Composition Hierarchies of Linear Weighted Extended Top-Down Tree Transducers

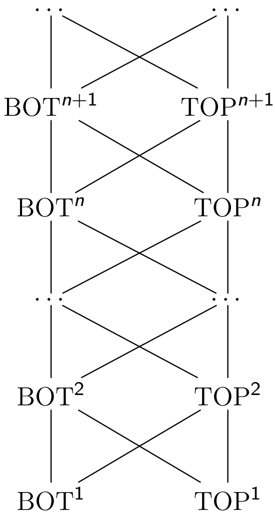
Zoltán Fülöp and Andreas Maletti

September 26, 2018

Composition hierarchy

Theorem (Engelfriet 1981)

Composition of tree transducers yields a proper hierarchy
(of transformations and output languages)



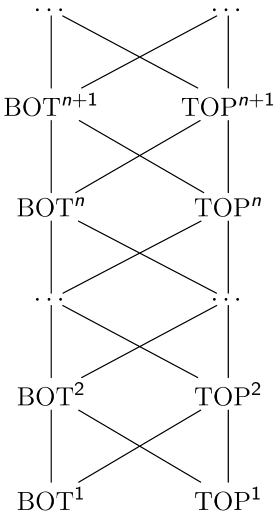
Composition hierarchy

Theorem (Engelfriet 1981)

Composition of tree transducers yields a proper hierarchy
(of transformations and output languages)

Theorem (Fülöp et al. 2004; M. 2006)

Composition of weighted tree transducers yields a proper hierarchy over non-rings



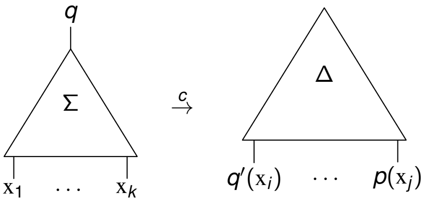
Open question

Is the hierarchy of transformations also proper for rings?

(for fields?)

Extended tree transducer

Weighted extended top-down tree transducer (WXTT)
$M = (Q, \Sigma, \Delta, I, R)$ with finitely many rules

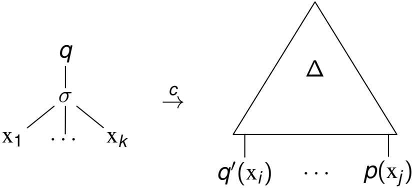


- states $q, q', p \in Q$
- variable indices $i, j \in \{1, \ldots, k\}$

[Arnold, Dauchet: *Bi-transductions de forêts.* Proc. ICALP 1976]
[Graehl, Knight: *Training tree transducers.* Proc. NAACL 2004]

Tree transducer

Weighted top-down tree transducer (WTT) if all rules

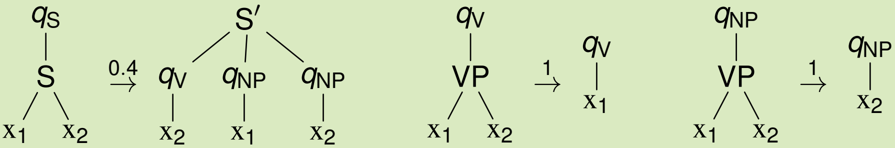


[Rounds: *Mappings and grammars on trees.* Math. Syst. Theory, 1970]
[Thatcher: *Generalized sequential machine maps.* J. Comput. Syst. Sci., 1970]

Derivations

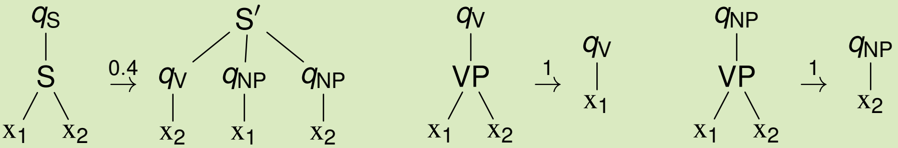
Example

States $\{q_S, q_V, q_{NP}\}$ of which only q_S has non-zero initial weight

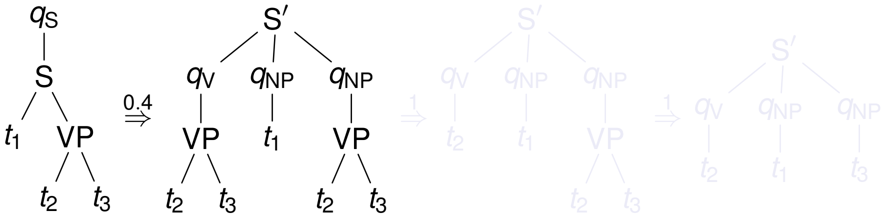


Example

States $\{q_S, q_V, q_{NP}\}$ of which only q_S has non-zero initial weight

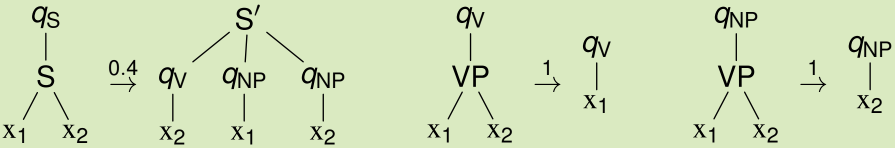


Derivation:

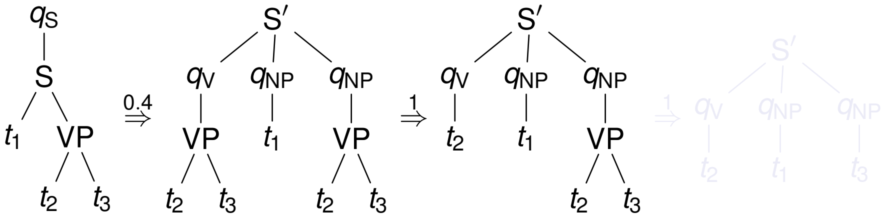


Example

States $\{q_S, q_V, q_{NP}\}$ of which only q_S has non-zero initial weight

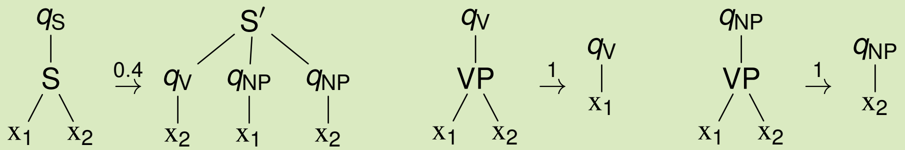


Derivation:

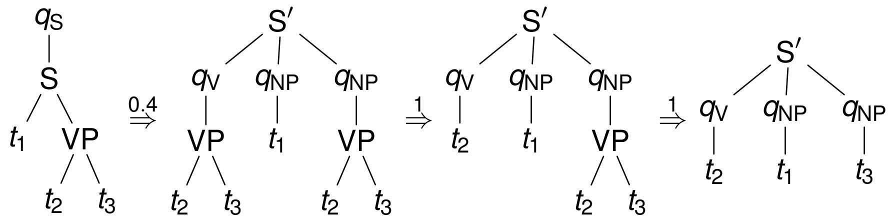


Example

States $\{q_S, q_V, q_{NP}\}$ of which only q_S has non-zero initial weight



Derivation:



Computed transformation

Computed transformation ($t \in T_\Sigma$ and $u \in T_\Delta$):

$$M(t, u) = \sum_{\substack{q \in Q \\ q(t) \overset{c_1}{\Rightarrow} \cdots \overset{c_n}{\Rightarrow} u \\ \text{left-most derivation}}} I(q) \cdot c_1 \cdot \ldots \cdot c_n$$

Computed transformation

Computed transformation ($t \in T_\Sigma$ and $u \in T_\Delta$):

$$M(t, u) = \sum_{\substack{q \in Q \\ q(t) \overset{c_1}{\Rightarrow} \cdots \overset{c_n}{\Rightarrow} u \\ \text{left-most derivation}}} I(q) \cdot c_1 \cdot \ldots \cdot c_n$$

Composition of transformations ($\tau \colon T_\Sigma \times T_\Delta \to \mathbb{K}$ and $\tau' \colon T_\Delta \times T_\Gamma \to \mathbb{K}$):

$$(\tau \,;\, \tau')(t, u) = \sum_{s \in T_\Delta} \tau(t, s) \cdot \tau'(s, u)$$

(Both these sums will be finite in all considered instances.)

Extended vs. non-extended top-down tree transducer

(In the absence of ε-rules)

Expressive power of

Extended top-down tree transducer $=$ top-down tree transducer

Extended vs. non-extended top-down tree transducer

(In the absence of ε-rules)
Expressive power of

Extended top-down tree transducer $=$ top-down tree transducer

but does not generalize to standard subclasses:

- simple if (in each rule)
  - exactly the same variables occur in left and right hand side
  - no variable occurs twice in the right hand side
  - both sides contain an input/output symbol

Extended vs. non-extended top-down tree transducer

In the unweighted setting:

Theorem (Engelfriet 1975; Baker 1979)

Simple top-down tree transformations are closed under composition, so the hierarchy collapses to the first level

Extended vs. non-extended top-down tree transducer

In the unweighted setting:

Theorem (Engelfriet 1975; Baker 1979)

Simple top-down tree transformations are closed under composition, so the hierarchy collapses to the first level

Theorem (Arnold, Dauchet 1982)

Simple extended top-down tree transf. are not closed under composition, but hierarchy collapses to the second level

Let's generalize this result to the weighted setting (for rings)

Normal form for weighted transducers

Theorem (Normal form)

$$\text{s-wXTT}^n = \text{REL} \,;\, \text{Bfus-wXTT}^n$$

Every chain of n weighted simple transducers can equivalently be presented as a chain of

- *a weighted relabeling and*
- *a chain of n Boolean functional unambiguous simple transducers.*

- Boolean = utilizing only the unit weights 0 and 1
- functional = computing a partial function
- unambiguous = having at most one derivation per input-output pair

Normal form for weighted transducers

Proof.

Achieved by induction using the decomposition:

$$\text{s-wXTT} \subseteq \text{REL} \; ; \text{Bfus-wXTT}$$

Normal form for weighted transducers

Proof.

Achieved by induction using the decomposition:

$$\text{s-wXTT} \subseteq \text{REL} \; ; \text{Bfus-wXTT}$$

and the compositions:

$$\text{s-wXTT} \; ; \text{REL} \subseteq \text{s-wXTT} \quad \text{and} \quad \text{REL} \; ; \text{s-wXTT} \subseteq \text{s-wXTT}$$

Normal form for weighted transducers

Proof.

Achieved by induction using the decomposition:

$$\text{s-wXTT} \subseteq \text{REL} \,;\, \text{Bfus-wXTT}$$

and the compositions:

$$\text{s-wXTT} \,;\, \text{REL} \subseteq \text{s-wXTT} \quad \text{and} \quad \text{REL} \,;\, \text{s-wXTT} \subseteq \text{s-wXTT}$$

$$\begin{aligned}
\text{s-wXTT}^{n+1} &= \text{s-wXTT}^{n} \,;\, \text{s-wXTT} \\
&\subseteq \text{s-wXTT}^{n} \,;\, \text{REL} \,;\, \text{Bfus-wXTT} \\
&\subseteq \text{s-wXTT}^{n} \,;\, \text{Bfus-wXTT} \\
&\subseteq \text{REL} \,;\, \text{Bfus-wXTT}^{n} \,;\, \text{Bfus-wXTT} \\
&= \text{REL} \,;\, \text{Bfus-wXTT}^{n+1}
\end{aligned}$$

□

Main theorem

Theorem

$$\text{s-wXTT}^3 = \text{s-wXTT}^2$$

The composition hierarchy of simple extended top-down tree transf. collapses to the second level

Main theorem

Theorem

$$\text{s-wXTT}^3 = \text{s-wXTT}^2$$

The composition hierarchy of simple extended top-down tree transf. collapses to the second level

Proof idea:

$$\begin{aligned}
\text{s-wXTT}^3 &= \text{REL} \, ; \text{Bfus-wXTT}^3 && \text{(normal form)} \\
&\subseteq \text{REL} \, ; \underbrace{\text{Bs-wXTT}^2}_{\text{functional}} && \text{(unweighted result)} \\
&\subseteq \text{REL} \, ; \text{Bfus-wXTT}^2 && (\ldots\text{to be seen} \ldots) \\
&\subseteq \text{s-wXTT}^2 && \text{(normal form)}
\end{aligned}$$

Uniformizer lemma

function f uniformizer of relation R if $f \subseteq R$ and $\mathrm{dom}(f) = \mathrm{dom}(R)$

Lemma

Given relations $R_1, \ldots, R_n$ and functions $f_1, \ldots, f_n$ such that

- $R_1 ; \cdots ; R_n$ *is functional*
- $\mathrm{range}(R_j) \subseteq \mathrm{dom}(R_{j+1})$ *for all j*
- f_j *is a uniformizer of R_j for all j*

then $f_1 ; \cdots ; f_n = R_1 ; \cdots ; R_n$

Back to the proof

Lemma (Benedikt, Engelfriet, Maneth 2017)

Relations of Bs-wXTT *have uniformizers in* Bfs-wXTT

Back to the proof

Lemma (Benedikt, Engelfriet, Maneth 2017)

Relations of Bs-wXTT *have uniformizers in* Bfs-wXTT

Proof of main theorem:

$$
\begin{aligned}
\text{s-wXTT}^3 &= \text{REL} \,;\, \text{Bfus-wXTT}^3 && \text{(normal form)} \\
&\subseteq \text{REL} \,;\, \underbrace{\text{Bs-wXTT}^2}_{\text{functional}} && \text{(unweighted result)} \\
&\subseteq \text{REL} \,;\, \text{Bfs-wXTT}^2 && \text{(uniformizer lemma)} \\
&\subseteq \text{REL} \,;\, \text{Bfus-wXTT}^2 && (\text{Bfs-wXTT} = \text{Bfus-wXTT}) \\
&\subseteq \text{s-wXTT}^2 && \text{(normal form)}
\end{aligned}
$$

That's all folks!

Theorem (Full version of main theorem)

For all commutative semirings

- $\not\varepsilon\text{nsl-XTT}^3 = \not\varepsilon\text{nsl-XTT}^2$

That's all folks!

Theorem (Full version of main theorem)

For all commutative semirings

- $\not\in$nsl-XTT^3 $=$ $\not\in$nsl-XTT^2
- $\not\in$sl-XTT^3 $=$ $\not\in$sl-XTT^2

That's all folks!

Theorem (Full version of main theorem)

For all commutative semirings

- $\not\subseteq\text{nsl-XTT}^3 = \not\subseteq\text{nsl-XTT}^2$
- $\not\subseteq\text{sl-XTT}^3 = \not\subseteq\text{sl-XTT}^2$
- $(\not\subseteq\text{sl-XTT}^{\text{R}})^3 = (\not\subseteq\text{sl-XTT}^{\text{R}})^2$

That's all folks!

Theorem (Full version of main theorem)

For all commutative semirings

- $\not\varepsilon\text{nsl-XTT}^3 = \not\varepsilon\text{nsl-XTT}^2$
- $\not\varepsilon\text{sl-XTT}^3 = \not\varepsilon\text{sl-XTT}^2$
- $(\not\varepsilon\text{sl-XTT}^R)^3 = (\not\varepsilon\text{sl-XTT}^R)^2$
- $(\not\varepsilon\text{l-XTT}^R)^4 = (\not\varepsilon\text{l-XTT}^R)^3$

That's all folks!

Theorem (Full version of main theorem)

For all commutative semirings

- $\not\varepsilon\text{nsl-XTT}^3 = \not\varepsilon\text{nsl-XTT}^2$
- $\not\varepsilon\text{sl-XTT}^3 = \not\varepsilon\text{sl-XTT}^2$
- $(\not\varepsilon\text{sl-XTT}^R)^3 = (\not\varepsilon\text{sl-XTT}^R)^2$
- $(\not\varepsilon\text{l-XTT}^R)^4 = (\not\varepsilon\text{l-XTT}^R)^3$
- $\not\varepsilon\text{l-XTT}^5 = \not\varepsilon\text{l-XTT}^4$

That's all folks!

Theorem (Full version of main theorem)

For all commutative semirings

- $\not\varepsilon\text{nsl-XTT}^3 = \not\varepsilon\text{nsl-XTT}^2$
- $\not\varepsilon\text{sl-XTT}^3 = \not\varepsilon\text{sl-XTT}^2$
- $(\not\varepsilon\text{sl-XTT}^{\text{R}})^3 = (\not\varepsilon\text{sl-XTT}^{\text{R}})^2$
- $(\not\varepsilon\text{l-XTT}^{\text{R}})^4 = (\not\varepsilon\text{l-XTT}^{\text{R}})^3$
- $\not\varepsilon\text{l-XTT}^5 = \not\varepsilon\text{l-XTT}^4$

Thank you for your attention!