Introduction to *Weighted Tree Automata* and *Tree Series Transducers*

Andreas Maletti

Technische Universität Dresden
Department of Computer Science

June 22, 2006

Table of Contents

Tree

Definition (Ranked, labelled, ordered trees)

Let Σ ranked alphabet and X set. Smallest set T such that

- $X \subseteq T$
- $\sigma(t_1, \ldots, t_k) \in T$ for every $k \in \mathbb{N}$, $\sigma \in \Sigma_k$, and $t_1, \ldots, t_k \in T$

denoted by $T_\Sigma(X)$.

Convention

- Elements of $T_\Sigma(X)$ are called Σ-trees indexed by X
- $T_\Sigma = T_\Sigma(\emptyset)$
- $t|_x$ number of occurrences of $x \in X$ in $t \in T_\Sigma(X)$

Tree Properties

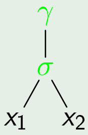
Definition

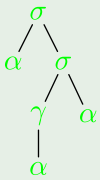
Let $t \in T_\Sigma(X)$

- t linear in X, if every $x \in X$ occurs at most once in t
- t nondeleting in X, if every $x \in X$ occurs at least once in t

Example

$\Sigma = \{\alpha^{(0)}, \gamma^{(1)}, \sigma^{(2)}\}$ and $X = \{x_1, x_2\}$





- Both trees linear in X
- Left tree nondeleting in X

Semiring

Definition

A semiring $\mathcal{A}$ is an algebraic structure $(A, +, \cdot, 0, 1)$ such that

- $(A, +, 0)$ is a commutative monoid
- $(A, \cdot, 1)$ is a monoid
- $\cdot$ distributes (both-sided) over $+$
- 0 is absorbing with respect to $\cdot$

Example

- Booleans $(\{0, 1\}, \vee, \wedge, 0, 1)$
- Natural numbers $(\mathbb{N}, +, \cdot, 0, 1)$
- Probabilities $([0, 1], \max, \cdot, 0, 1)$
- Positive reals $([0, \infty), +, \cdot, 0, 1)$
- Subsets $(\mathfrak{P}(A), \cup, \cap, \emptyset, A)$

Semiring Properties

Definition

Semiring $(A, +, \cdot, 0, 1)$ is

- commutative, if $a \cdot b = b \cdot a$
- idempotent, if $a + a = a$
- semifield, if $(A \setminus \{0\}, \cdot, 1)$ is a group
- locally finite, if $\langle B \rangle$ is finite for every finite $B \subseteq A$

Example

semiring	commutative	idempotent	semifield	locally finite
$\{0, 1\}$	yes	yes	yes	yes
$\mathbb{N}$	yes	NO	NO	NO
$[0, 1]$	yes	yes	NO	NO
$\mathbb{R}_+$	yes	NO	yes	NO
$\mathfrak{P}(A)$	yes	yes	yes	yes

Tree Languages

Definition

- Any $L \subseteq T_\Sigma(X)$ is a tree language
- Can be seen as mapping $\psi \colon T_\Sigma(X) \to \{0, 1\}$

$$t \in L \quad \Longleftrightarrow \quad \psi(t) = 1$$

- Set of tree languages $B\langle\!\langle T_\Sigma(X) \rangle\!\rangle$ where $B = \{0, 1\}$

Tree Series

Definition

- Tree series is mapping $\psi \colon T_\Sigma(X) \to A$
- Set of tree series $A\langle\langle T_\Sigma(X)\rangle\rangle$

Example

- Height $\mathrm{height} \colon T_\Sigma(X) \to \mathbb{N}$ is a tree series
- Size $\mathrm{size} \colon T_\Sigma(X) \to \mathbb{N}$ is a tree series

Conventions

- A usually endowed with semiring structure
- $\widetilde{0}$ is tree series that maps every tree to 0
- $\psi(t)$ written as (ψ, t)

Tree Series

Definition

Let $\psi \in A\langle\langle T_\Sigma(X)\rangle\rangle$

$$\mathrm{supp}(\psi) = \{t \in T_\Sigma(X) \mid (\psi, t) \neq 0\}$$

i.e. set of nonzero-weighted trees

Example

Let $L \subseteq T_\Sigma(X)$ and ψ characteristic mapping for L

$$\mathrm{supp}(\psi) = L$$

Definition

- ψ linear in X, if t linear in X for every $t \in \mathrm{supp}(\psi)$
- ψ nondeleting in X, if t nondeleting in X for every $t \in \mathrm{supp}(\psi)$

Tree Series

Convention

- tree series ψ written as $\sum_{t \in T_\Sigma(X)} (\psi, t)\ t$
- $(\psi + \varphi, t) = (\psi, t) + (\varphi, t)$
- $(a \cdot \psi, t) = a \cdot (\psi, t)$

Example

- $\psi = 5\,\alpha \ + \ 23\,\gamma(\alpha) \ + \ 1\,\sigma(\alpha, \alpha)$
- $\varphi = 2\,\alpha \ + \ 10\,\sigma(\gamma(\alpha), \alpha)$
- scalar multiplication $2 \cdot \psi$ gives

$$2 \cdot \psi = 10\,\alpha \ + \ 46\,\gamma(\alpha) \ + \ 2\,\sigma(\alpha, \alpha)$$

- sum $\psi + \varphi$ gives

$$\psi + \varphi = 7\,\alpha \ + \ 23\,\gamma(\alpha) \ + \ 1\,\sigma(\alpha, \alpha) \ + \ 10\,\sigma(\gamma(\alpha), \alpha)$$

Weighted Rewrite Rules

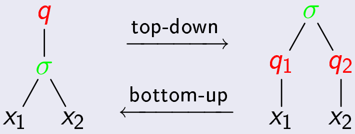
- top-down

$$q(\sigma(x_1, \ldots, x_k)) \xrightarrow{a} \sigma(q_1(x_1), \ldots, q_k(x_k))$$

- bottom-up

$$\sigma(q_1(x_1), \ldots, q_k(x_k)) \xrightarrow{a} q(\sigma(x_1, \ldots, x_k))$$

Illustration



Tree Representation

Definition

A tree representation $\mu = (\mu_k)_{k \in \mathbb{N}}$ over Σ, Q, and $\mathcal{A}$ consists of mappings

$$\mu_k : \Sigma_k \longrightarrow A^{Q \times Q^k}$$

Example

Tree representation μ over $\{\alpha^{(0)}, \gamma^{(1)}, \sigma^{(2)}\}$, $\{1, 2\}$, and $(\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$

$$\mu_0(\alpha)_1 = 1 \qquad \mu_0(\alpha)_2 = 0$$

$$\mu_1(\gamma)_{1,1} = 1 \qquad \mu_1(\gamma)_{2,2} = 0$$

$$\mu_2(\sigma)_{1,12} = 1 \qquad \mu_2(\sigma)_{2,22} = 0$$

$$\mu_2(\sigma)_{1,21} = 1$$

Tree Representation

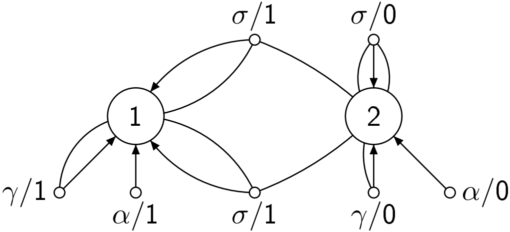
- Just a transition matrix for every input symbol

$$\mu_0(\alpha) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\mu_1(\gamma) = \begin{pmatrix} 1 & -\infty \\ -\infty & 0 \end{pmatrix}$$

$$\mu_2(\sigma) = \begin{pmatrix} -\infty & 1 & 1 & -\infty \\ -\infty & -\infty & -\infty & 0 \end{pmatrix}$$

- Graphically represented:



Weighted Tree Automata

Definition

$(Q, \Sigma, \mathcal{A}, F, \mu)$ weighted tree automaton

- Q finite, nonempty set of states
- Σ ranked alphabet
- $\mathcal{A} = (A, +, \cdot, 0, 1)$ semiring
- $F\colon Q \to A$ final distribution
- μ tree representation over Σ, Q, and $\mathcal{A}$

Note

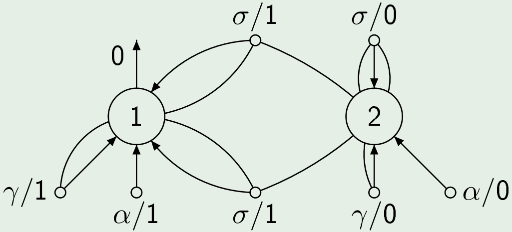
- top-down and bottom-up equivalent
- weight for leaving the system

Weighted Tree Automata

Example

$M_{\mathrm{ht}} = (Q, \Sigma, \mathcal{A}, F, \mu)$ with

- $Q = \{1, 2\}$
- $\Sigma = \{\alpha^{(0)}, \gamma^{(1)}, \sigma^{(2)}\}$
- $\mathcal{A} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$
- $F_1 = 0$ and $F_2 = -\infty$
- μ as shown before



Semantics of WTA

Definition (Initial algebra semantics)

$M = (Q, \Sigma, \mathcal{A}, F, \mu)$ wta

- $h_\mu \colon T_\Sigma \to A^Q$ given by

$$h_\mu(\sigma(t_1, \ldots, t_k))_q = \sum_{q_1,\ldots,q_k \in Q} \mu_k(\sigma)_{q,q_1\cdots q_k} \cdot h_\mu(t_1)_{q_1} \cdot \ldots \cdot h_\mu(t_k)_{q_k}$$

- Semantics of M, denoted by $\|M\| \in A\langle\langle T_\Sigma \rangle\rangle$

$$(\|M\|, t) = \sum_{q \in Q} F_q \cdot h_\mu(t)_q$$

Example

$\|M_{\mathrm{ht}}\| = \text{height}$

Alternative Semantics

Definition (Run semantics)

$M = (Q, \Sigma, \mathcal{A}, F, \mu)$ wta

- set of runs: $R_M = T_{\langle \Sigma, Q \rangle}$
- runs on t: $R_M(t) = \{r \in R_M \mid \pi_1(r) = t\}$
- runs on t ending in q: $R_M(t, q) = \{r \in R_M(t) \mid \pi_2(r(\varepsilon)) = q\}$
- weight of a run $r = \langle \sigma, q \rangle (r_1, \ldots, r_k)$

$$c_M(r) = \mu_k(\sigma)_{q, q_1 \cdots q_k} \cdot c_M(r_1) \cdot \ldots \cdot c_M(r_k)$$

  where $q_i = \pi_2(r(i))$

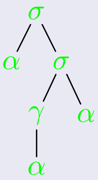
Semantics of M, denoted by $|M| \in A\langle\!\langle T_\Sigma \rangle\!\rangle$

$$(|M|, t) = \sum_{q \in Q} F_q \cdot \Big(\sum_{r \in R_M(t,q)} c_M(r) \Big)$$

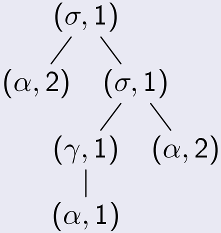
Illustration of Run Semantics

Illustration

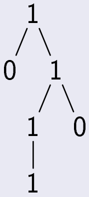
Using the semiring $(\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$



input tree



run



weight

This run has weight $1 + 0 + 1 + 1 + 0 + 1 = 4$

Equivalence of Initial Algebra and Run Semantics

Theorem

$M = (Q, \Sigma, \mathcal{A}, F, \mu)$ *wta*

$$|M| = \|M\|$$

Definition

- $A^{\mathrm{rec}}\langle\langle T_\Sigma \rangle\rangle$ class of recognizable tree series
- $A^{\mathrm{rec}}_{\mathrm{d}}\langle\langle T_\Sigma \rangle\rangle$ class of deterministically (bottom-up) recognizable tree series

Major Theorems—Determinization

Theorem (Borchardt, Vogler 03)

$\mathcal{A}$ locally finite semifield

$$A^{\mathrm{rec}}\langle\langle T_\Sigma \rangle\rangle = A_{\mathrm{d}}^{\mathrm{rec}}\langle\langle T_\Sigma \rangle\rangle$$

(proved for automata with final states)

Theorem (Borchardt 04)

$\mathcal{A}$ locally finite semiring

$$A^{\mathrm{rec}}\langle\langle T_\Sigma \rangle\rangle = A_{\mathrm{d}}^{\mathrm{rec}}\langle\langle T_\Sigma \rangle\rangle$$

(proved for automata with final weights)

Kleene Characterization

Definition (Rational Tree Series)

- Polynomials (finite support tree series) are rational
- Closed under:
  - sum
  - scalar product
  - top-concatenation
  - α-concatenation ($\alpha \in \Sigma_0$)
  - α-Kleene-star ($\alpha \in \Sigma_0$)

Theorem (Droste, Pech, Vogler 05)

$\mathcal{A}$ commutative semiring

$$A^{\mathrm{rec}}\langle\langle T_\Sigma \rangle\rangle = A^{\mathrm{rat}}\langle\langle T_\Sigma \rangle\rangle$$

Major Theorems

Further Characterizations

- Myhill-Nerode characterization [Borchardt 04]
- Systems of Equations [Kuich 97, Bozapalidis 99]
- Weighted automata using fixpoints [Kuich 97]
- Syntactic algebras [Bozapalidis 91]

Tree Series Substitutions

Definition (ε- or pure substitution)

$\psi \in A\langle\langle T_\Sigma(X_k)\rangle\rangle$ and $\psi_1, \ldots, \psi_k \in A\langle\langle T_\Sigma\rangle\rangle$

$$(\psi \overset{\varepsilon}{\leftarrow} (\psi_1, \ldots, \psi_k), u) = \sum_{\substack{t \in T_\Sigma(X_k) \\ t_1, \ldots, t_k \in T_\Sigma \\ u = t[t_1, \ldots, t_k]}} (\psi, t) \cdot (\psi_1, t_1) \cdot \ldots \cdot (\psi_k, t_k)$$

Definition (o-substitution)

$\psi \in A\langle\langle T_\Sigma(X_k)\rangle\rangle$ and $\psi_1, \ldots, \psi_k \in A\langle\langle T_\Sigma\rangle\rangle$

$$\begin{aligned}&(\psi \overset{\mathrm{o}}{\leftarrow} (\psi_1, \ldots, \psi_k), u) \\ &= \sum_{\substack{t \in T_\Sigma(X_k) \\ t_1, \ldots, t_k \in T_\Sigma \\ u = t[t_1, \ldots, t_k]}} (\psi, t) \cdot (\psi_1, t_1)^{|t|_{x_1}} \cdot \ldots \cdot (\psi_k, t_k)^{|t|_{x_k}}\end{aligned}$$

Tree Series Transducer

Definition

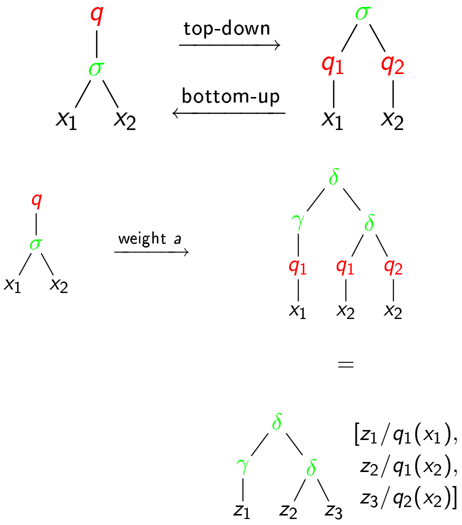
$M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tree series transducer, if

- Q finite set of states
- Σ and Δ input and output ranked alphabet
- $\mathcal{A}$ semiring
- $F \subseteq Q$ set of designated states
- $\mu = (\mu_k)_{k \in \mathbb{N}}$ with

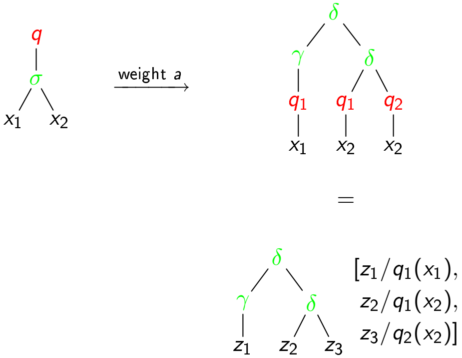
$$\mu_k : \Sigma_k \to A\langle\langle T_\Delta(X) \rangle\rangle^{Q \times Q(X_k)^*}$$

Looks bad, but wait!

Top-down Tree Series Transducer



Top-down Tree Series Transducer



Corresponding Tree Representation Entry

$$(\mu_k(\sigma)_{q,q_1(x_1)q_1(x_2)q_2(x_2)}, \delta(\gamma(z_1), \delta(z_2, z_3))) = a$$

Semantics of TST

Definition (Initial Algebra Semantics)

$M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst, $\eta \in \{\varepsilon, \mathrm{o}\}$

- $h^\eta_\mu \colon T_\Sigma \to A\langle\!\langle T_\Delta \rangle\!\rangle^Q$

$$h^\eta_\mu(\sigma(t_1, \dots, t_k))_q = \sum_{w = q_1(x_{i_1}) \cdots q_n(x_{i_n}) \in Q(X_k)^*} \mu_k(\sigma)_{q,w} \overset{\eta}{\leftarrow} (h^\eta_\mu(t_{i_1})_{q_1}, \dots, h^\eta_\mu(t_{i_n})_{q_n})$$

- Semantics of M, denoted $\|M\|^\eta \colon T_\Sigma \to A\langle\!\langle T_\Delta \rangle\!\rangle$

$$(\|M\|^\eta, t) = \sum_{q \in F} h^\eta_\mu(t)_q$$

Run Semantics and Properties of TST

Note

For pure (ε-) substitution there exists an equivalent run semantics [Fülöp, Vogler 04]

Definition

$(Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ top-down TST

- deterministic, if there is at most one rule with a given left hand side and at most one initial state
- linear, if (for every rule) every variable appears at most once in the right hand side
- nondeleting, if (for every rule) variables that occur in the left hand side also occur in the right hand side

Classes of Transformations

Definition

denotation	class of transformations computed by	substitution
$x\text{-TOP}_{\varepsilon}(\mathcal{A})$	top-down TST with properties x	ε-subst.
$x\text{-TOP}_{\mathrm{o}}(\mathcal{A})$	top-down TST with properties x	o-subst.
$x\text{-BOT}_{\varepsilon}(\mathcal{A})$	bottom-up TST with properties x	ε-subst.
$x\text{-BOT}_{\mathrm{o}}(\mathcal{A})$	bottom-up TST with properties x	o-subst.

Note

Bottom-up TST process input tree from the leaves toward the root.

Composition of Transformations

Definition

Let

- $\varphi \colon T_\Sigma \times T_\Delta \to A$
- $\psi \colon T_\Delta \times T_\Gamma \to A$

Composition of φ and ψ

$$(\varphi \,;\psi) \colon T_\Sigma \times T_\Gamma \to A$$

$$(t, v) \mapsto \sum_{u \in T_\Delta} \varphi(t, u) \cdot \psi(u, v)$$

Composition Results

Theorem (see [Kuich 99] and [Engelfriet et al 02])

$\mathcal{A}$ *commutative semiring*

$$\text{nlp-BOT}(\mathcal{A}) \mathbin{;} \text{p-BOT}(\mathcal{A}) = \text{p-BOT}(\mathcal{A})$$

$$\text{p-BOT}(\mathcal{A}) \mathbin{;} \text{bdth-BOT}(\mathcal{A}) = \text{p-BOT}(\mathcal{A})$$

Theorem

$\mathcal{A}$ *commutative semiring*

$$\text{lp-BOT}(\mathcal{A}) \mathbin{;} \text{p-BOT}(\mathcal{A}) = \text{p-BOT}(\mathcal{A})$$

$$\text{p-BOT}(\mathcal{A}) \mathbin{;} \text{bd-BOT}(\mathcal{A}) = \text{p-BOT}(\mathcal{A})$$

$$\text{bdt-TOP}(\mathcal{A}) \mathbin{;} \text{lp-TOP}(\mathcal{A}) \subseteq \text{p-TOP}(\mathcal{A})$$

References

Björn Borchardt.
The Theory of Recognizable Tree Series.
PhD thesis, Technische Universität Dresden, 2005.

Joost Engelfriet, Zoltán Fülöp, and Heiko Vogler.
Bottom-up and top-down tree series transformations.
J. Autom. Lang. Combin., 7(1):11–70, 2002.

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

Andreas Maletti.
The Power of Tree Series Transducers.
PhD thesis, Technische Universität Dresden, 2006.