Weighted Tree Automata I.– A Kleene theorem for wta over semirings

Zoltán Fülöp

University of Szeged

Department of Foundations of Computer Science

fulop@inf.u-szeged.hu

November 3, 2009

Contents

Trees (= terms)

Ranked alphabet : (Σ, rank) with $\mathrm{rank} : \Sigma \rightarrow \mathbb{N}$

$\Sigma^{(k)} = \{\sigma \in \Sigma \mid \mathrm{rank}(\sigma) = k\}$

The set of trees (terms) over Σ and a set Z is the smallest set U satisfying:

(i) $\Sigma^{(0)} \cup Z \subseteq U$,

(ii) if $k \geq 1$, $\sigma \in \Sigma^{(k)}$, $t_1, \ldots, t_k \in T_\Sigma(Z)$, then $\sigma(t_1, \ldots, t_k) \in U$.

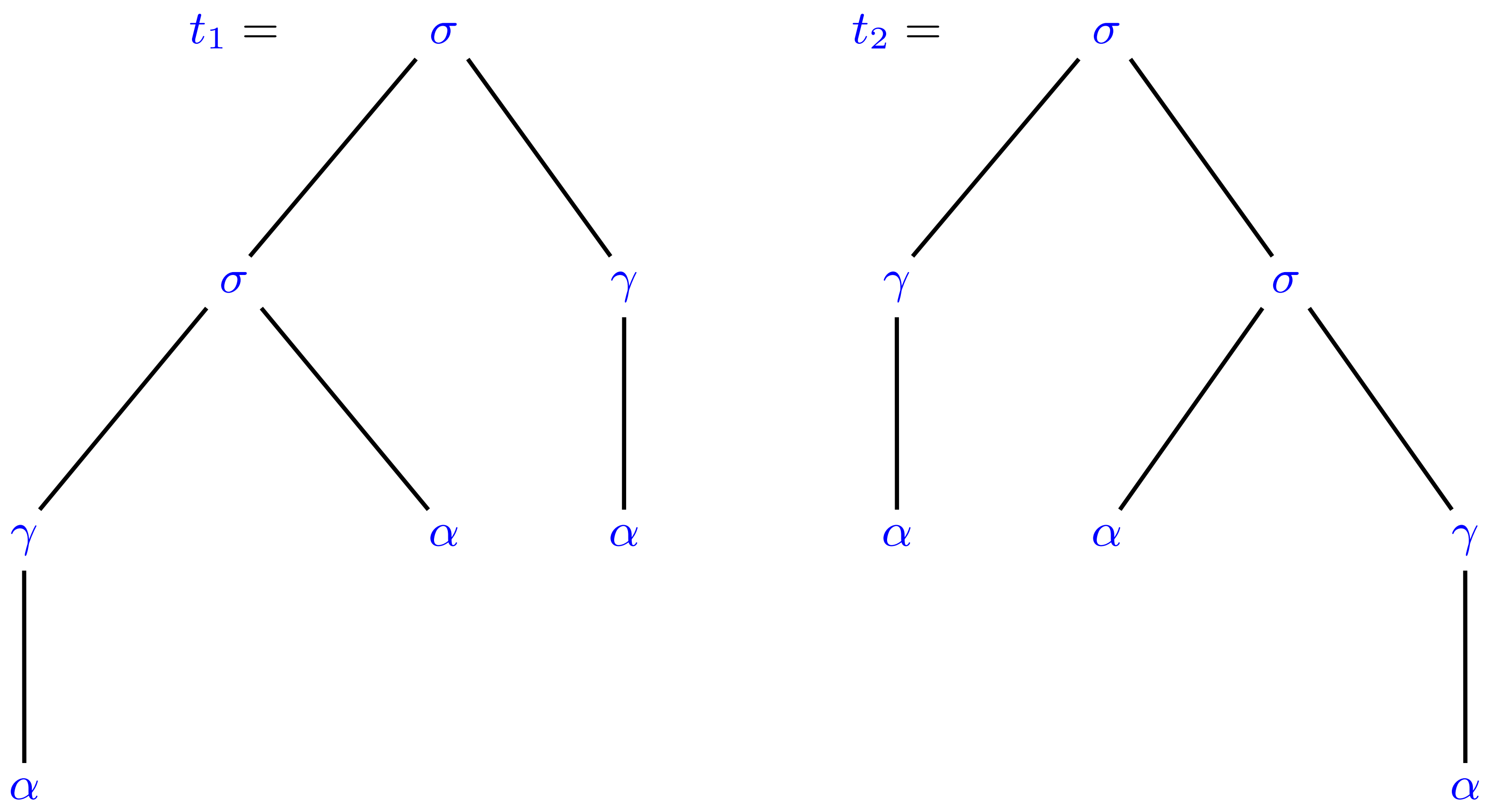
We denote this set by $T_\Sigma(Z)$

Note: $T_\Sigma(Z) = \emptyset$ iff $\Sigma^{(0)} \cup Z = \emptyset$.

Tree language : $L \subseteq T_\Sigma(Z)$ (or: $L : T_\Sigma(Z) \rightarrow \{0, 1\}$) .

Trees (= terms)

Example: $\Sigma = \{\sigma^{(2)}, \gamma^{(1)}, \alpha^{(0)}\}$, $Z = \emptyset$



$t_1 = \sigma(\sigma(\gamma(\alpha), \alpha), \gamma(\alpha))$

$t_2 = \sigma(\gamma(\alpha), \sigma(\alpha, \gamma(\alpha)))$

Terms ($=$ trees)

Positions in trees:

$\mathrm{pos} : T_\Sigma(Z) \to \mathcal{P}(\mathbb{N}^*)$ such that, for every $t \in T_\Sigma(Z)$,

(i) if $t \in (\Sigma^{(0)} \cup Z)$, then $\mathrm{pos}(t) = \{\varepsilon\}$

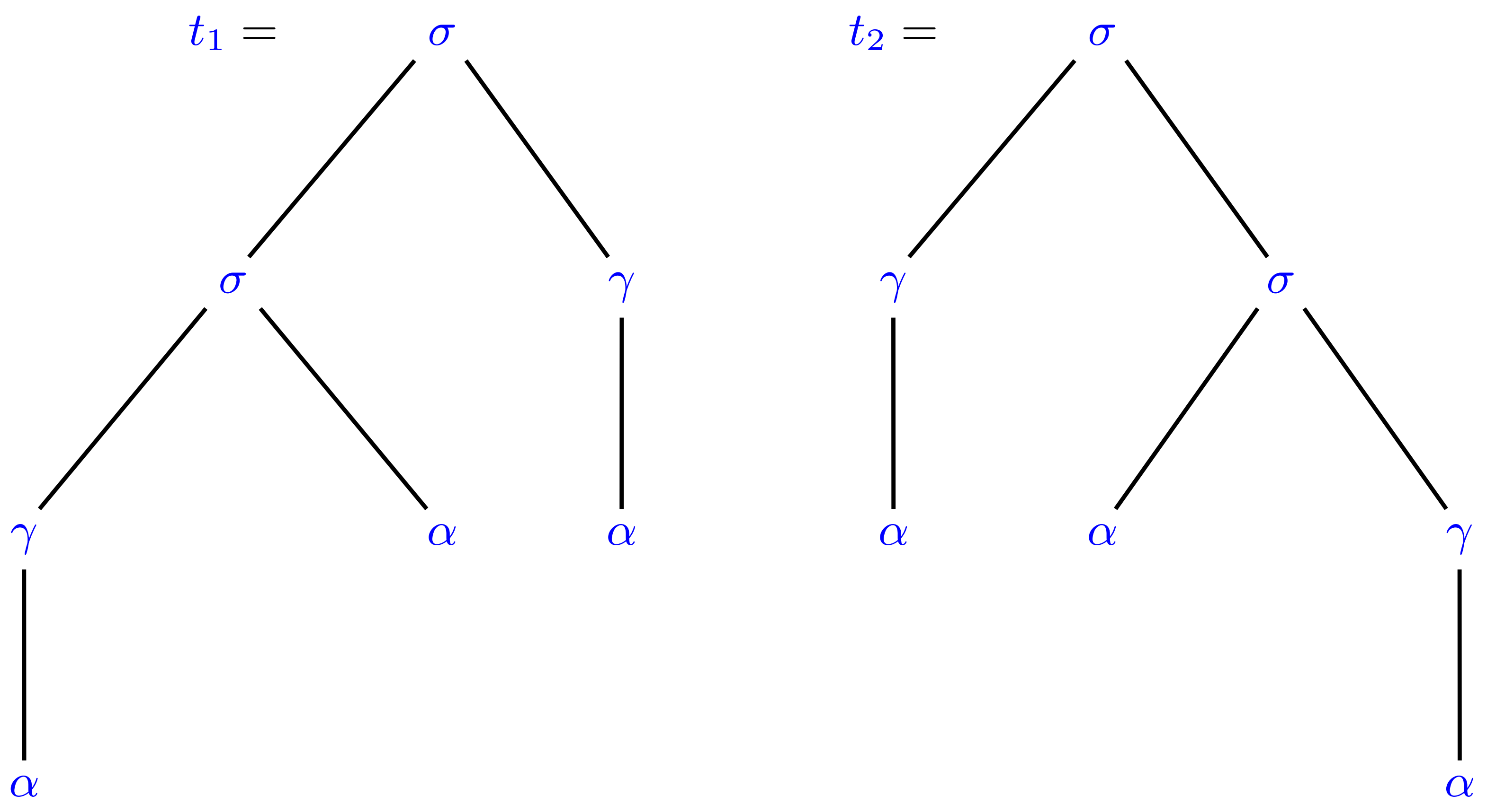
(ii) if $t = \sigma(t_1, \ldots, t_k)$, then

$$\mathrm{pos}(t) = \{\varepsilon\} \cup \{iw \mid 1 \leq i \leq k, w \in \mathrm{pos}(t_i)\}.$$

The label of a tree $t \in T_\Sigma(Z)$ at position w is denoted by $t(w)$.

The height of a tree is denoted be $\mathrm{height}(t)$.

Trees (= terms)



$\mathrm{pos}(t_1) = \{\varepsilon, 1, 11, 111, 12, 2, 21\}$

$t_1(\varepsilon) = \sigma, t_1(11) = \gamma, t_1(12) = \alpha$

$\mathrm{pos}(t_2) = \{\varepsilon, 1, 11, 2, 21, 22, 221\}$

$t_2(2) = \sigma, t_2(22) = \gamma$

Tree Automata

Syntax

A *tree automaton (over Σ and Z)* is a tuple $M = (Q, \Sigma, Z, F, \delta, \nu)$, where

- Q is a finite set (*states*),
- Σ is a ranked alphabet (*input ranked alphabet*),
- Z is a finite set (*variables*),
- $F \subseteq Q$ is a set (*final states*), and
- δ is a family $(\delta_k | k \geq 0)$ of mappings, where $\delta_k \subseteq Q^k \times \Sigma^{(k)} \times Q$ (*transitions*),
- $\nu : Z \to \mathcal{P}(Q)$ is a mapping (*the variate assignment*).

Note: a transition has the form $(q_1, \ldots, q_k, \sigma, q)$.

Tree Automata

Semantics

$M = (Q, \Sigma, Z, F, \delta, \nu)$ a tree automaton, $t \in T_\Sigma(Z)$

- a run of M on t is a mapping $r : \mathrm{pos}(t) \to Q$ such that for every $w \in \mathrm{pos}(t)$ we have

  - if $t(w) = z$, for some $z \in Z$, then $r(w) \in \nu(z)$,
  - otherwise (if $t(w) = \sigma$ for some $\sigma \in \Sigma^{(k)}, k \geq 0$), then $(r(w1), \ldots, r(wk), t(w), r(w)) \in \delta_k$

- a run r on t is successful if $r(\varepsilon) \in F$

- the set of successful runs of M on t is $R_M(t)$

The *tree language recognized by* M is

$$L_M = \{t \in T_\Sigma(Z) \mid R_M(t) \neq \emptyset\}.$$

Tree Automata

Example

$\Sigma = \{\sigma^{(2)}, \gamma^{(1)}, \alpha^{(0)}\}, Z = \emptyset$, the tree language

$$L = \{s \in T_\Sigma \mid \sigma(\bullet, \alpha) \text{ occurs in } s\} \text{ is recognizable}$$

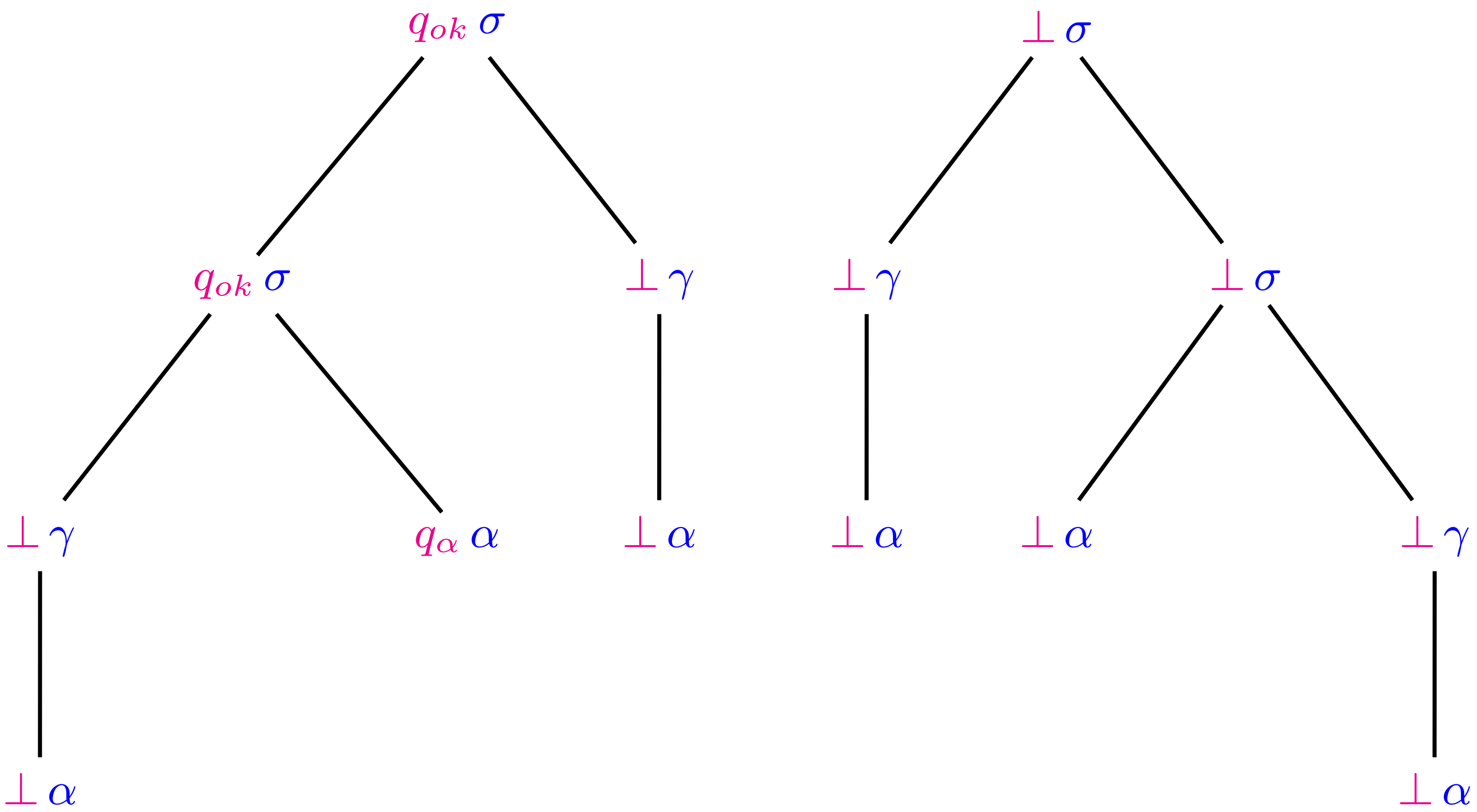
Let $M = (Q, \Sigma, F, \delta)$, where

- $Q = \{\perp, q_\alpha, q_{ok}\}$,
- $F = \{q_{ok}\}$,
- - $\delta_0 : (\alpha, \perp), (\alpha, q_\alpha)$,
  - $\delta_2 : (\perp, q_\alpha, \sigma, q_{ok}), (\perp, q_{ok}, \sigma, q_{ok}), (q_{ok}, \perp, \sigma, q_{ok}), (\perp, \perp, \sigma, \perp)$,
  - $\delta_1 : (q_{ok}, \gamma, q_{ok})$, $(\perp, \gamma, \perp)$.

Then $L_M = L$.

Tree Automata

Example

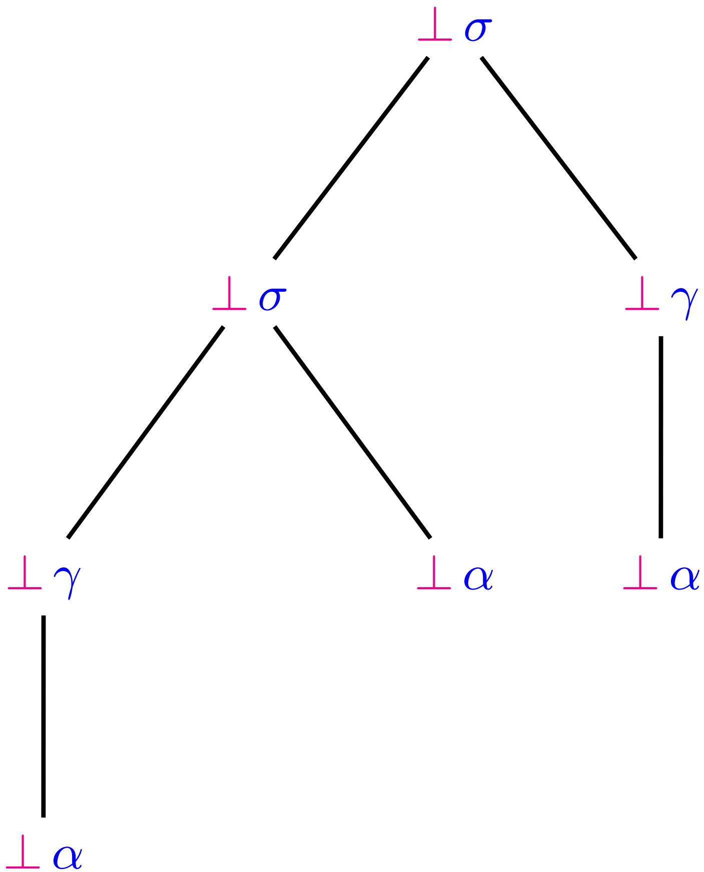


A successful run.

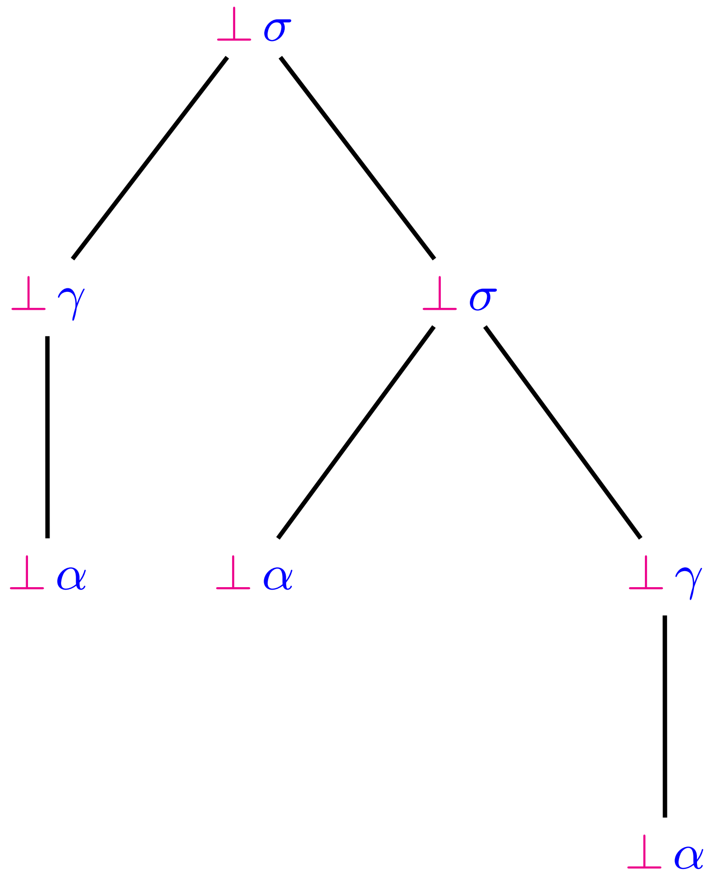
A not successful run.

Tree Automata

Example



A not successful run.



A not successful run.

Semirings

Semiring : $(K, +, \cdot, 0, 1)$

- $(K, +, 0)$ is a commutative monoid,
- $(K, \cdot, 1)$ is a monoid,

  and for every $a, b, c \in K$:
$$(a + b) \cdot c = (a \cdot c) + (b \cdot c)$$
$$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$
$$a \cdot 0 = 0 \cdot a = 0.$$

K is commutative if $(K, \cdot, 1)$ is a commutative monoid.

Examples :

- Boolean semiring : $\mathbb{B} = (\{0, 1\}, \vee, \wedge, 0, 1)$
- semiring of natural numbers : $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$
- tropical semiring : $\mathrm{Trop} = (\mathbb{N} \cup \{\infty\}, min, +, \infty, 0)$
- arctic semiring : $\mathrm{Arct} = (\mathbb{N} \cup \{-\infty\}, max, +, -\infty, 0)$

Tree series

(Tree language : $L : T_\Sigma(Z) \to \{0,1\}$)

Tree series : $S : T_\Sigma(Z) \to K$, where $(K, +, \cdot, 0, 1)$ is a semiring

Examples of tree series:

$\mathrm{height} : T_\Sigma \to \mathbb{N}$, in $\underline{\mathrm{Arct}} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$

$\mathrm{size}_\sigma : T_\Sigma \to \mathbb{N}$, in $\underline{\mathbb{N}} = (\mathbb{N}, +, \cdot, 0, 1)$

$\mathrm{size} : T_\Sigma \to \mathbb{N}$, in $\underline{\mathbb{N}} = (\mathbb{N}, +, \cdot, 0, 1)$

$\#_{\sigma(\bullet,\alpha)} : T_\Sigma \to \mathbb{N}$, in $\underline{\mathbb{N}} = (\mathbb{N}, +, \cdot, 0, 1)$

$\mathrm{shortest}_\alpha : T_\Sigma \to \mathbb{N}$, in $\underline{\mathrm{Trop}} = (\mathbb{N} \cup \{-\infty\}, \min, +, -\infty, 0)$

$\mathrm{yield} : T_\Sigma \to \mathcal{P}(\Sigma^*)$, in $\mathrm{Lang}_\Sigma = (\mathcal{P}(\Sigma^*), \cup, \cdot, \emptyset, \{\varepsilon\})$

$\mathrm{pos} : T_\Sigma \to \mathcal{P}(\mathbb{N}^*)$, in $\mathrm{Lang}_\mathbb{N}$

$\mathrm{pos}_{\sigma(\bullet,\alpha)} : T_\Sigma \to \mathcal{P}(\mathbb{N}^*)$, in $\mathrm{Lang}_\mathbb{N}$

Weighted tree automata (wta) over semirings

Syntax

A wta (over Σ, Z and K) is a system $M = (Q, \Sigma, Z, K, F, \delta, \nu)$, where

- K is a commutative semiring,
- $F : Q \to K$ is the root weight,
- $\delta = (\delta_k \mid k \geq 0)$ is the family of transition mappings, where $\delta_k : Q^k \times \Sigma^{(k)} \times Q \to K$,
- $\nu : Z \times Q \to K$ is the variable assignment.

Note: $\delta(q_1, \ldots, q_k, \sigma, q) \in K$ is the weight of the transition $(q_1, \ldots, q_k, \sigma, q)$.

Wta over semirings

<u>Semantics</u>

$M = (Q, \Sigma, Z, K, F, \delta, \nu)$ a wta, $t \in T_\Sigma(Z)$

- a run of M on t is a mapping $r : \mathrm{pos}(t) \to Q$

- the set of runs of M on t is $R_M(t)$

- for $w \in \mathrm{pos}(t)$, the weight $\mathrm{wt}(t, r, w)$ of w in t under r

  - if $t(w) = z$ for some $z \in Z$, then $\mathrm{wt}(t, r, w) = \nu(z, r(w))$
  - otherwise (if $t(w) = \sigma$ for some $\sigma \in \Sigma^{(k)}, k \geq 0$) $\mathrm{wt}(t, r, w) = \delta_k(r(w1), \ldots, r(wk), t(w), r(w))$
  - the weight of r is $\mathrm{wt}(t, r) = \prod_{w \in pos(t)} \mathrm{wt}(t, r, w)$.

The tree series $S_M : T_\Sigma(Z) \to K$ recognized by M is defined by

$$S_M(t) = \sum_{r \in R_M(t)} \mathrm{wt}(t, r) \cdot F(r(\varepsilon)).$$

Wta over semirings

Example

$\Sigma = \{\sigma^{(2)}, \gamma^{(1)}, \alpha^{(0)}\}, Z = \emptyset$, the semiring is $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$

The tree series $\#_{\sigma(\bullet,\alpha)} : T_\Sigma \to \mathbb{N}$ is recognizable.

Let $M = (Q, \Sigma, \mathbb{N}, F, \delta)$ the wta, where

- $Q = \{\bot, q_\alpha, q_{ok}\}$,
- $F(\bot) = 0, F(q_\alpha) = 0, F(q_{ok}) = 1$,
- - $\delta_0(\alpha, \bot) = \delta_0(\alpha, q_\alpha) = 1$,
  - $\delta_2(\bot, q_\alpha, \sigma, q_{ok}) = \delta_2(\bot, q_{ok}, \sigma, q_{ok}) = \delta_2(q_{ok}, \bot, \sigma, q_{ok}) = \delta_2(\bot, \bot, \sigma, \bot) = 1$,
  - $\delta_1(q_{ok}, \gamma, q_{ok}) = \delta_1(\bot, \gamma, \bot) = 1$

Then $S_M = \#_{\sigma(\bullet,\alpha)}$.

Wta over semirings

Example

$\Sigma = \{\sigma^{(2)}, \alpha^{(0)}\}$, the is a semiring $\mathrm{Arct} = (\mathbb{N} \cup \{-\infty\}, max, +, -\infty, 0)$

The wta $M = (Q, \Sigma, \mathrm{Arct}, F, \delta)$ recognizes the tree series height, where

- $Q = \{p_1, p_2\}$,
- $F(p_1) = 0$ and $F(p_2) = -\infty$.

Moreover, let

$$\begin{array}{lllll} \delta_0(\alpha, p_1) & = & \delta_0(\alpha, p_2) & = & 0, \\ \delta_2(p_1, p_2, \sigma, p_1) & = & \delta_2(p_2, p_1, \sigma, p_1) & = & 1, \\ \delta_2(p_2, p_2, \sigma, p_2) & = & 0, & & \end{array}$$

and for every other transition (q_1, q_2, σ, q) we have $\delta_2(q_1, q_2, \sigma, q) = -\infty$.

Then $S_M = \mathrm{height}$.

Tree automata = wta over the Boolean semiring $\mathbb{B}$

$\mathbb{B} = (\{0,1\}, \vee, \wedge, 0, 1)$ is the Boolean semiring

A wta (over Σ, Z and $\mathbb{B}$) has the form $M = (Q, \Sigma, Z, \mathbb{B}, F, \delta, \nu)$, where

- $F : Q \to \{0,1\}$ is the root weight,
- $\delta = (\delta_k \mid k \geq 0)$ is a family of transition mappings, where $\delta_k : Q^k \times \Sigma^{(k)} \times Q \to \{0,1\}$,
- $\nu : Z \times Q \to \{0,1\}$ is the variable assignment.

$t \in T_\Sigma(Z)$ is a tree, $r : \mathrm{pos}(t) \to Q$ is a run on t

The weight of r is $\mathrm{wt}(t,r) = \prod_{w \in pos(t)} \mathrm{wt}(t,r,w)$.

The tree series $S_M : T_\Sigma(Z) \to \{0,1\}$ recognized by M is defined by

$$S_M(t) = \sum_{r \in R_M(t)} \mathrm{wt}(t,r) \cdot F(r(\varepsilon)).$$

Wta over semirings

We denote the class of tree series recognizable by wta over Σ, Z and K by

$$\mathrm{Rec}(\Sigma, Z, K).$$

Tree series

For a tree series $S : T_\Sigma(Z) \to K$ and $t \in T_\Sigma(Z)$, we write (S, t) for $S(t)$.

We write S in the form $S = \sum_{t \in T_\Sigma(Z)} (S, t).t$.

The set of tree series over Σ, Z, and K is denoted by $K\langle\langle T_\Sigma(Z)\rangle\rangle$.

The *support* of S is $\mathrm{supp}(S) = \{t \in T_\Sigma(Z) \mid (S, t) \neq 0\}$.

The tree series S is *polynomial* if $\mathrm{supp}(S)$ is finite.

We write a polynomial tree series S in the form $S = a_1.t_1 + \ldots + a_n.t_n$, where $\mathrm{supp}(S) = \{t_1, \ldots, t_n\}$ and $(S, t_i) = a_i$.

The set of polynomial tree series over Σ, Z, and K is denoted by $K\langle T_\Sigma(Z)\rangle$.

Constant tree series: $\exists(a \in K) : (S, t) = a$ for all $t \in T_\Sigma(Z)$;
it is also denoted by $\widetilde{a}$.

Operations on trees series

K is a (commutative) semiring.

Let $a \in K$, and $S, T \in K\langle\langle T_\Sigma(Z)\rangle\rangle$

- scalar multiplication: $(aS, t) = a \cdot (S, t)$
- sum: $(S + T, t) = (S, t) + (T, t)$

for $t \in T_\Sigma(Z)$.

Let $\sigma \in \Sigma^{(k)}, k \geq 0$, and $S_1, \ldots, S_k \in K\langle\langle T_\Sigma(Z)\rangle\rangle$

- top concatenation: $(\mathrm{top}_\sigma(S_1, \ldots, S_k), t) = (S_1, t_1) \cdot \ldots \cdot (S_k, t_k)$ if $t = \sigma(t_1, \ldots, t_k)$ and $(\mathrm{top}_\sigma(S_1, \ldots, S_k), t) = 0$ otherwise.

Note: $\mathrm{top}_\alpha = 1.\alpha$ for $\alpha \in \Sigma^{(0)}$.

Operations on trees series

Let $t \in T_\Sigma(Z)$ and $S, T \in K\langle\langle T_\Sigma(Z)\rangle\rangle$

- z-concatenation: $t \circ_z T$

$$t \circ_z T = \begin{cases} T & \text{if } t = z \\ 1.z' & \text{if } t = z' \neq z \\ \mathrm{top}_\sigma(t_1 \circ_z T, \ldots, t_k \circ_z T) & \text{if } t = \sigma(t_1, \ldots, t_k) \end{cases}$$

- and $S \circ_z T = \sum_{t \in T_\Sigma(Z)} (S, t)(t \circ_z T)$

- the mth z iteration: $S_z^0 = \widetilde{0}$ and $S_z^{m+1} = S_z^m \circ_z S + 1.z$

Operations on trees series

A tree series $S \in K\langle\langle T_\Sigma(Z)\rangle\rangle$ is z-proper, if $(S, z) = 0$.

If S is z-proper, then $(S_z^{m+1}, t) = (S_z^m, t)$ for any $m \geq \text{height}(t) + 1$ and $t \in T_\Sigma(Z)$.

For $S \in K\langle\langle T_\Sigma(Z)\rangle\rangle$ we define the z-Kleene star S_z^* of S as follows: if S is z-proper, then $(S_z^*, t) = (S_z^{\text{height}(t)+1}, t)$, otherwise $S_z^* = \tilde{0}$.

Rational (regular) operations: scalar multiplication, sum, $\text{top}_\sigma, (\sigma \in \Sigma)$, z-concatenation, and z-Kleene star $(z \in Z)$.

Rational expressions and their semantics

The set of *rational tree series expressions over* Σ, Z *and* K, denoted by $\mathrm{RatExp}(\Sigma, Z, K)$, is the smallest set R which satisfies Conditions (1)-(6). For every $\eta \in \mathrm{RatExp}(\Sigma, Z, K)$ we define $[\![\eta]\!] \in K\langle\langle T_\Sigma(Z)\rangle\rangle$ simultaneously.

1. For every $z \in Z$, the expression $z \in R$, and $[\![z]\!] = 1.z$.
2. For every $k \geq 0$, $\sigma \in \Sigma^{(k)}$, and $\eta_1, \ldots, \eta_k \in R$, the expression $\sigma(\eta_1, \ldots, \eta_k) \in R$ and $[\![\sigma(\eta_1, \ldots, \eta_k)]\!] = \mathrm{top}_\sigma([\![\eta_1]\!], \ldots, [\![\eta_k]\!])$.
3. For every $\eta \in R$ and $a \in K$, the expression $(a\eta) \in R$ and $[\![(a\eta)]\!] = a[\![\eta]\!]$.
4. For every $\eta_1, \eta_2 \in R$, the expression $(\eta_1 + \eta_2) \in R$ and $[\![(\eta_1 + \eta_2)]\!] = [\![\eta_1]\!] + [\![\eta_2]\!]$.
5. For every $\eta_1, \eta_2 \in R$ and $z \in Z$, the expression $(\eta_1 \circ_z \eta_2) \in R$ and $[\![(\eta_1 \circ_z \eta_2)]\!] = [\![\eta_1]\!] \circ_z [\![\eta_2]\!]$.
6. For every $\eta \in R$ and $z \in Z$, the expression $(\eta_z^*) \in R$ and $[\![(\eta_z^*)]\!] = [\![\eta]\!]_z^*$.

Rational tree series

A tree series $S \in K\langle\langle T_\Sigma(Z)\rangle\rangle$ is a rational (over Σ, Z and K) if there is an $\eta \in \mathrm{RatExp}(\Sigma, Z, K)$ such that $S = [\![\eta]\!]$. The *class of all rational tree series over* Σ*,* Z *and* K is denoted by $\mathrm{Rat}(\Sigma, Z, K)$.

Note: every polynomial is a rational tree series (note that $\widetilde{0} = [\![0\alpha]\!]$ for any $\alpha \in \Sigma^{(0)}$). Thus $\mathrm{Rat}(\Sigma, Z, K)$ is the smallest subclass of $K\langle\langle T_\Sigma(Z)\rangle\rangle$ that contains $K\langle T_\Sigma(Z)\rangle$, and is closed under the rational operations.

Rational tree series

Example

Let $\Sigma = \{\sigma^{(2)}, \gamma^{(1)}, \alpha^{(0)}\}, Z = \{z\}$. We show that $\#_{\sigma(\bullet,\alpha)} \in \mathrm{Rat}(\Sigma, Z, \mathbb{N})$.

We define the rational expressions $\eta, \eta_1, \eta_2 \in \mathrm{RatExp}(\Delta, Z, \mathbb{N})$ by

$$\begin{array}{lcl} \eta & = & \eta_1 \circ_z \sigma(z,\alpha) \circ_z \eta_2 \\ \eta_1 & = & \big(\gamma(z) + \sigma(\eta_2, z) + \sigma(z, \eta_2)\big)_z^* \\ \eta_2 & = & \big(\gamma(z) + \sigma(z,z)\big)_z^* \circ_z \alpha \end{array}$$

It is obvious that $[\![\eta_1]\!], [\![\eta_2]\!] \in \mathbb{N}\langle\!\langle T_\Sigma(Z)\rangle\!\rangle$ with $[\![\eta_1]\!] = 1_{(\mathbb{N}, C_\Sigma)}$ and $[\![\eta_2]\!] = 1_{(\mathbb{N}, T_\Sigma(Z))}$. Then $[\![\eta]\!]|_{T_\Sigma} = \#_{\sigma(z,\alpha)}$.

Recognizable $\Rightarrow$ rational

Let $M = (Q, \Sigma, Z, K, F, \delta, \nu)$ a wta with Boolean root weight.

We will show that $S_M \in \mathrm{Rat}(\Sigma, Z \cup Q, K)$.

For every $P \subseteq Q$ and $q \in Q$ we define the tree series $S_M(P, q) \in K\langle\langle T_\Sigma(Z \cup Q)\rangle\rangle$ such that for every $t \in T_\Sigma(Z \cup Q)$,

$$(S_M(P,q), t) = \begin{cases} \sum_{r \in R_M^P(t,q)} \mathrm{wt}(t, r) & \text{if } t \in T_\Sigma(Z \cup Q) \setminus Q \\ 0 & \text{if } t \in Q \end{cases}$$

where $R_M^P(t, q)$ is the set of all those runs $r \in R_M(t)$ for which (i) $r(\varepsilon) = q$, (ii) $r(w) \in P$ for every $w \in \mathrm{pos}(t) \setminus (\mathrm{pos}_Q(t) \cup \{\varepsilon\})$, and (iii) $r(w) = t(w)$ for every $w \in \mathrm{pos}_Q(t)$.

(Here $\mathrm{wt}(t, r, w) = 1$ if $t(w) \in Q$.)

Recognizable $\Rightarrow$ rational

Lemma 1. Let $P \subseteq Q$, $q \in Q$, and $p \in Q \setminus P$. Then

$$S_M(P \cup \{p\}, q) = S_M(P, q) \circ_p S_M(P, p)_p^*.$$

Lemma 2. $S_M = [\![\eta]\!]|_{T_\Sigma(Z)}$ for some $\eta \in \mathrm{Rat}(\Sigma, Z \cup Q, K)$.

We prove by induction on $|P|$ that for every $P \subseteq Q$ and $q \in Q$, the tree series $S_M(P, q)$ is in $\mathrm{Rat}(\Sigma, Z \cup Q, S)$.

For the induction base, i.e., $P = \emptyset$, we have

$$S_M(\emptyset, q) = \sum_{\substack{k \geq 0, \sigma \in \Sigma^{(k)} \\ q_1, \ldots, q_k \in Q}} \delta_k(q_1 \ldots q_k, \sigma, q).\sigma(q_1, \ldots, q_k),$$

which is a polynomial, and hence $S_M(\emptyset, q)$ is rational. ...

Recognizable $\Rightarrow$ rational

...

For the induction step, we assume that $S_M(P, q)$ is rational for every $q \in Q$. Now let $p \in Q \setminus P$. Then it follows from the Lemma 1. that also $S_M(P \cup \{p\}, q)$.

Finally

$$S_M = \sum_{\substack{q \in Q \\ F(q)=1}} (\ldots(S_M(Q, q) \circ_{q_1} \widetilde{0}) \circ_{q_2} \widetilde{0} \ldots) \circ_{q_n} \widetilde{0}|_{T_\Sigma(Z)}.$$

Theorem. $\mathrm{Rec}(\Sigma, Z, K) \subseteq \mathrm{Rat}(\Sigma, \mathrm{fin}, K)|_{T_\Sigma(Z)}$.

Rational $\Rightarrow$ recognizable

Facts:

1) Monomials are in $\mathrm{Rec}(\Sigma, Z, K)$

2) $\mathrm{Rec}(\Sigma, Z, K)$ is closed under rational operations. (Many details, commutativity is used!)

Hence $\mathrm{Rec}(\Sigma, Z, K)$ is a class of tree series that contains polynomials and is closed under rational operations.

Theorem. $\mathrm{Rat}(\Sigma, Z, K) \subseteq \mathrm{Rec}(\Sigma, Z, K)$.

Normal form theorems for wta

Input: $M = (Q, \Sigma, Z, K, F, \delta, \nu)$

Output: $M' = (Q', \Sigma, Z, K, F', \delta', \nu)$ such that $S_M = S_{M'}$ and there is a state $q_f \in Q'$ such that

- $F'(q_f) = 1$ and $F'(q) = 0$ for every $q \neq q_f$,
- if $q_i = q_f$, then $\delta'(q_1, \ldots, q_k, \sigma, q) = 0$.

Construction: $Q' = Q \cup \{q_f\}$, where q_f is a new state.

For every $\sigma \in \Sigma^{(k)}$, $w \in (Q')^k$, and $q \in Q'$, we define

$$\delta'_k(w, \sigma, q) = \begin{cases} \delta_k(w, \sigma, q) & \text{if } w \in Q^k, q \in Q \\ \sum_{q \in Q} \delta_k(w, \sigma, q) \cdot F(q) & \text{if } w \in Q^k, q = q_f \\ 0 & \text{otherwise.} \end{cases}$$

Normal form theorems for wta

Input: $M = (Q, \Sigma, Z, K, F, \delta, \nu), z \in Z$, s.t. S_M is z-proper.

Output: $M' = (Q', \Sigma, Z, K, F', \delta', \nu')$ such that $S_M = S_{M'}$ and there is a state $q_0 \in Q'$ such that

- $\nu(z, q_0) = 1$,
- for every $q_0 \neq q \in Q$: $\nu(z, q) = 0$,
- for every $\sigma \in \Sigma^{(k)}$ and $z' \neq z$: $\delta'_k(\ldots, \sigma, q_0) = \nu'(z', q_0) = 0$.

References

[1] A. Alexandrakis and S. Bozapalidis. Weighted grammars and Kleene's theorem. *Information Processing Letters*, 24(1):1–4, January 1987.

[2] M. Droste, Chr. Pech, and H. Vogler. A Kleene theorem for weighted tree automata. *Theory Comput. Syst.*, 38:1–38, 2005.

[3] Z. Ésik and W. Kuich. Formal Tree Series. *Journal of Automata, Languages and Combinatorics*, 8:219–285, 2003.

[4] Z. Fülöp, H. Vogler, Weighted tree automata and tree transducers, in: M. Droste, W. Kuich, H. Vogler (Eds.), Handbook of Weighted Automata, Chapter 9, Springer-Verlag 2009.