Characterizing Tree Valuation Languages by Multioperator Weighted Tree Languages

Markus Teichmann** and Johannes Osterholzer*

Faculty of Computer Science, Technische Universität Dresden

Abstract. Weighted tree languages can be a) recognized by weighted tree automata over tree valuation monoids, b) defined by monadic second order logic over tree valuation monoids, c) recognized by weighted tree automata over multioperator monoids, and d) defined by multioperator expressions. We show that the classes a) and b) are characterized by c) and d), respectively, using a special multioperator monoid.

1 Introduction

The fundamental result of Büchi, Elgot, and Trakhtenbrot states the equivalence of *recognizability* by classical finite-state string automata and *definability* by means of formulas of monadic second-order logic (mso). This equivalence result has been extended, in particular, into two directions: (1) from string automata to finite-state tree automata and (2) from the unweighted to the weighted case. In this paper, we focus on weighted tree languages over *tree valuation monoids* (tv-monoid) and over *multioperator monoids* (m-monoid).

A tv-monoid is a commutative monoid $(D, +, 0)$ equipped with a valuation function Val which maps each unranked tree over elements of the carrier set D to an element of D. The weight of a run of a tv-monoid weighted tree automaton (tv-wta) is determined by applying the valuation function to a tree generated from the input tree by replacing every node by the weight of the transition taken at this position. In [1] the Büchi-like characterization has been proved for tv-wta using a syntactically restricted tree valuation weighted mso (tv-mso), i.e. $\mathrm{Rec}(\Sigma, \mathbb{D}) = \mathrm{Def}(\Sigma, \mathbb{D})$ where $\mathbb{D}$ is a product tv-monoid 'with multiplication'.

An m-monoid is a commutative monoid $(A, +, 0)$ equipped with a set Ω of (arbitrary) operations on A. To each transition of an m-monoid weighted tree automaton (m-wta) an operation is assigned which has the same arity as the transition. The weight of a run is obtained by evaluating the operations inductively according to their occurrences in that run. In [2] the Büchi-like characterization has been proved for m-wta using multioperator expressions (m-expressions), i.e. automata and logics are equally powerful: $\mathrm{Rec}(\Sigma, \mathcal{A}) = \mathrm{Def}(\Sigma, \mathcal{A})$.

In this paper, we construct (cf. Constr. 1) for any tv-monoid $\mathcal{D}$ (regular product tv-monoid $\mathbb{D}$) an m-monoid $\mathcal{A}_{\mathcal{D}}$ ($\mathcal{A}_{\mathbb{D}}$) s.t. the following holds (cf. Fig. 1):

1. $\mathrm{Rec}(\Sigma, \mathcal{D}) = \pi_1^2(\mathrm{Rec}(\Sigma, \mathcal{A}_{\mathcal{D}}))$ (cf. Thm. 3) and
2. $\mathrm{Def}(\Sigma, \mathbb{D}) = \pi_1^2(\mathrm{Def}(\Sigma, \mathcal{A}_{\mathbb{D}}))$ (cf. Thm. 5).

* (Partially) **Supported by DFG Graduiertenkolleg 1763 (QuantLA)

$$
\begin{tikzcd}
\text{tv-wta} \arrow[r, "[\![\cdot]\!]"] \arrow[d, leftrightarrow, "\text{related}"] & \mathrm{Rec}(\Sigma, \mathbb{D}) \arrow[r, equal, "{[1,\ \text{Thm. } 5.5]}"] \arrow[d, equal, "\text{Thm. } 3"'] & \mathrm{Def}(\Sigma, \mathbb{D}) \arrow[d, equal, "\text{Thm. } 5"] & \text{tv-mso formulas} \arrow[l, "[\![\cdot]\!]"'] \arrow[d, shift left, "t(\cdot)"] \\
\text{m-wta} \arrow[r, "[\![\cdot]\!]"'] & \pi_1^2(\mathrm{Rec}(\Sigma, \mathcal{A}_{\mathbb{D}})) \arrow[r, equal, "{[2,\ \text{Thm. } 4.1]}"'] & \pi_1^2(\mathrm{Def}(\Sigma, \mathcal{A}_{\mathbb{D}})) & \text{m-expressions} \arrow[l, "[\![\cdot]\!]"] \arrow[u, shift left, "t'(\cdot)"]
\end{tikzcd}
$$

Fig. 1. Overview of proven results (based on [2, p. 275])

2 Main Results

Construction 1. Given a tv-monoid $\mathcal{D} = (D, +, 0, \mathrm{Val})$ we construct the *corresponding m-monoid* $\mathcal{A}_{\mathcal{D}} = (D \times T_D^u, \oplus, (0, 0), \Omega)$ as follows. Let $(T_D^u, \tilde{+}, 0)$ be the monoid over unranked trees labeled by elements of D where the operation $\tilde{+}$ is the additive superposition of two unranked trees from T_D^u using $+$. Then, $\oplus$ is the direct product of $+$ and $\tilde{+}$. We set $\Omega = \{\mathrm{valtop}_d^{(k)} \mid d \in D, k \in \mathbb{N}\}$ and, for $d \in D$, $k \in \mathbb{N}$, and $(d_1, \xi_1), \ldots, (d_k, \xi_k) \in D \times T_D^u$ with $\xi = d(\xi_1, \ldots, \xi_k)$, we define

$$
\mathrm{valtop}_d^{(k)}((d_1, \xi_1), \ldots, (d_k, \xi_k)) = \begin{cases} \big(\mathrm{Val}(\xi), \xi\big) & \text{if } d \neq 0 \text{ and } \\ & \xi_1, \ldots, \xi_k \text{ are 0-free,} \\ (0, 0) & \text{otherwise.} \end{cases}
$$

Definition 2. *Let $\mathcal{N} = (Q, \mu, F)$ be a tv-wta over Σ and $\mathcal{D}$, and $\mathcal{M} = (Q', \delta, F')$ be an m-wta over Σ and $\mathcal{A}_{\mathcal{D}}$. We call $\mathcal{N}$ and $\mathcal{M}$* related *if*

- *$Q = Q'$ and $F = F'$;*
- *for every $k \in \mathbb{N}$, $\sigma \in \Sigma^{(k)}$, and $q, q_1, \ldots, q_k \in Q$ it holds that $\delta_\sigma(q_1 \ldots q_k, q) = \mathrm{valtop}_d$ iff $\mu_\sigma(q_1 \ldots q_k, q) = d$.*

Related automata recognize the same weighted tree language (disregarding the projection) and since the definition can be seen as a construction we conclude:

Theorem 3. *Let $\mathcal{D}$ be a tv-monoid, $\mathcal{A}_{\mathcal{D}}$ the corresponding m-monoid, and Σ be a finite ranked alphabet. The following holds:*

1. *For every tv-wta $\mathcal{N}$ over Σ and $\mathcal{D}$ there is an m-wta $\mathcal{M}$ over Σ and $\mathcal{A}_{\mathcal{D}}$ such that $[\![\mathcal{N}]\!] = \pi_1^2([\![\mathcal{M}]\!])$.*
2. *For every m-wta $\mathcal{M}$ over Σ and $\mathcal{A}_{\mathcal{D}}$ there is a tv-wta $\mathcal{N}$ over Σ and $\mathcal{D}$ such that $\pi_1^2([\![\mathcal{M}]\!]) = [\![\mathcal{N}]\!]$.*

On the level of the logics we provide transformation functions $t(\cdot)$ and $t'(\cdot)$ between $\mathrm{Def}(\Sigma, \mathbb{D})$ and $\mathrm{Def}(\Sigma, \mathcal{A}_{\mathbb{D}})$ and vice versa, respectively. We only sketch the transformation $t(\cdot)$:

Construction 4. Let φ be a syntactically restricted[1] tv-mso formula over Σ and a regular product tv-monoid $\mathbb{D}$. We construct the *corresponding m-expression $t(\varphi)$ over Σ and $\mathcal{A}_{\mathbb{D}}$* inductively on the structure of φ:

[1] formally, φ is $\forall$-restricted and strongly $\wedge$-restricted

– For every Boolean tv-mso formula β: $t(\beta) = \beta \triangleright t(1)$.
– For every $d \in D$: Since $\mathbb{D}$ is regular, there is a tv-wta $\mathcal{N}_d = (Q, \mu, F)$ which assigns d to every input tree. The formula $t(d)$ simulates the semantics of this automaton.
– $t(\varphi_1 \vee \varphi_2) = t(\varphi_1) + t(\varphi_2)$.
– $\varphi = \varphi_1 \wedge \varphi_2$ is strongly $\wedge$-restricted:
  - φ_1 or φ_2 Boolean: Assume φ_1 is Boolean then $t(\varphi_1 \wedge \varphi_2) = \varphi_1 \triangleright t(\varphi_2)$
  - φ_1 and φ_2 are almost Boolean, i.e. $\mathrm{step}(\varphi_i) = (a_1^i, \psi_1^i) \dots (a_{n_i}^i, \psi_{n_i}^i)$ then we set $t(\varphi_1 \wedge \varphi_2) = \sum^{+}_{\substack{i \in [n_1] \\ j \in [n_2]}} (\psi_i^1 \wedge \psi_j^2) \triangleright t\big(a_i^1 \diamond a_j^2\big)$.
– $t(\exists x.\psi) = \sum_x t(\psi)$
– $t(\exists X.\psi) = \sum_X t(\psi)$
– $\varphi = \forall x.\psi$ is $\forall$-restricted: ψ almost Boolean, i.e. $\mathrm{step}(\psi) = (d_1, \psi_1) \dots (d_n, \psi_n)$. Let $\mathcal{U} = \{X_1, \dots, X_n\}$ and ω^ψ be a $\Sigma_{\mathcal{U}}$-family such that for every $U \subseteq \mathcal{U}$, $k \in \mathbb{N}$, and $\sigma \in \Sigma^{(k)}$ we have $(\omega^\psi)_{(\sigma,U)} = \mathrm{valtop}^{(k)}_{d_U}$ where $d_U = \sum_{\substack{i \in [n] \\ X_i \in U}} d_i$.

  We define, similarly to the proof in [2, Lemma 5.10],

$$t(\forall x.\psi) = \textstyle\sum_{X_1} \dots \sum_{X_n} \big(\forall x.(\bigwedge_{i \in [n]} (x \in X_i) \Leftrightarrow \psi_i)\big) \triangleright H(\omega^\psi) \ .$$

The transformation $t'(\cdot)$ in the opposite direction uses similar considerations for common syntactic constructs like disjunction. The semantics of the structure $H(\omega)$ is simulated using a more complex formula enumerating symbol-variable combinations. Using both transformation functions we can conclude:

Theorem 5. *Let $\mathbb{D}$ be a regular product tv-monoid, $\mathcal{A}_{\mathbb{D}}$ the corresponding m-monoid, and Σ a finite alphabet. Then the following holds:*

1. *For every $\forall$-restricted and strongly $\wedge$-restricted tv-mso formula φ there is an m-expression $t(\varphi)$ over Σ and $\mathcal{A}_{\mathbb{D}}$ such that $[\![\varphi]\!] = \pi_1^2([\![t(\varphi)]\!])$.*
2. *For every m-expression e over Σ and $\mathcal{A}_{\mathbb{D}}$ there is a tv-mso formula $t'(e)$ over Σ and $\mathbb{D}$ such that $\pi_1^2([\![e]\!]) = [\![t'(e)]\!]$.*

An alternative proof for the Büchi-like theorem [1, Thm. 5.5] arises utilizing our result and the Büchi-characterization [2, Thm. 4.1] (cf. Fig. 1).

References

1. M. Droste, D. Götze, S. Märcker, and I. Meinecke. Weighted tree automata over valuation monoids and their characterization by weighted logics. In *Algebraic Foundations in Computer Science*, volume 7020 of *Lecture Notes in Comput. Sci.*, pages 30–55. Springer Berlin Heidelberg, 2011.
2. Z. Fülöp, T. Stüber, and H. Vogler. A Büchi-like theorem for weighted tree automata over multioperator monoids. *Theory Comput. Syst.*, 50(2):241–278, 2012. Published online 28.10.2010.