Preservation of Recognizability for o-substitution

Andreas Maletti

Department of Computer Science
TECHNISCHE UNIVERSITÄT DRESDEN

March 27, 2006

Applications

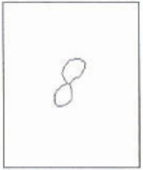
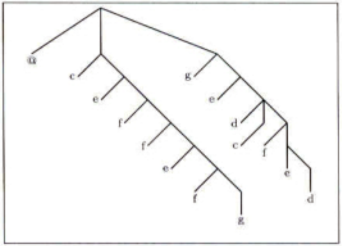
. . . of (weighted/probabilistic) tree automata:

- Syntactic Pattern Matching (e.g. handwritten digit recognition)
  [López, Piñaga: Syntactic Pattern Recognition by Error Correcting Analysis on Tree Automata, 2000]
- Tree Banks
  [Liakata, Pulman: Learning Theories from Text, 2004]

. . . of tree series transducers:

- Code Selection
  [Borchardt: Code Selection by Tree Series Transducers, 2004]
- Natural Language Processing
  [Graehl, Knight: Training Tree Transducers, 2004]

Syntactic Pattern Recognition

- Source image:  Thinned image: 
- Representation as tree: 
- Which transformations preserve finite-state recognizability?

Tree Series Substitution Respecting Occurrences

Used Notation:

- $T_\Sigma(V)$: set of V-indexed trees (terms) formed using the ranked alphabet Σ
- $T_\Sigma = T_\Sigma(\emptyset)$
- $A\langle\langle T\rangle\rangle$: set of mappings $\psi\colon T \longrightarrow A$
- (ψ, t) denotes $\psi(t)$
- $\text{supp}(\psi) = \{\, t \in T \mid (\psi, t) \neq 0 \,\}$
- $\mathrm{Z}_n = \{\mathrm{z}_1, \ldots, \mathrm{z}_n\}$

Definition:

Let $\psi, \psi_1, \ldots, \psi_n \in A\langle\langle T_\Sigma(\mathrm{Z}_n)\rangle\rangle$.

$$\psi \overset{\circ}{\leftarrow} (\psi_1, \ldots, \psi_n) = \sum_{\substack{t \in \text{supp}(\psi),\\ t_1 \in \text{supp}(\psi_1),\\ \ldots,\\ t_n \in \text{supp}(\psi_n)}} (\psi, t) \cdot (\psi_1, t_1)^{|t|_{\mathrm{z}_1}} \cdot \ldots \cdot (\psi_n, t_n)^{|t|_{\mathrm{z}_n}} \; t[t_1, \ldots, t_n]$$

Tree Series Substitution Respecting Occurrences

Used Notation:

- $T_\Sigma(V)$: set of V-indexed trees (terms) formed using the ranked alphabet Σ
- $T_\Sigma = T_\Sigma(\emptyset)$
- $A\langle\langle T\rangle\rangle$: set of mappings $\psi\colon T \longrightarrow A$
- (ψ, t) denotes $\psi(t)$
- $\mathrm{supp}(\psi) = \{\, t \in T \mid (\psi, t) \neq 0 \,\}$
- $\mathrm{Z}_n = \{\mathrm{z}_1, \dots, \mathrm{z}_n\}$

Definition:

Let $\psi, \psi_1, \dots, \psi_n \in A\langle\langle T_\Sigma(\mathrm{Z}_n)\rangle\rangle$.

$$\psi \overset{\circ}{\leftarrow} (\psi_1, \dots, \psi_n) = \sum_{\substack{t \in \mathrm{supp}(\psi), \\ t_1 \in \mathrm{supp}(\psi_1), \\ \dots, \\ t_n \in \mathrm{supp}(\psi_n)}} (\psi, t) \cdot (\psi_1, t_1)^{|t|_{\mathrm{z}_1}} \cdot \ldots \cdot (\psi_n, t_n)^{|t|_{\mathrm{z}_n}} \; t[t_1, \dots, t_n]$$

Notes on Substitution

- introduced in
  [Fülöp, Vogler: Tree Series Transformations that Respect Copying, 2003]
- potentially infinite sum
- usually only considered for polynomial (i.e. finite support) tree series or in complete semirings (that have an infinite summation)

Example:

Let $\Delta = \{\delta^{(2)}, \alpha^{(0)}\}$ and $\psi \in \mathbb{N}\langle\langle T_\Delta(\mathrm{Z}_1)\rangle\rangle$ be

$$\psi = \max_{t \in T_\Delta(\mathrm{Z}_1)} |t|_\delta \; t$$

- $\psi \overset{\circ}{\longleftarrow} (\psi)$ undefined in $\mathbb{A} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$ [not a complete semiring]
- $\psi \overset{\circ}{\longleftarrow} (\psi) = \psi$ in $\mathbb{A}_\infty = (\mathbb{N} \cup \{\infty, -\infty\}, \max, +, -\infty, 0)$ [a complete semiring]

Notes on Substitution

- introduced in
  [Fülöp, Vogler: Tree Series Transformations that Respect Copying, 2003]
- potentially infinite sum
- usually only considered for polynomial (i.e. finite support) tree series or in complete semirings (that have an infinite summation)

Example:

Let $\Delta = \{\delta^{(2)}, \alpha^{(0)}\}$ and $\psi \in \mathbb{N}\langle\langle T_\Delta(\mathrm{Z}_1)\rangle\rangle$ be

$$\psi = \max_{t \in T_\Delta(\mathrm{Z}_1)} |t|_\delta \; t$$

- $\psi \overset{\circ}{\leftarrow} (\psi)$ undefined in $\mathbb{A} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$ [not a complete semiring]
- $\psi \overset{\circ}{\leftarrow} (\psi) = \psi$ in $\mathbb{A}_\infty = (\mathbb{N} \cup \{\infty, -\infty\}, \max, +, -\infty, 0)$ [a complete semiring]

Weighted Tree Automata

Definition:

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

- Q finite set (of *states*)
- Σ ranked alphabet
- $\mathcal{A} = (A, +, \cdot, 0, 1)$ semiring
- $F : Q \longrightarrow A$ (*final distribution*)
- $\mu = (\mu_k)_{k \in \mathbb{N}}$ with $\mu_k : \Sigma_k \longrightarrow A^{Q \times Q^k}$

Example:

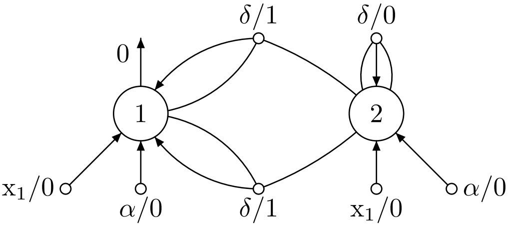
- $Q = \{1, 2\}$
- $\Sigma = \{\delta^{(2)}, \alpha^{(0)}, x_1^{(0)}\}$
- $\mathcal{A} = \mathbb{A}_\infty$
- $F(1) = 0$, $F(2) = -\infty$
- μ see graphic below

Weighted Tree Automata

Definition:

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

- Q finite set (of *states*)
- Σ ranked alphabet
- $\mathcal{A} = (A, +, \cdot, 0, 1)$ semiring
- $F \colon Q \longrightarrow A$ (*final distribution*)
- $\mu = (\mu_k)_{k \in \mathbb{N}}$ with $\mu_k \colon \Sigma_k \longrightarrow A^{Q \times Q^k}$



Example:

- $Q = \{1, 2\}$
- $\Sigma = \{\delta^{(2)}, \alpha^{(0)}, x_1^{(0)}\}$
- $\mathcal{A} = \mathbb{A}_\infty$
- $F(1) = 0$, $F(2) = -\infty$
- μ see graphic below

Recognizable Tree Series

Definition:

Let $M = (Q, \Sigma, \mathcal{A}, F, \mu)$ weighted tree automaton. Define $h_\mu \colon T_\Sigma \longrightarrow A^Q$

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

Tree series computed by M is $\|M\|$

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

Definition:

Tree series $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable, if there exists wta M with $\|M\| = \psi$.

Recognizable Tree Series

Definition:

Let $M = (Q, \Sigma, \mathcal{A}, F, \mu)$ weighted tree automaton. Define $h_\mu : T_\Sigma \longrightarrow A^Q$

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

Tree series computed by M is $\|M\|$

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

Definition:

Tree series $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable, if there exists wta M with $\|M\| = \psi$.

Preservation of Recognizability

Theorem:

- $\mathcal{A}$ commutative, idempotent, and continuous
- $\psi \in A\langle\langle T_\Sigma(\mathrm{Z}_n)\rangle\rangle$ recognizable and linear
- $\psi_1, \ldots, \psi_n \in A\langle\langle T_\Sigma\rangle\rangle$ recognizable

Then $\psi \overset{\circ}{\longleftarrow} (\psi_1, \ldots, \psi_n)$ is recognizable.

Definition:

- $\mathcal{A}$ commutative: $a \cdot b = b \cdot a$ for all $a, b \in A$
- $\mathcal{A}$ idempotent: $a + a = a$ for all $a \in A$
- $\mathcal{A}$ continuous: $\mathcal{A}$ complete and $\sum$ preserves certain suprema
- ψ linear: every $t \in \mathrm{supp}(\psi)$ linear (i.e., $\mathrm{z}_1, \ldots, \mathrm{z}_n$ occur at most once)

Preservation of Recognizability

Theorem:

- $\mathcal{A}$ commutative, idempotent, and continuous
- $\psi \in A\langle\langle T_\Sigma(\mathrm{Z}_n)\rangle\rangle$ recognizable and linear
- $\psi_1, \ldots, \psi_n \in A\langle\langle T_\Sigma\rangle\rangle$ recognizable

Then $\psi \overset{\circ}{\leftarrow} (\psi_1, \ldots, \psi_n)$ is recognizable.

Definition:

- $\mathcal{A}$ commutative: $a \cdot b = b \cdot a$ for all $a, b \in A$
- $\mathcal{A}$ idempotent: $a + a = a$ for all $a \in A$
- $\mathcal{A}$ continuous: $\mathcal{A}$ complete and $\sum$ preserves certain suprema
- ψ linear: every $t \in \mathrm{supp}(\psi)$ linear (i.e., $\mathrm{z}_1, \ldots, \mathrm{z}_n$ occur at most once)

Preservation of Recognizability

Theorem:

- $\mathcal{A}$ commutative, idempotent, and continuous
- $\psi \in A\langle\langle T_\Sigma(\mathrm{Z}_n)\rangle\rangle$ recognizable and linear
- $\psi_1, \ldots, \psi_n \in A\langle\langle T_\Sigma\rangle\rangle$ recognizable

Then $\psi \overset{\circ}{\longleftarrow} (\psi_1, \ldots, \psi_n)$ is recognizable.

Definition:

- $\mathcal{A}$ commutative: $a \cdot b = b \cdot a$ for all $a, b \in A$
- $\mathcal{A}$ idempotent: $a + a = a$ for all $a \in A$
- $\mathcal{A}$ continuous: $\mathcal{A}$ complete and $\sum$ preserves certain suprema
- ψ linear: every $t \in \mathrm{supp}(\psi)$ linear (i.e., $\mathrm{z}_1, \ldots, \mathrm{z}_n$ occur at most once)

Preservation of Recognizability

Theorem:

- $\mathcal{A}$ commutative, idempotent, and continuous
- $\psi \in A\langle\langle T_\Sigma(Z_n)\rangle\rangle$ recognizable and linear
- $\psi_1, \ldots, \psi_n \in A\langle\langle T_\Sigma\rangle\rangle$ recognizable

Then $\psi \overset{\circ}{\leftarrow} (\psi_1, \ldots, \psi_n)$ is recognizable.

Definition:

- $\mathcal{A}$ commutative: $a \cdot b = b \cdot a$ for all $a, b \in A$
- $\mathcal{A}$ idempotent: $a + a = a$ for all $a \in A$
- $\mathcal{A}$ continuous: $\mathcal{A}$ complete and $\sum$ preserves certain suprema
- ψ linear: every $t \in \mathrm{supp}(\psi)$ linear (i.e., $z_1, \ldots, z_n$ occur at most once)

Preservation of Recognizability

Theorem:

- $\mathcal{A}$ commutative, idempotent, and continuous
- $\psi \in A\langle\langle T_\Sigma(\mathsf{Z}_n)\rangle\rangle$ recognizable and linear
- $\psi_1, \ldots, \psi_n \in A\langle\langle T_\Sigma\rangle\rangle$ recognizable

Then $\psi \stackrel{\circ}{\leftarrow} (\psi_1, \ldots, \psi_n)$ is recognizable.

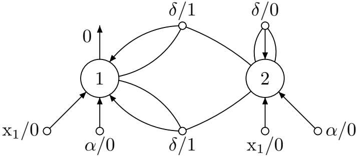
Definition:

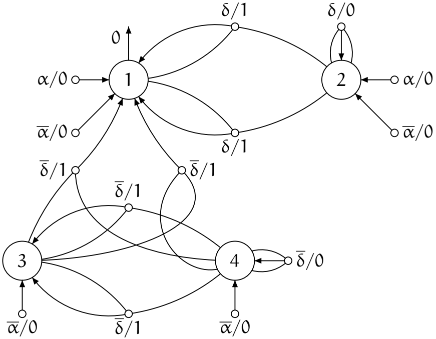
- $\mathcal{A}$ commutative: $a \cdot b = b \cdot a$ for all $a, b \in A$
- $\mathcal{A}$ idempotent: $a + a = a$ for all $a \in A$
- $\mathcal{A}$ continuous: $\mathcal{A}$ complete and $\sum$ preserves certain suprema
- ψ linear: every $t \in \mathrm{supp}(\psi)$ linear (i.e., $\mathsf{z}_1, \ldots, \mathsf{z}_n$ occur at most once)

Illustration

Proof Idea:

Illustrated on $\psi \overset{\circ}{\leftarrow} (\psi)$



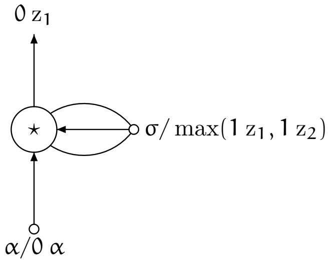


Syntax

Definition:

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

- Q finite set (of *states*)
- Σ and Δ ranked alphabets
- $\mathcal{A}$ semiring
- $F \colon Q \longrightarrow A\langle\langle C_\Delta(\mathrm{Z}_1)\rangle\rangle$
- $\mu = (\mu_k)_{k \in \mathbb{N}}$ with
  $\mu_k \colon \Sigma_k \longrightarrow A\langle\langle T_\Delta(\mathrm{Z}_n)\rangle\rangle^{Q \times Q(\mathrm{Z}_k)^*}$



Definition:

Tst $(Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ recognizable, if

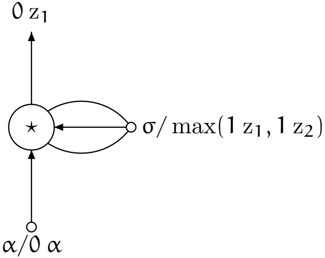
- $F(q)$ recognizable for every $q \in Q$
- $\mu_k(\sigma)_{q,w}$ recognizable for all k, σ, q, and w

Syntax

Definition:

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

- Q finite set (of *states*)
- Σ and Δ ranked alphabets
- $\mathcal{A}$ semiring
- $F \colon Q \longrightarrow A\langle\langle C_\Delta(Z_1)\rangle\rangle$
- $\mu = (\mu_k)_{k \in \mathbb{N}}$ with
  $\mu_k \colon \Sigma_k \longrightarrow A\langle\langle T_\Delta(Z_n)\rangle\rangle^{Q \times Q(Z_k)^*}$



Definition:

Tst $(Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ recognizable, if

- $F(q)$ recognizable for every $q \in Q$
- $\mu_k(\sigma)_{q,w}$ recognizable for all k, σ, q, and w

Semantics

Definition:

Let $(Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst. Define $h_\mu^\circ : T_\Sigma \longrightarrow A\langle\langle T_\Delta \rangle\rangle^Q$

$$h_\mu^\circ(\sigma(t_1, \ldots, t_k))_q = \sum_{\substack{w \in Q(Z_k)^* \\ w = q_1(z_{i_1}) \cdots q_n(z_{i_n})}} \mu_k(\sigma)_{q,w} \overset{\circ}{\leftarrow} (h_\mu^\circ(t_{i_1})_{q_1}, \ldots, h_\mu^\circ(t_{i_n})_{q_n})$$

Example:

Input tree: $\sigma(\sigma(\alpha, \alpha), \alpha)$

$$h_\mu^\circ(\alpha)_\star = 0\ \alpha$$

$$h_\mu^\circ(\sigma(\alpha, \alpha))_\star = \max(1\ z_1, 1\ z_2) \overset{\circ}{\leftarrow} (0\ \alpha, 0\ \alpha) = 1\ \alpha$$

$$h_\mu^\circ(\sigma(\sigma(\alpha, \alpha), \alpha))_\star = \max(1\ z_1, 1\ z_2) \overset{\circ}{\leftarrow} (1\ \alpha, 0\ \alpha) = 2\ \alpha$$

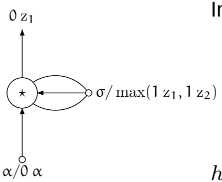
Semantics

Definition:

Let $(Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst. Define $h^\circ_\mu : T_\Sigma \longrightarrow A\langle\langle T_\Delta \rangle\rangle^Q$

$$h^\circ_\mu(\sigma(t_1, \dots, t_k))_q = \sum_{\substack{w \in Q(Z_k)^* \\ w = q_1(z_{i_1}) \cdots q_n(z_{i_n})}} \mu_k(\sigma)_{q,w} \overset{\circ}{\leftarrow} (h^\circ_\mu(t_{i_1})_{q_1}, \dots, h^\circ_\mu(t_{i_n})_{q_n})$$

Example:



Input tree: $\sigma(\sigma(\alpha, \alpha), \alpha)$

$$h^\circ_\mu(\alpha)_\star = 0\ \alpha$$

$$h^\circ_\mu(\sigma(\alpha, \alpha))_\star = \max(1\ z_1, 1\ z_2) \overset{\circ}{\leftarrow} (0\ \alpha, 0\ \alpha) = 1\ \alpha$$

$$h^\circ_\mu(\sigma(\sigma(\alpha, \alpha), \alpha))_\star = \max(1\ z_1, 1\ z_2) \overset{\circ}{\leftarrow} (1\ \alpha, 0\ \alpha) = 2\ \alpha$$

Semantics

Definition:

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst. Transformation computed by M

Tree Level $\|M\| \colon T_\Sigma \longrightarrow A\langle\langle T_\Delta \rangle\rangle$:

$$\|M\|(t) = \sum_{q \in Q} F_q \overset{\circ}{\leftarrow} (h^\circ_\mu(t)_q)$$

Series Level $\|M\| \colon A\langle\langle T_\Sigma \rangle\rangle \longrightarrow A\langle\langle T_\Delta \rangle\rangle$:

$$\|M\|(\psi) = \sum_{t \in \mathsf{supp}(\psi)} (\psi, t) \cdot \|M\|(t)$$

Example:

Let $M =$ Then $\|M\|(t) = \mathsf{height}(t)\ \alpha$

Semantics

Definition:

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst. Transformation computed by M

Tree Level $\|M\| \colon T_\Sigma \longrightarrow A\langle\langle T_\Delta \rangle\rangle$:

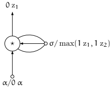
$$\|M\|(t) = \sum_{q \in Q} F_q \overset{\circ}{\leftarrow} (h_\mu^\circ(t)_q)$$

Series Level $\|M\| \colon A\langle\langle T_\Sigma \rangle\rangle \longrightarrow A\langle\langle T_\Delta \rangle\rangle$:

$$\|M\|(\psi) = \sum_{t \in \mathrm{supp}(\psi)} (\psi, t) \cdot \|M\|(t)$$

Example:

Let $M =$



Then $\|M\|(t) = \mathrm{height}(t)\ \alpha$

Tree Series Transducers and Recognizability

Theorem:

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative, idempotent, and continuous
- M recognizable and output-linear

Then $\|M\|(t)$ is recognizable for every $t \in T_\Sigma$.

Definition:

M output-linear: $\mu_k(\sigma)_{q,w}$ linear for all k, σ, q, and w.

Question:

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative, idempotent, and continuous
- M recognizable and output-linear
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Is $\|M\|(\psi)$ recognizable?

Tree Series Transducers and Recognizability

Theorem:

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative, idempotent, and continuous
- M recognizable and output-linear

Then $\|M\|(t)$ is recognizable for every $t \in T_\Sigma$.

Definition:

M output-linear: $\mu_k(\sigma)_{q,w}$ linear for all k, σ, q, and w.

Question:

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative, idempotent, and continuous
- M recognizable and output-linear
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Is $\|M\|(\psi)$ recognizable?

Tree Series Transducers and Recognizability

Theorem:

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative, idempotent, and continuous
- M recognizable and output-linear

Then $\|M\|(t)$ is recognizable for every $t \in T_\Sigma$.

Definition:

M output-linear: $\mu_k(\sigma)_{q,w}$ linear for all k, σ, q, and w.

Question:

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative, idempotent, and continuous
- M recognizable and output-linear
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Is $\|M\|(\psi)$ recognizable?

An Answer

Theorem:

[Kuich: Tree Transducers and Formal Tree Series, 1999]

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative and continuous
- M recognizable, input-linear and -nondeleting, top-down
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Then $\|M\|(\psi)$ is recognizable!

Definition:

- M input-linear: w linear for all w such that $\text{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$
- M input-nondeleting: w nondeleting (every variable from Z_k occurs at least once) for all w such that $\text{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$
- M top-down: $\mu_k(\sigma)_{q,w}$ nondeleting and linear

An Answer

Theorem:

[Kuich: Tree Transducers and Formal Tree Series, 1999]

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative and continuous
- M recognizable, input-linear and -nondeleting, top-down
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Then $\|M\|(\psi)$ is recognizable!

Definition:

- M input-linear: w linear for all w such that $\text{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$
- M input-nondeleting: w nondeleting (every variable from Z_k occurs at least once) for all w such that $\text{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$
- M top-down: $\mu_k(\sigma)_{q,w}$ nondeleting and linear

An Answer

Theorem:

[Kuich: Tree Transducers and Formal Tree Series, 1999]

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative and continuous
- M recognizable, input-linear and -nondeleting, top-down
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Then $\|M\|(\psi)$ is recognizable!

Definition:

- M input-linear: w linear for all w such that $\text{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$
- M input-nondeleting: w nondeleting (every variable from Z_k occurs at least once) for all w such that $\text{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$
- M top-down: $\mu_k(\sigma)_{q,w}$ nondeleting and linear

An Answer

Theorem:

[Kuich: Tree Transducers and Formal Tree Series, 1999]

Let $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ tst.

- $\mathcal{A}$ commutative and continuous
- M recognizable, input-linear and -nondeleting, top-down
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Then $\|M\|(\psi)$ is recognizable!

Definition:

- M input-linear: w linear for all w such that $\mathrm{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$
- M input-nondeleting: w nondeleting (every variable from Z_k occurs at least once) for all w such that $\mathrm{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$
- M top-down: $\mu_k(\sigma)_{q,w}$ nondeleting and linear

The Downside

Observation:

There exist single-state input-linear top-down tst M such that $\|M\|(\psi)$ is not recognizable albeit ψ is recognizable.

Problem:

Sequential execution preserves weight a'!

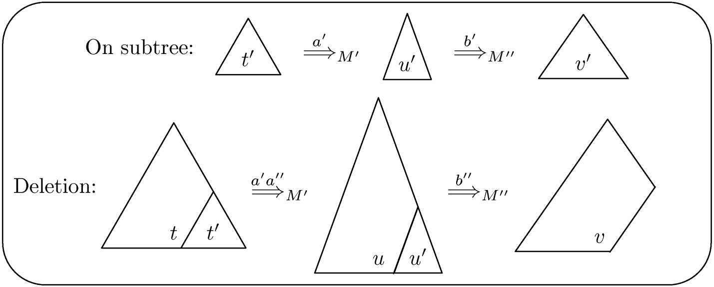
The Downside

Observation:

There exist single-state input-linear top-down tst M such that $\|M\|(\psi)$ is not recognizable albeit ψ is recognizable.

Problem:

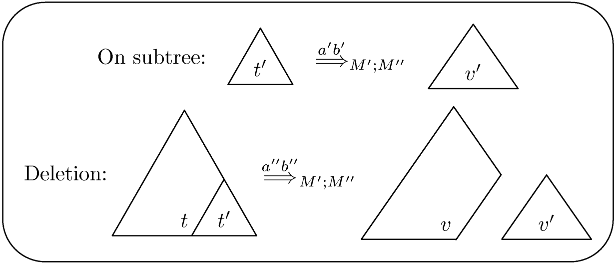
Sequential execution preserves weight a'!



The Downside

Problem:

Deletion neglects the weight a' in the composition!



But: Distinction between 0 and 1 is preserved.

Partial Solution

Idea:

Use boolean weights!

Theorem:

[Borchardt: A Pumping Lemma and Decidability Problems for Recognizable Tree Series, 2004]

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

Then there exists a wta M' with boolean tree representation such that $\|M'\| = \|M\|$.

Definition:

- $\mathcal{A}$ locally finite: closure of finite sets under $+$ and $\cdot$ still finite
- μ boolean: $\mu_k(\sigma)_{q,q_1,\ldots,q_k} \in \{0, 1\}$ for all $k, \sigma, q, q_1, \ldots, q_k$

Partial Solution

Idea:

Use boolean weights!

Theorem:

[Borchardt: A Pumping Lemma and Decidability Problems for Recognizable Tree Series, 2004]

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

Then there exists a wta M' with boolean tree representation such that $\|M'\| = \|M\|$.

Definition:

- $\mathcal{A}$ locally finite: closure of finite sets under $+$ and $\cdot$ still finite
- μ boolean: $\mu_k(\sigma)_{q,q_1,\ldots,q_k} \in \{0, 1\}$ for all $k, \sigma, q, q_1, \ldots, q_k$

Partial Solution

Idea:
Use boolean weights!

Theorem:
[Borchardt: A Pumping Lemma and Decidability Problems for Recognizable Tree Series, 2004]

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

Then there exists a wta M' with boolean tree representation such that $\|M'\| = \|M\|$.

Definition:

- $\mathcal{A}$ locally finite: closure of finite sets under $+$ and $\cdot$ still finite
- μ boolean: $\mu_k(\sigma)_{q,q_1,\dots,q_k} \in \{0, 1\}$ for all k, σ, $q, q_1, \dots, q_k$

Partial Solution

Idea:
Use boolean weights!

Theorem:
[Borchardt: A Pumping Lemma and Decidability Problems for Recognizable Tree Series, 2004]

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

Then there exists a wta M' with boolean tree representation such that $\|M'\| = \|M\|$.

Definition:

- $\mathcal{A}$ locally finite: closure of finite sets under $+$ and $\cdot$ still finite
- μ boolean: $\mu_k(\sigma)_{q,q_1,\dots,q_k} \in \{0, 1\}$ for all k, σ, $q, q_1, \dots, q_k$

The Result

Main Theorem:

- $\mathcal{A}$ commutative, idempotent, continuous, and locally finite
- $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ recognizable and linear bottom-up tst
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Then $\|M\|(\psi)$ is recognizable.

Definition:

M bottom-up: $w = q_1(\mathrm{z}_1) \cdots q_k(\mathrm{z}_k)$ for every w such that $\mathrm{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$

Remaining Question:

What transformations can be realized by such tst?

The Result

Main Theorem:

- $\mathcal{A}$ commutative, idempotent, continuous, and locally finite
- $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ recognizable and linear bottom-up tst
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Then $\|M\|(\psi)$ is recognizable.

Definition:

M bottom-up: $w = q_1(\mathrm{z}_1) \cdots q_k(\mathrm{z}_k)$ for every w such that $\mathrm{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$

Remaining Question:

What transformations can be realized by such tst?

The Result

Main Theorem:

- $\mathcal{A}$ commutative, idempotent, continuous, and locally finite
- $M = (Q, \Sigma, \Delta, \mathcal{A}, F, \mu)$ recognizable and linear bottom-up tst
- $\psi \in A\langle\langle T_\Sigma \rangle\rangle$ recognizable

Then $\|M\|(\psi)$ is recognizable.

Definition:

M bottom-up: $w = q_1(\mathrm{z}_1) \cdots q_k(\mathrm{z}_k)$ for every w such that $\mathrm{supp}(\mu_k(\sigma)_{q,w}) \neq \emptyset$

Remaining Question:

What transformations can be realized by such tst?

References

- D. López, I. Piñaga: *Syntactic Pattern Recognition by Error Correcting Analysis on Tree Automata*, LNCS 1876, p. 133–142, 2000
- B. Borchardt: *The Theory of Recognizable Tree Series*, Akademische Abhandlungen zur Informatik, 2005
- W. Kuich: *Tree Transducers and Formal Tree Series*, Acta Cybernet. 14(1), p. 135–149, 1999
- J. Engelfriet, Z. Fülöp, H. Vogler: *Bottom-up and Top-down Tree Series Transformations*, J. Autom. Lang. Combin. 8(2), p. 219–285, 2003
- Z. Fülöp, H. Vogler: *Tree Series Transformations that Respect Copying*, Theory Comput. Syst. 36(3), p. 247–293, 2003

Thank you for your attention!