = Q' \cup Q''$;
- $F_{q'} = \mathrm{id}$ and $F_p = \mathbf{0}$ for every $p \in Q \setminus \{q'\}$;
- $\nu(z)_q = \mathbf{0}$ and $\nu(x)_q = \nu'(x)_q$ for every $x \in Z$ such that $x \neq z$ and $q \in Q'$;
- $\nu(y)_p = \nu''(y)_p$ for every $y \in Y$ and $p \in Q''$;
- $\nu(y)_q = \nu'(z)_q(\nu''(y)_{q''})$ for every $y \in Y$ and $q \in Q'$;
- $\mu_k(\sigma)_{q_1 \ldots q_k, q} = \mu'_k(\sigma)_{q_1 \ldots q_k, q}$ for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q, q_1, \ldots, q_k \in Q'$;
- $\mu_k(\delta)_{p_1 \ldots p_k, p} = \mu''_k(\delta)_{p_1 \ldots p_k, p}$ for every $k \geq 0$, $\delta \in \Delta^{(k)}$, and $p, p_1, \ldots, p_k \in Q''$; and
- $\mu_k(\delta)_{p_1 \ldots p_k, q} = \nu'(z)_q(\mu''_k(\delta)_{p_1 \ldots p_k, q''})$ for every $k \in \mathbb{N}$, $\delta \in \Delta^{(k)}$, $q \in Q'$, and $p_1, \ldots, p_k \in Q''$.
- All remaining entries in ν and μ are $\mathbf{0}$.

Clearly, M is a wta with terminating state q'. We claim that $S(M) = \psi' \cdot_z \theta(\psi'')$. In order to prove this, we first prove that

$$M(t[z \leftarrow (s_1, \ldots, s_n)])_q = M'(t)_q \circ_{t,z} \big((\theta(\psi''), s_1), \ldots, (\theta(\psi''), s_n)\big) \qquad (\dagger)$$

for every $n \geq 0$, $t \in T_\Sigma(Z)$ with $|t|_z = n$, $q \in Q'$, and $s_1, \ldots, s_n \in T_\Delta(Y)$. We prove this statement by induction on t.

Induction base: Suppose that $t = z$ and thus $n = 1$. Moreover, suppose that $s_1 = y$ for some $y \in Y$. Then $M(z[z \leftarrow (y)])_q = M(y)_q = \nu'(z)_q(\nu''(y)_{q''})$ by Lemma 3.13. Since q'' is the terminating state of M'' and by Lemma 3.13, the latter equals $M'(z)_q((S(M''), y))$, which in turn is equal to $M'(z)_q \circ_{\{\varepsilon\},\{\varepsilon\}} ((\theta(\psi''), y))$.

Now suppose that $t = z$ (and hence $n = 1$) and $s_1 = \delta(t_1, \ldots, t_k)$ for some $k \geq 0$, $\delta \in \Delta^{(k)}$, and $t_1, \ldots, t_k \in T_\Delta(Y)$. Then $M(z[z \leftarrow (\delta(t_1, \ldots, t_k))])_q$ equals

$$M(\delta(t_1, \ldots, t_k))_q = \bigoplus_{q_1, \ldots, q_k \in Q} \mu_k(\delta)_{q_1 \ldots q_k, q}(M(t_1)_{q_1}, \ldots, M(t_k)_{q_k})$$

by Lemma 3.13. Clearly, $M(t)_p = M''(t)_p$ for every $t \in T_\Delta(Y)$ and $p \in Q''$. So by definition of $\mu_k(\delta)$, the latter equals

$$\bigoplus_{p_1, \ldots, p_k \in Q''} \Big(\nu'(z)_q(\mu''_k(\delta)_{p_1 \ldots p_k, q''})\Big)(M''(t_1)_{p_1}, \ldots, M''(t_k)_{p_k}),$$

which by distributivity and associativity (see Observations 3.3 and 3.4) and the inductive definition of $S(M'')$ (see Lemma 3.13) can be rewritten to

$$M'(z)_q((S(M''), \delta(t_1, \ldots, t_k))) = M'(z)_q \circ_{\{\varepsilon\},\{\varepsilon\}} ((\theta(\psi''), \delta(t_1, \ldots, t_k))) \ .$$

We complete the induction base with the case that $t = x$ for some $x \in Z$ such that $x \neq z$ and thereby $n = 0$. Then $M(x[z \leftarrow ()])_q = M(x)_q = M'(x)_q = M'(x)_q \circ_{\{\varepsilon\},\emptyset} ()$.

Induction step: Let $t = \sigma(t_1, \ldots, t_k)$ for some $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $t_1, \ldots, t_k \in T_\Sigma(Z)$. For every $i \in [k]$ let $m_i = \sum_{j \in [i]} |t_j|_z$. Then

$$\begin{aligned}
& M(\sigma(t_1, \ldots, t_k)[z \leftarrow (s_1, \ldots, s_n)])_q \\
= {} & M(\sigma(t_1[z \leftarrow (s_1, \ldots, s_{m_1})], \ldots, t_k[z \leftarrow (s_{m_{k-1}+1}, \ldots, s_n)]))_q \\
= {} & \quad \text{(by Lemma 3.13)} \\
& \bigoplus_{q_1, \ldots, q_k \in Q} \mu_k(\sigma)_{q_1 \ldots q_k, q}\big(M(t_1[z \leftarrow (s_1, \ldots, s_{m_1})])_{q_1}, \ldots, \\
& \qquad\qquad M(t_k[z \leftarrow (s_{m_{k-1}+1}, \ldots, s_n)])_{q_k}\big) \\
= {} & \quad \text{(by induction hypothesis and because } \mu_k(\sigma)_{q_1 \ldots q_k, q} \neq \mathbf{0} \text{ only if } q_1, \ldots, q_k \in Q'\text{)} \\
& \bigoplus_{q_1, \ldots, q_k \in Q'} \mu'_k(\sigma)_{q_1 \ldots q_k, q}\big(M'(t_1)_{q_1} \circ_{t_1,z} ((\theta(\psi''), s_1), \ldots, (\theta(\psi''), s_{m_1})), \ldots, \\
& \qquad\qquad M'(t_k)_{q_k} \circ_{t_k,z} ((\theta(\psi''), s_{m_{k-1}+1}), \ldots, (\theta(\psi''), s_n))\big) \\
= {} & \quad \text{(by Proposition 5.2)} \\
& \Big(\bigoplus_{q_1, \ldots, q_k \in Q'} \mu'_k(\sigma)_{q_1 \ldots q_k, q}(M'(t_1)_{q_1}, \ldots, M'(t_k)_{q_k})\Big) \circ_{t,z} ((\theta(\psi''), s_1), \ldots, (\theta(\psi''), s_n)) \\
= {} & \quad \text{(by Lemma 3.13)} \\
& M'(\sigma(t_1, \ldots, t_k))_q \circ_{t,z} ((\theta(\psi''), s_1), \ldots, (\theta(\psi''), s_n)).
\end{aligned}$$

We have thus proved Equation (†). It remains to prove that $S(M) = \psi' \cdot_z \theta(\psi'')$. Let $T_\Sigma(Z)[z \leftarrow T_\Delta(Y)]$ denote the set of all trees $t[z \leftarrow (s_1, \dots, s_n)]$ where $t \in T_\Sigma(Z)$, $|t|_z = n$, and $s_i \in T_\Delta(Y)$. Then clearly, $\mathrm{supp}(S(M)) \subseteq T_\Sigma(Z)[z \leftarrow T_\Delta(Y)]$ and $\mathrm{supp}(\psi' \cdot_z \theta(\psi'')) \subseteq T_\Sigma(Z)[z \leftarrow T_\Delta(Y)]$. Moreover, for each $u \in T_\Sigma(Z)[z \leftarrow T_\Delta(Y)]$ there exists a unique decomposition into $t \in T_\Sigma(Z)$ and $s_1, \dots, s_n \in T_\Delta(Y)$ such that $u = t[z \leftarrow (s_1, \dots, s_n)]$. Then

$$
\begin{aligned}
&\big(S(M), t[z \leftarrow (s_1, \dots, s_n)]\big)\\
= &\qquad \text{(by Observation 4.7 because } q' \text{ is the terminating state of } M)\\
&M(t[z \leftarrow (s_1, \dots, s_n)])_{q'}\\
= &\qquad \text{(by Equation (†))}\\
&M'(t)_{q'} \circ_{t,z} \big((\theta(\psi''), s_1), \dots, (\theta(\psi''), s_n)\big)\\
= &\qquad \text{(by Observation 4.7 because } q' \text{ is the terminating state of } M')\\
&(S(M'), t) \circ_{t,z} \big((\theta(\psi''), s_1), \dots, (\theta(\psi''), s_n)\big)\\
= &\qquad \text{(by definition of } \cdot_z \text{ because the decomposition is unique)}\\
&\big(\psi' \cdot_z \theta(\psi''), t[z \leftarrow (s_1, \dots, s_n)]\big).
\end{aligned}
$$

Hence $\psi' \cdot_z \theta(\psi'') \in \mathrm{Rec}(\Sigma \cup \Delta, Z \cup Y, \underline{A})$. Now consider the additional relabeling θ' induced by the mappings (f', g') where $f' \colon \Sigma \cup \Delta \to \Sigma$ and $g' \colon Z \cup Y \to Z$ are defined for every $s \in \Sigma \cup \Delta$ and $x \in Z \cup Y$ by

$$
f'(s) = \begin{cases} s & \text{if } s \in \Sigma,\\ f^{-1}(s) & \text{if } s \in \Delta, \end{cases} \qquad \text{and} \qquad g'(x) = \begin{cases} x & \text{if } x \in Z,\\ g^{-1}(x) & \text{if } x \in Y. \end{cases}
$$

Clearly, $\theta'(\psi' \cdot_z \theta(\psi'')) = \psi' \cdot_z \psi''$ and, since recognizable uniform tree valuations are closed under relabeling (see Theorem 7.1), we proved the theorem. □

Finally, we consider the KLEENE-star. We will prove closure under z-KLEENE-star provided that the underlying DM-monoid is $(1, \star)$-composition closed and sum closed. However, we first prove the closure in unary-composition closed and sum closed DM-monoids and later use this result for the proof of the correctness of the construction that uses the relaxed requirements.

Lemma 7.8. Let $z \in Z$ and $\underline{A}$ be a unary-composition closed and sum closed DM-monoid. The set $\mathrm{Rec}(\Sigma, Z, \underline{A})$ is closed under z-KLEENE-star.

Proof. If $\psi \in \mathrm{Rec}(\Sigma, Z, \underline{A})$ is not z-proper, then $\psi_z^* = \widetilde{\mathbf{0}}$, which is trivially recognizable. So, let $\psi \in \mathrm{Rec}(\Sigma, Z, \underline{A})$ be z-proper; i.e., we have $(\psi, z) = \mathbf{0}$. Moreover, let $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ be a wta such that $S(M) = \psi$. Without loss of generality, let us suppose that M has a private z-initial variable state p and a terminating state p' (see Lemma 4.10). Note that $p \neq p'$ because ψ is z-proper. We construct the wta $M' = (Q', \Sigma, Z, \underline{A}, F', \mu', \nu')$ as follows:

- $Q' = Q \setminus \{p'\}$;
- $F'_p = \mathrm{id}$ and $F'_q = \mathbf{0}$ for every $q \in Q' \setminus \{p\}$;
- $\nu'(x)_q = \nu(x)_q$ for every $x \in Z$ and $q \in Q'$ with $q \neq p$;
- $\nu'(z)_p = \mathrm{id}$ and $\nu'(x)_p = \nu(x)_{p'}$ for every $x \in Z$ with $x \neq z$;
- $\mu'_k(\sigma)_{q_1 \dots q_k, q} = \mu_k(\sigma)_{q_1 \dots q_k, q}$ for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q, q_1, \dots, q_k \in Q'$ with $q \neq p$; and

- $\mu'_k(\sigma)_{q_1\dots q_k,p} = \mu_k(\sigma)_{q_1\dots q_k,p'}$ for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q_1, \dots, q_k \in Q'$.

For every $t \in T_\Sigma(Z)$, we have $(S(M'), t) = (S(M')_p, t) = (S(M')_p^{\{z\},Q'} \oplus^{\mathrm{u}} \mathrm{id}.z, t)$ where the last equality is due to the definition of $S(M')_p^{\{z\},Q'}$. Clearly, p is the $\{z\}$-private z-initial state, so by Lemma 6.6 we obtain

$$\Big(\big(S(M')_p^{\{z\},Q'\setminus\{p\}} \cdot_z \big(S(M')_p^{\{z\},Q'\setminus\{p\}}\big)^*_z\big) \oplus^{\mathrm{u}} \mathrm{id}.z, t\Big) = \Big(\big(S(M) \cdot_z S(M)^*_z\big) \oplus^{\mathrm{u}} \mathrm{id}.z, t\Big)$$

because $(S(M), z) = \mathbf{0}$ and $S(M')_p^{\{z\},Q'\setminus\{p\}} = S(M)_{p'} = S(M)$ (note that the runs in $R_{M'}^{\{z\},Q'\setminus\{p\}}(t,p)$ are the same as the runs in $R_M(t,p')$, apart from the labels of the roots). Finally, by Lemma 5.5, the latter is equal to $(S(M)^*_z, t)$. □

Let us show now that $(\star, 1)$-composition closedness of the DM-monoid is actually not necessary for the previous statement. We chose to present the matter in this way because the proof of the statement with the relaxed condition is now easier. We present a construction that utilizes only $(1, \star)$-compositions and sum and then show that the construction is correct by showing the resulting automaton recognizes the same uniform tree valuation as the automaton in the previous lemma. This can be done since each DM-monoid can be extended to a unary-composition closed and sum closed DM-monoid (see Lemma 4.4).

Lemma 7.9. Let $z \in Z$ and $\underline{A}$ be a $(1, \star)$-composition closed and sum closed DM-monoid. The set $\mathrm{Rec}(\Sigma, Z, \underline{A})$ is closed under z-Kleene-star.

Proof. Again, if $\psi \in \mathrm{Rec}(\Sigma, Z, \underline{A})$ is not z-proper, then ψ^*_z is trivially recognizable. In the sequel, let $\psi \in \mathrm{Rec}(\Sigma, Z, \underline{A})$ be z-proper; i.e., we have $(\psi, z) = \mathbf{0}$. Moreover, let $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ be a wta such that $S(M) = \psi$. Without loss of generality, suppose that M has the terminating state p (see Lemma 4.8). Let $f\colon Q \to P$ be a bijection for some set P such that $P \cap Q = \emptyset$. Let $g\colon Q \cup P \to Q$ be the mapping defined for every $q \in Q \cup P$ by

$$g(q) = \begin{cases} q & \text{if } q \in Q, \\ f^{-1}(q) & \text{if } q \in P. \end{cases}$$

We construct the wta $M' = (Q', \Sigma, Z, \underline{A}, F', \mu', \nu')$ as follows.

- $Q' = Q \cup P$;
- $F'_p = \mathrm{id}$ and $F'_q = \mathbf{0}$ for every $q \in Q \cup P$ with $q \neq p$;
- $\nu'(z)_p = \mathrm{id}$, $\nu'(z)_q = \mathbf{0}$ for every $q \in Q \setminus \{p\}$, and $\nu'(z)_{f(q)} = \nu(z)_q$ for every $q \in Q$;
- $\nu'(x)_q = \nu(x)_q$ and $\nu'(x)_{f(q)} = \nu(z)_q(\nu(x)_p)$ for every $x \in Z$ with $x \neq z$ and $q \in Q$;
- for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, $q \in Q$ and $q_1, \dots, q_k \in Q \cup P$

$$\mu'_k(\sigma)_{q_1\dots q_k,q} = \mu_k(\sigma)_{g(q_1)\dots g(q_k),q}$$
$$\mu'_k(\sigma)_{q_1\dots q_k,f(q)} = \nu(z)_q(\mu_k(\sigma)_{g(q_1)\dots g(q_k),p}).$$

It remains to prove that $S(M') = S(M)^*_z$. For this we show that $S(M') = S(M'')$ where $M'' = (Q, \Sigma, Z, \underline{B}, F'', \mu'', \nu'')$ is constructed according to the proof of Lemma 7.8. Note that this requires several steps. First the wta M can be seen as a wta over Σ, Z, and $\underline{B}$, where $\underline{B}$ is the sum closed and unary-composition closed DM-monoid constructed from $\underline{A}$ in Lemma 4.4. To this wta we apply Lemma 4.5; note that p is a terminating state of the resulting wta because $(\psi, z) = \mathbf{0}$ and hence $\nu(z)_p = \mathbf{0}$. Then, we apply the construction in Lemma 7.8 to the resulting wta and obtain the wta M''. We repeat the combined constructions of Lemmata 4.5 and 7.8 for the convenience of the proof (note that in M'', for convenience, we renamed the state z into p). The wta $M'' = (Q, \Sigma, Z, \underline{B}, F'', \mu'', \nu'')$ is given by

- $F''_p = \mathrm{id}$ and $F''_q = \mathbf{0}$ for every $q \in Q \setminus \{p\}$;
- $\nu''(z)_p = \mathrm{id}$ and $\nu''(z)_q = \mathbf{0}$ for every $q \in Q \setminus \{p\}$;
- $\nu''(x)_q = \nu(x)_q$ for every $x \in Z$ with $x \neq z$ and $q \in Q$;
- for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q, q_1, \ldots, q_k \in Q$

$$\mu''_k(\sigma)_{q_1 \ldots q_k, q} = \bigoplus_{\substack{p_1,\ldots,p_k \in Q \setminus \{p\}, \\ (\forall i \in [k])\colon p_i = q_i \text{ if } q_i \neq p}} \mu_k(\sigma)_{p_1 \ldots p_k, q}(f_{p_1,q_1}, \ldots, f_{p_k,q_k})$$

where for every $p', q' \in Q$

$$f_{p',q'} = \begin{cases} \nu(z)_{p'} & \text{if } q' = p, \\ \mathrm{id} & \text{otherwise.} \end{cases}$$

We first prove that $M''(t)_q = M'(t)_q$ and $\nu(z)_q(M''(t)_p) = M'(t)_{f(q)}$ for every $q \in Q$ and $t \in T_\Sigma(Z)$. This is achieved by induction on t.

<u>Induction base:</u> Since this is immediate we leave it to the reader.

<u>Induction step:</u> Let $t = \sigma(t_1, \ldots, t_k)$ for some $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $t_1, \ldots, t_k \in T_\Sigma(Z)$. Then we have $M''(\sigma(t_1, \ldots, t_k))_q = \bigoplus_{q_1,\ldots,q_k \in Q} \mu''_k(\sigma)_{q_1 \ldots q_k, q}(M''(t_1)_{q_1}, \ldots, M''(t_k)_{q_k})$ by Lemma 3.13. We continue as follows:

$$\begin{aligned}
& \bigoplus_{q_1,\ldots,q_k \in Q} \Bigg(\bigoplus_{\substack{p_1,\ldots,p_k \in Q \setminus \{p\}, \\ (\forall i \in [k])\colon p_i = q_i \text{ if } q_i \neq p}} \mu_k(\sigma)_{p_1 \ldots p_k, q}(f_{p_1,q_1}, \ldots, f_{p_k,q_k}) \Bigg) (M''(t_1)_{q_1}, \ldots, M''(t_k)_{q_k}) \\
= \quad & \text{(by Observations 3.2 and 3.4)} \\
& \bigoplus_{\substack{q_1,\ldots,q_k \in Q \\ p_1,\ldots,p_k \in Q \setminus \{p\} \\ (\forall i \in [k])\colon p_i = q_i \text{ if } q_i \neq p}} \mu_k(\sigma)_{p_1 \ldots p_k, q}(f_{p_1,q_1}(M''(t_1)_{q_1}), \ldots, f_{p_k,q_k}(M''(t_k)_{q_k})) \\
= \quad & \text{(by definition of } f_{p_i,q_i} \text{ and induction hypothesis)} \\
& \bigoplus_{\substack{q_1,\ldots,q_k \in Q \\ p_1,\ldots,p_k \in Q \setminus \{p\} \\ (\forall i \in [k])\colon p_i = q_i \text{ if } q_i \neq p}} \mu_k(\sigma)_{p_1 \ldots p_k, q} \Bigg(\left\{ \begin{matrix} M'(t_1)_{f(p_1)} & \text{if } q_1 = p \\ M'(t_1)_{p_1} & \text{otherwise} \end{matrix} \right\}, \\
& \qquad \ldots, \\
& \qquad \left\{ \begin{matrix} M'(t_k)_{f(p_k)} & \text{if } q_k = p \\ M'(t_k)_{p_k} & \text{otherwise} \end{matrix} \right\} \Bigg) \\
= \quad & \text{(by case analysis)} \\
& \bigoplus_{q_1,\ldots,q_k \in (Q \cup P) \setminus \{p, f(p)\}} \mu_k(\sigma)_{g(q_1) \ldots g(q_k), q}(M'(t_1)_{q_1}, \ldots, M'(t_k)_{q_k}) \\
= \quad & \text{(by definition of } \mu'\text{)} \\
& \bigoplus_{q_1,\ldots,q_k \in (Q \cup P) \setminus \{p, f(p)\}} \mu'_k(\sigma)_{q_1 \ldots q_k, q}(M'(t_1)_{q_1}, \ldots, M'(t_k)_{q_k})
\end{aligned}$$

$$
\begin{aligned}
= \quad & \text{(because } \mu'_k(\sigma)_{q_1\ldots q_k,q} = \mathbf{0} \text{ if there exists an } i \text{ such that } g(q_i) = p\text{;} \\
& \text{this holds because } p \text{ is a terminating state)} \\
& \bigoplus_{q_1,\ldots,q_k \in Q\cup P} \mu'_k(\sigma)_{q_1\ldots q_k,q}(M'(t_1)_{q_1},\ldots,M'(t_k)_{q_k}) = M'(\sigma(t_1,\ldots,t_k))_q
\end{aligned}
$$

where the last step is by Lemma 3.13. Let us continue with the second equality:

$$
\begin{aligned}
& \nu(z)_q(M''(t)_p) \\
= \quad & \text{(by the previous chain of equations)} \\
& \nu(z)_q\Big(\bigoplus_{q_1,\ldots,q_k \in (Q\cup P)\setminus\{p,f(p)\}} \mu_k(\sigma)_{g(q_1)\ldots g(q_k),p}(M'(t_1)_{q_1},\ldots,M'(t_k)_{q_k})\Big) \\
= \quad & \text{(by Observations 3.3 and 3.4)} \\
& \bigoplus_{q_1,\ldots,q_k \in (Q\cup P)\setminus\{p,f(p)\}} \big(\nu(z)_q(\mu_k(\sigma)_{g(q_1)\ldots g(q_k),p})\big)(M'(t_1)_{q_1},\ldots,M'(t_k)_{q_k}) \\
= \quad & \text{(by definition of } \mu'\text{)} \\
& \bigoplus_{q_1,\ldots,q_k \in (Q\cup P)\setminus\{p,f(p)\}} \mu'_k(\sigma)_{q_1\ldots q_k,f(q)}(M'(t_1)_{q_1},\ldots,M'(t_k)_{q_k}) \\
= \quad & \text{(by the last two steps in the previous chain of equations)} \\
& M'(\sigma(t_1,\ldots,t_k))_{f(q)}.
\end{aligned}
$$

Now, for every $t \in T_\Sigma(Z)$, we have $(S(M''),t) = M''(t)_p = M'(t)_p = (S(M'),t)$. Thus $S(M') = S(M'') = S(M)^*_z$ by Lemma 7.8, which proves the statement. $\square$

Theorem 7.10. Let $\underline{A}$ be a $(1,\star)$-composition closed and sum closed DM-monoid. Moreover, let Z be a finite set. Then $\mathrm{Rec}(\Sigma,Z,\underline{A})$ contains the uniform tree valuation $\omega.z$ for every $z \in Z$ and $\omega \in \Omega^{(1)}$, and it is closed under the rational operations as defined in Definition 5.3. In particular, $\mathrm{Rat}(\Sigma,Z,\underline{A}) \subseteq \mathrm{Rec}(\Sigma,Z,\underline{A})$.

Proof. The statement follows from Lemmata 7.2, 7.5, 7.6, 7.7, and 7.9. $\square$

8 The main result and the special case of semirings

Now we put the analysis and synthesis of automata together and prove KLEENE's result for wta with variables over M-monoids. Then we instantiate this to the case of semirings.

First, we define the concept of lifting in order to have type correct results (as discussed in the Introduction). For this, let $\psi \in \mathrm{Uvals}(\Sigma,Z,\underline{A})$ and Q be a set. We extend ψ to the mapping $\mathrm{lift}_Q(\psi) \in \mathrm{Uvals}(\Sigma,Z\cup Q,\underline{A})$ by defining that for every $t \in T_\Sigma(Z\cup Q)$

$$
(\mathrm{lift}_Q(\psi),t) = \begin{cases} (\psi,t) & \text{if } t \in T_\Sigma(Z), \\ \mathbf{0} & \text{otherwise.} \end{cases}
$$

Also, let us assume that there is a countable infinite set Θ such that every finite set (in particular, state sets Q and variable sets Z) can be chosen as subset of Θ. Then, for every $\psi \in \mathrm{Uvals}(\Sigma,Z,\underline{A})$, the lifting to Θ results in $\mathrm{lift}_\Theta(\psi) \in \mathrm{Uvals}(\Sigma,\Theta,\underline{A})$. Henceforth we will drop Θ from lift_Θ and just write lift.

As second technical preparation for KLEENE's result, we consider again Theorem 6.8. In the right hand side of that statement, we have restricted the semantics $[\![\eta]\!]$ to $T_\Sigma(Z)$. This was necessary in order to have the same functional type on both sides of the equation. Clearly, there is also the dual way, i.e., extend the mapping $S(M)$ by lift_Q. Let us now show that as long as Q is finite, there exists no difference (so far as rationality is concerned) in the two approaches.

Lemma 8.1. Let $\psi \in \mathrm{Uvals}(\Sigma, Z, \underline{A})$ and Q be a finite set. There exists a rational expression $\eta \in \mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ such that $[\![\eta]\!]|_{T_\Sigma(Z)} = \psi$ if and only if there exists a rational expression $\eta' \in \mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ such that $[\![\eta']\!] = \mathrm{lift}_Q(\psi)$.

Proof. The proof of the if-direction is trivial because if there exists $\eta' \in \mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ such that $[\![\eta']\!] = \mathrm{lift}_Q(\psi)$ then in particular $[\![\eta']\!]|_{T_\Sigma(Z)} = \psi$. For the opposite direction, assume that there exists $\eta \in \mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ such that $[\![\eta]\!]|_{T_\Sigma(Z)} = \psi$. Moreover, let $Q = \{q_1, \ldots, q_k\}$ and let $\eta' = (\cdots(\eta \cdot_{q_1} \mathbf{0}.q_1)\cdots) \cdot_{q_k} \mathbf{0}.q_k$. Clearly, η' is in $\mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ and $[\![\eta']\!] = \mathrm{lift}_Q(\psi)$. □

Now we can prove KLEENE's result for wta with variables over M-monoids. Here we assume that the function lift is extended to classes of uniform tree valuations in the usual way.

Theorem 8.2. For every $(1,\star)$-composition closed and sum closed DM-monoid $\underline{A}$, we have that $\mathrm{lift}(\mathrm{Rec}(\Sigma, \mathrm{fin}, \underline{A})) = \mathrm{lift}(\mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A}))$.

Proof. First we prove the inclusion from right to left. Let $Z \subseteq \Theta$ be a finite set. Then, by Theorem 7.10, we have $\mathrm{Rat}(\Sigma, Z, \underline{A}) \subseteq \mathrm{Rec}(\Sigma, Z, \underline{A})$, and thus the application of lift to both sides yields the desired inclusion.

Second we prove the inclusion from left to right. Let $\varphi \in \mathrm{lift}(\mathrm{Rec}(\Sigma, \mathrm{fin}, \underline{A}))$. Then there exists a finite set Z and $\psi \in \mathrm{Rec}(\Sigma, Z, \underline{A})$ such that $\varphi = \mathrm{lift}(\psi)$. Since ψ is recognizable, there exists a wta $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ such that $S(M) = \psi$. Without loss of generality, we can assume that $Z \cap Q = \emptyset$. We note that $\mathrm{lift}(\psi) = \mathrm{lift}(\mathrm{lift}_Q(\psi))$.

By Theorem 6.8 and Lemma 8.1, there exists $\eta \in \mathrm{RatExp}(\Sigma, Z \cup Q, \underline{A})$ such that $[\![\eta]\!] = \mathrm{lift}_Q(S(M)) = \mathrm{lift}_Q(\psi)$. Thus, $\mathrm{lift}_Q(\psi) \in \mathrm{Rat}(\Sigma, Z \cup Q, \underline{A}) \subseteq \mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A})$. Applying lift we obtain that $\varphi \in \mathrm{lift}(\mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A}))$. □

We can also derive KLEENE's theorem for $\underline{A}$-wta of [17], i.e., wta over DM-monoids (without variables).

Theorem 8.3. For every $(1,\star)$-composition closed and sum closed DM-monoid $\underline{A}$, we have that $\mathrm{Rec}(\Sigma, \emptyset, \underline{A}) = \mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A})|_{T_\Sigma}$.

Proof. By Theorems 6.8 and 7.10, we have that

$$\mathrm{Rec}(\Sigma, \emptyset, \underline{A}) \subseteq \mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{A})|_{T_\Sigma} \subseteq \mathrm{Rec}(\Sigma, \mathrm{fin}, \underline{A})|_{T_\Sigma} \subseteq \mathrm{Rec}(\Sigma, \emptyset, \underline{A})$$

where the last inclusion can be seen as follows. Let $\psi \in \mathrm{Rec}(\Sigma, \mathrm{fin}, \underline{A})|_{T_\Sigma}$. Thus, there exist a wta $M = (Q, \Sigma, Z, \underline{A}, F, \mu, \nu)$ such that $\psi = S(M)|_{T_\Sigma}$. It is easy to see that for the wta $N = (Q, \Sigma, \emptyset, \underline{A}, F, \mu, \emptyset)$ we have that $S(N) = S(M)|_{T_\Sigma}$. Thus $\psi \in \mathrm{Rec}(\Sigma, \emptyset, \underline{A})$. □

In the next part of this section, we show how to obtain KLEENE's result for recognizable tree series over an *arbitrary semiring* from our KLEENE result for wta over DM-monoids. To this end, we need some preparations. First we recall the definition of a wta over an arbitrary semiring. In fact, we adapt Definition 4.1 of [7], where a wta over a commutative semiring was defined. However, we deviate in the definition of the weight of a run on an input tree.

Let $\underline{K} = (K, \oplus, \odot, \mathbf{0}, \mathbf{1})$ be an arbitrary semiring. A wta over $\underline{K}$ is a tuple $N = (Q, \Sigma, \delta, F)$, where Q is a finite set of states, Σ is the input ranked alphabet, $F \subseteq Q$ is the set of final states, and $\delta = (\delta_\sigma \mid \sigma \in \Sigma)$ is a family of state behaviors with weights, such that, for every $k \geq 0$ and $\sigma \in \Sigma^{(k)}$, we have a mapping $\delta_\sigma : Q^k \times Q \to K$. The equation $\delta_\sigma(q_1, \ldots, q_k, q) = a$ means that the weight (or: cost) of the transition $(q_1, \ldots, q_k) \xrightarrow{\sigma} q$ is a. The semantics of N is the tree series $S(N) \in K\langle\langle T_\Sigma \rangle\rangle$, which is defined as follows. For an input tree t and a state $q \in Q$, a q-run r on t is a tree $r \in T_{\langle \Sigma, Q\rangle}$ such that $\pi_1(r) = t$ and $\pi_2(r(\varepsilon)) = q$. Hence, $r = \langle \sigma, q\rangle(r_1, \ldots, r_k)$ for some $k \geq 0$, $\sigma \in \Sigma^{(k)}$, $q \in Q$, and q_i-runs r_i for some $q_i \in Q$, $1 \leq i \leq k$. The weight of r is defined by the equation $c_N(r) = c_N(r_1) \odot \cdots \odot c_N(r_k) \odot \delta_\sigma(q_1, \ldots, q_k, q)$. (Thus, referring to the discussion in the Introduction, we fix the postorder tree walk as product order. But we could as well take the preorder tree walk, and adapt the definition of mul_a in Definition 8.4 appropriately.) Finally, for every $t \in T_\Sigma$ we define

$$(S(N), t) = \bigoplus_{q \in F} \Big(\bigoplus_{r \in R_N(t,q)} c_N(r) \Big),$$

where $R_N(t, q)$ is the set of q-runs on t.

For a ranked alphabet Σ and a finite set Z (of nullary symbols) with $\Sigma \cap Z = \emptyset$, we denote the class of tree series which are recognized by wta over the ranked input alphabet $\Sigma \cup Z$ and semiring $\underline{K}$ by $\mathrm{Rec}_{\mathrm{sr}}(\Sigma, Z, \underline{K})$. We abbreviate $\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \emptyset, \underline{K})$ by $\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{K})$ and the union $\bigcup_{Z \text{ finite set}} \mathrm{Rec}_{\mathrm{sr}}(\Sigma, Z, \underline{K})$ by $\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \mathrm{fin}, \underline{K})$.

Next, for a semiring $\underline{K}$, we define an M-monoid $\underline{D}(\underline{K})$ such that we can express recognizability of $\underline{K}$ by recognizability over $\underline{D}(\underline{K})$. In particular, the right-multiplication of the k-fold product $a_1 \odot \cdots \odot a_k$ with an $a \in K$ is simulated by the k-ary operation $\mathrm{mul}_a^{(k)}$ of $\underline{D}(\underline{K})$ which is defined as follows.

Definition 8.4. Let $(K, \oplus, \odot, \mathbf{0}, \mathbf{1})$ be a semiring, $a \in K$, and $k \geq 0$ an integer. The *k-ary multiplication with a* is the mapping $\mathrm{mul}_a^{(k)} \colon K^k \to K$ that is defined as follows: for every $a_1, \ldots, a_k \in K$ we have $\mathrm{mul}_a^{(k)}(a_1, \ldots, a_k) = a_1 \odot \cdots \odot a_k \odot a$.

Note that $\mathrm{mul}_a^{(0)}() = a$. Next we simulate a semiring with the help of a DM-monoid.

Definition 8.5. Let $\underline{K} = (K, \oplus, \odot, \mathbf{0}, \mathbf{1})$ be a semiring. For every integer $k \geq 0$, let $\Omega^{(k)} = \{\mathrm{mul}_a^{(k)} \mid a \in K\}$. We denote by $\underline{D}(\underline{K})$ the M-monoid $(K, \oplus, \mathbf{0}, \Omega)$. Note that $\mathrm{id}_K = \mathrm{mul}_{\mathbf{1}}^{(1)}$ and $\mathbf{0}^k = \mathrm{mul}_{\mathbf{0}}^{(k)}$ for every $k \geq 0$.

Lemma 8.6. For every semiring $\underline{K}$, the M-monoid $\underline{D}(\underline{K})$ is distributive, sum closed, and $(1, \star)$-composition closed.

Proof. Let $\underline{K}$ and $\underline{D}(\underline{K})$ be as in Definition 8.5. We observe that $\underline{D}(\underline{K})$ is distributive because equations (d-M) and (a-M) easily follow from the semiring properties (d1-SR), (d2-SR), and (a-SR). It is sum closed and $(1, \star)$-composition closed because, obviously, for every $a, b \in K$, $\mathrm{mul}_a^{(k)} \oplus \mathrm{mul}_b^{(k)} = \mathrm{mul}_{a \oplus b}^{(k)}$ (using (d1-SR)) and $\mathrm{mul}_a^{(1)}(\mathrm{mul}_b^{(k)}) = \mathrm{mul}_{b \odot a}^{(k)}$, respectively. ☐

Next we show that, for every ranked alphabet Σ and semiring $\underline{K}$, the tree series recognizable by wta over $\underline{K}$ and the uniform tree valuations with empty variable set Z recognizable by wta over the DM-monoid $\underline{D}(\underline{K})$ in the sense of Definition 3.5 coincide. (Let us recall, that uniform tree valuations with empty variable set Z are in fact also tree series.) We will need the following concept.

Definition 8.7. Let $\underline{K} = (K, \oplus, \odot, \mathbf{0}, \mathbf{1})$ be a semiring. A wta $M = (Q, \Sigma, \delta, F)$ over $\underline{K}$ and a wta $N = (Q, \Sigma, \emptyset, \underline{D}(\underline{K}), \overline{F}, \mu, \emptyset)$ over the DM-monoid $\underline{D}(\underline{K})$ are *related* if the following conditions hold.

- For every $q \in Q$,
$$\overline{F}_q = \begin{cases} \mathrm{id} & \text{if } q \in F \\ \mathbf{0} & \text{otherwise.} \end{cases}$$
- For every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $q_1, \ldots, q_k, q \in Q$, $\mu_k(\sigma)_{q_1 \ldots q_k, q} = \mathrm{mul}_a^{(k)}$ if and only if $\delta_\sigma(q_1, \ldots, q_k, q) = a$.

The following observation is obvious, the proof is left as an exercise.

Observation 8.8. If M and N are related, then we have $S(M) = S(N)$.

Lemma 8.9. For every semiring $\underline{K}$, we have that $\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{K}) = \mathrm{Rec}(\Sigma, \emptyset, \underline{D}(\underline{K}))$.

Proof. The inclusion $\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{K}) \subseteq \mathrm{Rec}(\Sigma, \emptyset, \underline{D}(\underline{K}))$ is immediate from Observation 8.8. To prove $\mathrm{Rec}(\Sigma, \emptyset, \underline{D}(\underline{K})) \subseteq \mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{K})$, let us take a wta $M = (Q, \Sigma, \emptyset, \underline{D}(\underline{K}), F, \mu, \emptyset)$ over $\underline{D}(\underline{K})$. By Lemmas 8.6 and 4.8, there exists a wta $M' = (Q', \Sigma, \emptyset, \underline{D}(\underline{K}), F', \mu', \emptyset)$ with a terminating state such that $S(M) = S(M')$. Then, there also exists a wta N over the semiring $\underline{K}$ such that N and M' are related (cf. Definition 8.7). Moreover, by Observation 8.8, $S(N) = S(M')$. □

Now we prove KLEENE's theorem for tree series over an arbitrary semiring.

Theorem 8.10. For every semiring $\underline{K}$, we have that $\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{K}) = \mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{D}(\underline{K}))|_{T_\Sigma}$.

Proof. By Lemma 8.9, Lemma 8.6, and Theorem 8.3, we have
$$\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{K}) = \mathrm{Rec}(\Sigma, \emptyset, \underline{D}(\underline{K})) = \mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{D}(\underline{K}))|_{T_\Sigma} \ .$$ □

Now we turn to commutative semirings and show that KLEENE's theorem for commutative semirings [7] is a special case of the main result of this paper. For this, we will relate the rational expressions over a commutative semiring $\underline{K}$ to our rational expressions over $\underline{D}(\underline{K})$. In order to be able to do this, we slightly modify the definition of the rational expressions of both [7] and our paper. However, we manage these changes in a way that the results of both papers remain valid. First let us recall and modify the definition of rational tree series from [7]. Let Σ be a ranked alphabet, Z a finite set (of nullary symbols) with $\Sigma \cap Z = \emptyset$, and $\underline{K}$ a commutative semiring. By a rational expression over Σ, Z, and $\underline{K}$, we mean a rational expression over $\Sigma \cup Z$ and $\underline{K}$ in the sense of Definition 3.17 of [7] with the following two modifications. Firstly, as opposed to (5) and (6) of that definition, we allow the α-concatenation and the α-KLEENE star only for $\alpha \in Z$. Secondly, as opposed to (6), we allow to form the α-KLEENE star η_α^* for an arbitrary rational expression η. Note that in (6) of the discussed definition, η_α^* is a rational expression only if the semantics $[\![\eta]\!]_{\mathrm{sr}}$ of η is an α-proper tree series and in that case $([\![\eta_\alpha^*]\!]_{\mathrm{sr}}, t) = ([\![\eta]\!]_{\mathrm{sr},\alpha}^{\mathrm{ht}(t)+1}, t)$ for every $t \in T_\Sigma$. (Here $[\![\eta]\!]_{\mathrm{sr}}$ denotes the "semiring semantics" of η according to Definition 3.17 of [7].) Therefore we complete the definition of the semantics of η_α^* such that if $[\![\eta]\!]_{\mathrm{sr}}$ is not α-proper, then we define $[\![\eta_\alpha^*]\!]_{\mathrm{sr}} = \widetilde{\mathbf{0}}$. We denote by $\mathrm{Rat}_{\mathrm{sr}}(\Sigma, Z, \underline{K})$ the set of all rational tree series that are defined by these rational expressions over Σ, Z, and $\underline{K}$. Moreover, we abbreviate the class $\bigcup_{Z \text{ finite set}} \mathrm{Rat}_{\mathrm{sr}}(\Sigma, Z, \underline{K})$ by $\mathrm{Rat}_{\mathrm{sr}}(\Sigma, \mathrm{fin}, \underline{K})$. Now we recall KLEENE's theorem from [7] and give a sketch of the proof in order to demonstrate that the two changes we made in the definition of rational tree series have no effect on the correctness of the theorem. At the same time, we correct the small flaw (mentioned in the Introduction) by adding 'lift'.

Proposition 8.11 (Theorem 7.1 of [7]). For every commutative semiring $\underline{K}$, we have that $\text{lift}(\text{Rec}_{\text{sr}}(\Sigma, \text{fin}, \underline{K})) = \text{lift}(\text{Rat}_{\text{sr}}(\Sigma, \text{fin}, \underline{K}))$.

Proof. The proof of the inclusion from right to left is based on the fact that recognizable tree series are closed under the rational operations, which is shown in Theorem 6.8 of [7]. We have to check whether this theorem remains valid because we allowed to apply the α-KLEENE star not only to an α-proper but to an arbitrary tree series ψ. In fact, if $\psi \in \text{Rec}_{\text{sr}}(\Sigma, Z, \underline{K})$ is α-proper, then Lemma 6.7 of [7] proves that $\psi^*_{\widetilde{\alpha}}$ is recognizable. If, on the other hand, $\psi \in \text{Rec}_{\text{sr}}(\Sigma, Z, \underline{K})$ is not α-proper, then $\psi^*_{\alpha} = \widetilde{\mathbf{0}}$, which is obviously recognizable. Hence $\text{Rat}_{\text{sr}}(\Sigma, Z, \underline{K}) \subseteq \text{Rec}_{\text{sr}}(\Sigma, Z, \underline{K})$ follows, which verifies the desired inclusion. Moreover, by a careful reading of the proof of Theorem 5.2 of [7], we can see that for every $\psi \in \text{Rec}_{\text{sr}}(\Sigma, Z, \underline{K})$ there exists a finite set Q (in fact the state set of the wta recognizing ψ) such that $\text{lift}_Q(\psi) \in \text{Rat}_{\text{sr}}(\Sigma, Z \cup Q, \underline{K})$ because only q-concatenations and q-KLEENE stars with $q \in Q$ appear in the constructed rational expression. This proves that the left-hand side is a subset of the right-hand side. □

Next we show that Proposition 8.11 is in fact a consequence of our Theorem 8.10. For this, let us change the definition of rational expressions over Σ, Z, and $\underline{D}(\underline{K})$. We define $\text{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$ in the following way. We change (i) of Definition 5.6 such that, for every $z \in Z$, we have $z \in \text{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$ with semantics $[\![z]\!] = \text{id}.z$. Moreover, we change (ii) of the same definition such that for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and rational expressions $\eta_1, \ldots, \eta_k$, we have $\sigma(\eta_1, \ldots, \eta_k) \in \text{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$ with semantics $[\![\sigma(\eta_1, \ldots, \eta_k)]\!] = \text{top}_{\sigma, \text{mul}_1^{(k)}}([\![\eta_1]\!], \ldots, [\![\eta_k]\!])$. Then, we take over (iii), (iv), and (v) of Definition 5.6. Finally, for every $a \in K$ and rational expression η, we allow that $a\eta \in \text{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$ with semantics $[\![a\eta]\!] = \text{mul}_a^{(1)} \diamond [\![\eta]\!]$ (where $\diamond$ has been defined at the end of Section 5). By Observation 5.9, $[\![a\eta]\!]$ is a rational uniform tree valuation (if $[\![\eta]\!]$ is one). It is easily seen that $\text{Rat}(\Sigma, Z, \underline{D}(\underline{K})) = \text{Rat}'(\Sigma, Z, \underline{D}(\underline{K}))$, where the latter is the set of uniform tree evaluations defined by the rational expressions in $\text{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$; note, for instance, that $[\![\text{mul}_a^{(1)}.z]\!] = [\![az]\!]$ and $[\![\text{top}_{\sigma, \text{mul}_a^{(k)}}(\eta_1, \ldots, \eta_k)]\!] = [\![a\sigma(\eta_1, \ldots, \eta_k)]\!]$. On the other hand, with this definition of $\text{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$ and the modification of the rational expressions of [7], we have achieved that rational expressions over Σ, Z, and the semiring $\underline{K}$ are syntactically the same as rational expressions over Σ, Z, and the DM-monoid $\underline{D}(\underline{K})$. Thus we can relate rational tree series over Σ, Z, and $\underline{K}$ and rational uniform tree valuations over Σ, Z, and $\underline{D}(\underline{K})$. For this we will need the auxiliary function $\text{one} : \text{Uvals}(\Sigma, Z, \underline{D}(\underline{K})) \to K\langle\langle T_{\Sigma \cup Z}\rangle\rangle$ which is defined in the following way. For every $\psi \in \text{Uvals}(\Sigma, Z, \underline{D}(\underline{K}))$ and $t \in T_{\Sigma \cup Z}$, let $(\text{one}(\psi), t) = (\psi, t)(\mathbf{1}, \ldots, \mathbf{1})$, where the number of arguments $\mathbf{1}$ is $|t|_Z$. Note that we identify $T_\Sigma(Z)$ and $T_{\Sigma \cup Z}$ and that $(\text{one}(\psi), t) = (\psi, t)$ for every $t \in T_\Sigma$. We extend one to classes in the usual way.

Lemma 8.12. For every commutative semiring $\underline{K}$, we have that $\text{Rat}_{\text{sr}}(\Sigma, Z, \underline{K}) = \text{one}(\text{Rat}(\Sigma, Z, \underline{D}(\underline{K})))$.

Proof. It suffices to show that, for every $\eta \in \text{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$, we have $[\![\eta]\!]_{\text{sr}} = \text{one}([\![\eta]\!])$, where $[\![\eta]\!]_{\text{sr}}$ denotes the semantics of η according to Definition 3.17 of [7]. This follows from the statement (†): for every $\eta \in \text{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$, $t \in T_\Sigma(Z)$, and $a_1, \ldots, a_n \in K$, we have that

$$([\![\eta]\!], t)(a_1, \ldots, a_n) = ([\![\eta]\!]_{\text{sr}}, t) \odot a_1 \odot \ldots \odot a_n$$

which we prove by induction on η. The proof is obvious when η has one of the forms z, $\sigma(\eta_1, \ldots, \eta_k)$, $\eta_1 + \eta_2$, or $a\eta$. Therefore, let $\eta = \eta_1 \cdot_z \eta_2$ with

$z \in Z$ and $t \in T_\Sigma(Z)$. We abbreviate the sequence $a_1, \ldots, a_n$ by $\overline{a}$ and the set $\{(s, u_1, \ldots, u_m) \in T_\Sigma(Z)^{m+1} \mid m = |s|_z, t = s[z \leftarrow (u_1, \ldots, u_m)]\}$ by E. Then

$$
\begin{aligned}
&([\![\eta_1 \cdot_z \eta_2]\!], t)(\overline{a}) = ([\![\eta_1]\!] \cdot_z [\![\eta_2]\!], t)(\overline{a}) \\
&= \Big(\bigoplus_E^{\mathrm{u}} ([\![\eta_1]\!], s) \circ_{s,z} (([\![\eta_2]\!], u_1), \ldots, ([\![\eta_2]\!], u_m))\Big)(\overline{a}) \\
&= \bigoplus_E \big(([\![\eta_1]\!], s) \circ_{s,z} (([\![\eta_2]\!], u_1), \ldots, ([\![\eta_2]\!], u_m))\big)(\overline{a}) \qquad (\dagger\dagger).
\end{aligned}
$$

Now we apply the definition of $\circ_{s,z}$ which yields $([\![\eta_1]\!], s)(seq)$ where seq is a sequence of identities interleaved with the operations of the form $([\![\eta_2]\!], u_i)$. Now, due to the definition of composition, the a_i's are distributed over the entries in seq; in particular, the operations $([\![\eta_2]\!], u_i)$ are applied to appropriate subsequences of $\overline{a}$. For such applications we can use the induction hypothesis. Finally, we apply the induction hypothesis also to $([\![\eta_1]\!], s)$ with its arguments, and continue with:

$$
\begin{aligned}
&= \bigoplus_E ([\![\eta_1]\!]_{\mathrm{sr}}, s) \odot ([\![\eta_2]\!]_{\mathrm{sr}}, u_1) \odot \cdots \odot ([\![\eta_2]\!]_{\mathrm{sr}}, u_m) \odot \prod_{i=1}^{k} a_i \\
&= ([\![\eta_1 \cdot_z \eta_2]\!]_{\mathrm{sr}}, t) \odot \prod_{i=1}^{k} a_i.
\end{aligned}
$$

Now we consider the z-Kleene star. First we show the following statement $(\dagger\dagger\dagger)$ by induction on n. For every $\eta \in \mathrm{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$, $t \in T_\Sigma(Z)$, and $(\overline{a}) = (a_1, \ldots, a_n) \in K^n$, if $([\![\eta]\!], t)(\overline{a}) = ([\![\eta]\!]_{\mathrm{sr}}, t) \odot \prod_{i=1}^k a_i$, then $([\![\eta]\!]_z^n, t)(\overline{a}) = ([\![\eta]\!]_{\mathrm{sr},n}^n, t) \odot \prod_{i=1}^k a_i$ for every $z \in Z$ and $n \geq 0$. The case $n = 0$ is clear because $([\![\eta]\!]_z^0), t)(\overline{a}) = \widetilde{\mathbf{0}}(\overline{a}) = \mathbf{0}$. The induction step is proved as follows.

$$
([\![\eta]\!]_z^{n+1}, t)(\overline{a}) = ([\![\eta]\!] \cdot_z [\![\eta]\!]_z^n \oplus^{\mathrm{u}} \mathrm{id}.z, t)(\overline{a}) = ([\![\eta]\!] \cdot_z [\![\eta]\!]_z^n, t)(\overline{a}) \oplus (\mathrm{id}.z, t)(\overline{a}).
$$

The subexpression $([\![\eta]\!] \cdot_z [\![\eta]\!]_z^n, t)(\overline{a})$ is equal to $([\![\eta]\!]_{\mathrm{sr}} \cdot_z [\![\eta]\!]_{\mathrm{sr},z}^n, t) \odot \prod_{i=1}^n a_i$ which is proved in exactly the same way as $(\dagger)$ for the concatenation of η_1 and η_2 except that at $(\dagger\dagger)$ the induction hypothesis of statement $(\dagger\dagger\dagger)$ has to be applied at the inner subexpressions. Then we can continue with:

$$
([\![\eta]\!]_{\mathrm{sr}} \cdot_z [\![\eta]\!]_{\mathrm{sr},z}^n, t) \odot \prod_{i=1}^{n} a_i \oplus ([\![z]\!]_{\mathrm{sr}}, t) \odot \prod_{i=1}^{n} a_i = ([\![\eta]\!]_{\mathrm{sr},z}^{n+1}, t) \odot \prod_{i=1}^{n} a_i \ .
$$

Now, let $\eta \in \mathrm{RatExp}'(\Sigma, Z, \underline{D}(\underline{K}))$, $z \in Z$, and assume first that $[\![\eta]\!]$ is z-proper. Then

$$
([\![\eta_z^*]\!], t)(\overline{a}) = ([\![\eta]\!]_z^{\mathrm{ht}(t)+1}, t)(\overline{a}) = ([\![\eta]\!]_{\mathrm{sr},z}^{\mathrm{ht}(t)+1}, t) \odot \prod_{i=1}^{k} a_i = ([\![\eta_z^*]\!]_{\mathrm{sr}}, t) \odot \prod_{i=1}^{k} a_i \ .
$$

If $[\![\eta]\!]$ is not z-proper, then obviously $([\![\eta_z^*]\!], t)(\overline{a}) = \mathbf{0} = ([\![\eta_z^*]\!]_{\mathrm{sr}}, t) \odot \prod_{i=1}^n a_i$. This finishes the proof of our lemma. $\square$

Now we can show that the main result of [7] is a consequence of Theorem 8.10.

Proposition 8.13. For every commutative semiring $\underline{K}$, we have that $\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{K}) = \mathrm{Rat}_{\mathrm{sr}}(\Sigma, \mathrm{fin}, \underline{K})|_{T_\Sigma}$.

Proof. By Lemma 8.12 we obtain $\mathrm{Rat}_{\mathrm{sr}}(\Sigma, Z, \underline{K})|_{T_\Sigma} = \mathrm{Rat}(\Sigma, Z, \underline{D}(\underline{K}))|_{T_\Sigma}$ because, for every $\psi \in \mathrm{Rat}(\Sigma, Z, \underline{D}(\underline{K}))$ and $t \in T_\Sigma$, we have $(\mathrm{one}(\psi), t) = (\psi, t)$. Consequently, $\mathrm{Rat}_{\mathrm{sr}}(\Sigma, \mathrm{fin}, \underline{K})|_{T_\Sigma} = \mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{D}(\underline{K}))|_{T_\Sigma}$ and thus the statement easily follows from Theorem 8.10. □

Now it can be seen as follows that Proposition 8.13 implies Proposition 8.11. For the inclusion $\mathrm{lift}(\mathrm{Rec}_{\mathrm{sr}}(\Sigma, \mathrm{fin}, \underline{K})) \subseteq \mathrm{lift}(\mathrm{Rat}_{\mathrm{sr}}(\Sigma, \mathrm{fin}, \underline{K}))$ we first use Proposition 8.13, next that $\mathrm{Rat}_{\mathrm{sr}}(\Sigma \cup Z, Q, \underline{K}) \subseteq \mathrm{Rat}_{\mathrm{sr}}(\Sigma, Z \cup Q, \underline{K})$, and finally an analogue of Lemma 8.1 for $\mathrm{Rat}_{\mathrm{sr}}$. For the other inclusion we can prove that $\mathrm{Rat}_{\mathrm{sr}}(\Sigma, Z, \underline{K}) \subseteq \mathrm{Rec}_{\mathrm{sr}}(\Sigma, Z, \underline{K})$. This follows from $\mathrm{Rat}_{\mathrm{sr}}(\Sigma, Z, \underline{K}) \subseteq \mathrm{Rat}_{\mathrm{sr}}(\Sigma \cup Z, Q, \underline{K})|_{T_\Sigma(Z)}$ for some Q (in bijection with Z), and Proposition 8.13.

Finally, as an example of Theorem 8.10, we present a tree series, called post, which is recognizable by a wta over a non-commutative semiring, and we show how to construct a rational expression for post in the sense of Theorem 8.10. As semiring we choose the formal language semiring $\underline{\mathcal{P}(\Sigma^*)}$ over some alphabet Σ where $\underline{\mathcal{P}(\Sigma^*)} = (\mathcal{P}(\Sigma^*), \cup, \cdot, \emptyset, \{\varepsilon\})$, where $\cdot$, the multiplication in $\underline{\mathcal{P}(\Sigma^*)}$, is the concatenation of languages. Note that $\underline{\mathcal{P}(\Sigma^*)}$ is not commutative.

Now add also ranks to the symbols in Σ, i.e., Σ is a ranked alphabet. The tree series $\mathrm{post}\colon T_\Sigma \to \Sigma^*$ drops the parentheses (and) and the commas from its input tree t and shows the symbols of t in post order. More formally,

$$\mathrm{post}(\sigma(t_1, \ldots, t_k)) = \mathrm{post}(t_1) \cdot \ldots \cdot \mathrm{post}(t_k) \cdot \sigma$$

for every $k \geq 0$, $\sigma \in \Sigma^{(k)}$ and $t_1, \ldots, t_k \in T_\Sigma$.

Obviously, $\mathrm{post} \in \mathrm{Rec}_{\mathrm{sr}}(\Sigma, \underline{\mathcal{P}(\Sigma^*)})$ because there is a wta M over $\underline{\mathcal{P}(\Sigma^*)}$ such that $S(M) = \mathrm{post}$. Indeed, let $M = (Q, \Sigma, \delta, F)$ be the wta with $Q = \{q\}$, $F = Q$, and $\delta_\sigma(q, \ldots, q, q) = \{\sigma\}$ for every $k \geq 0$ and $\sigma \in \Sigma^{(k)}$. It should be clear that $S(M) = \mathrm{post}$.

Next we give a rational expression η in the sense of Definition 5.6 such that $[\![\eta]\!]|_{T_\Sigma} = S(M)$. For this, consider the DM-monoid $\underline{D}(\underline{\mathcal{P}(\Sigma^*)})$ and the wta $N = (Q, \Sigma, \emptyset, \underline{D}(\underline{\mathcal{P}(\Sigma^*)}), \overline{F}, \mu, \emptyset)$ such that M and N are related (cf. Definition 8.7). Note that $\mu_k(\sigma)_{q\ldots q,q} = \mathrm{mul}^{(k)}_{\{\sigma\}}$ for every $k \geq 0$ and $\sigma \in \Sigma^{(k)}$, and $\overline{F}_q = \mathrm{id}$. For the sake of simplicity, let $\Sigma = \{\sigma^{(2)}, \gamma^{(1)}, \alpha^{(0)}\}$.

By Observation 8.8, $S(M) = S(N)$. Thus it suffices to give a rational expression η such that $[\![\eta]\!]|_{T_\Sigma} = S(N)$. By Theorem 6.8, there is such a rational expression $\eta \in \mathrm{RatExp}(\Sigma, Q, \underline{D}(\underline{\mathcal{P}(\Sigma^*)}))$ and it has the form $\eta = \mathrm{id}.q \cdot_q (\eta_q + \mathrm{id}.q))$, where $\eta_q \in \mathrm{RatExp}(\Sigma, Q, \underline{D}(\underline{\mathcal{P}(\Sigma^*)}))$ with $[\![\eta_q]\!] = S(N')^{Q,Q}_q$ and $N' = (Q, \Sigma, \{q\}, \underline{D}(\underline{\mathcal{P}(\Sigma^*)}), \overline{F}, \mu, \nu)$ with $\nu(q)_q = \mathrm{id}$.

Next we give η_q. By Lemma 6.6

$$S(N')^{Q,Q}_q = S(N')^{Q,\emptyset}_q \cdot_q \left(S(N')^{Q,\emptyset}_q\right)^*_q,$$

hence if we find $\theta \in \mathrm{RatExp}(\Sigma, Q, \underline{D}(\underline{\mathcal{P}(\Sigma^*)}))$ such that $[\![\theta]\!] = S(N')^{Q,\emptyset}_q$, then we are ready because $\eta_q = \theta \cdot_q \theta^*_q$ is suitable. Now, by Lemma 6.7,

$$\theta = \mathrm{top}_{\sigma,\mathrm{mul}^{(2)}_{\{\sigma\}}}(\mathrm{id}.q\,,\,\mathrm{id}.q) + \mathrm{top}_{\gamma,\mathrm{mul}^{(1)}_{\{\gamma\}}}(\mathrm{id}.q) + \mathrm{top}_{\alpha,\mathrm{mul}^{(0)}_{\{\alpha\}}}().$$

This finishes the example.

Finally, we strongly conjecture that there is a semiring $\underline{K}$ such that

$$\mathrm{lift}(\mathrm{Rat}(\Sigma, \mathrm{fin}, \underline{D}(\underline{K}))) \setminus \mathrm{lift}(\mathrm{Rat}_{\mathrm{sr}}(\Sigma, \mathrm{fin}, \underline{K}))$$

contains a tree series $\psi\colon T_\Sigma \to K$. As an element of this difference set we might consider the tree series $S(M)$ of Example 3.11 (the semiring here is $\underline{\mathcal{P}(\Delta^*)}$ for some suitable Δ). We think that there is no semiring-rational expression which can express this tree series. In fact, the semiring-rational expressions as they are defined in Definition 3.17 of [7] also implement a particular product order (cf. the discussion in the Introduction), and this is probably too inflexible to express $S(M)$.

Acknowledgement The authors are very grateful to one of the anonymous referees who with his detailed report substantially contributed to this paper.

References

[1] J. Berstel and C. Reutenauer. Recognizable formal power series on trees. *Theoret. Comput. Sci.*, 18(2):115–148, 1982.

[2] S. Bozapalidis. Equational elements in additive algebras. *Theory of Computing Systems*, 32(1):1–33, 1999.

[3] S. Bozapalidis. Context-Free Series on Trees. *Information and Computation*, 169:186–229, 2001.

[4] S. Bozapalidis and A. Grammatikopoulou. Recognizable picture series. *J. Automata, Languages and Combinatorics*, 10:159–183, 2005.

[5] B. Courcelle. Equivalences and transformations of regular systems - applications to recursive program schemes and grammars. *Theoret. Comput. Sci.*, 42:1–122, 1986.

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

[7] M. Droste, Chr. Pech, and H. Vogler. A Kleene theorem for weighted tree automata. *Theory of Computing Systems*, 38:1–38, 2005.

[8] J. Engelfriet. Alternative Kleene theorem for weighted automata. Personal communication, 2003.

[9] J. Engelfriet, Z. Fülöp, and H. Vogler. Bottom-up and top-down tree series transformations. *J. Automata, Languages and Combinatorics*, 7:11–70, 2002.

[10] Z. Ésik and W. Kuich. Formal tree series. *J. of Automata, Languages, and Combinatorics*, 8(2):219–285, 2003.

[11] Z. Fülöp, Zs. Gazdag, and H. Vogler. Hierarchies of tree series transformations. *Theoret. Comput. Sci.*, 314:387–429, 2004.

[12] Z. Fülöp and H. Vogler. Comparison of several classes of weighted tree automata. Technical report, TU Dresden, 2006. TUD-FI06-08-Dez.2006.

[13] D. Giammarresi and A. Restivo. Two-dimensional languages. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages, Part III*, pages 215–268. Springer-Verlag, 1997.

[14] S. E. Kleene. Representation of events in nerve nets and finite automata. In C.E. Shannon and J. McCarthy, editors, *Automata Studies*, pages 3–42. Princeton University Press, Princeton, N.J., 1956.

[15] W. Kuich. Formal power series over trees. In S. Bozapalidis, editor, *3rd International Conference on Developments in Language Theory, DLT 1997, Thessaloniki, Greece, Proceedings*, pages 61–101. Aristotle University of Thessaloniki, 1998.

[16] A. Maletti. Hasse diagrams for classes of deterministic bottom-up tree-to-tree-series transformations. *Theoret. Comput. Sci.*, 339:200–240, 2005.

[17] A. Maletti. Relating tree series transducers and weighted tree automata. *Int. J. of Foundations of Computer Science*, 16(4):723–741, 2005.

[18] A. Maletti. Compositions of tree series transformations. *Theoret. Comput. Sci.*, 2006. unpublished.

[19] I. Mäurer. Rational and recognizable picture series. In *Conference on Algebraic Informatics, Thessaloniki*, April 2005.

[20] I. Mäurer. Characterizations of recognizable picture series. Submitted, 2006.

[21] E. Ochmanski. Regular behaviour of concurrent systems. *Bull. Europ. Assoc. for Theoret. Comp. Sci.*, 27:56–67, 1985.

[22] Chr. Pech. *Kleene-Type Results for Weighted Tree Automata.* PhD thesis, TU Dresden, 2003.

[23] Chr. Pech. Kleene's theorem for weighted tree-automata. In *14th International Symposium on Fundamentals of Computation Theory FCT 2003, Malmö, Sweden*, number 2751 in LNCS, pages 387–399. Springer-Verlag, 2003.

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

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