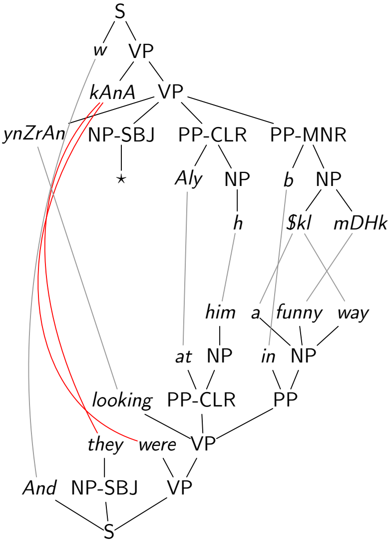
The Power of Regularity-Preserving Multi Bottom-up Tree Transducers

Andreas Maletti

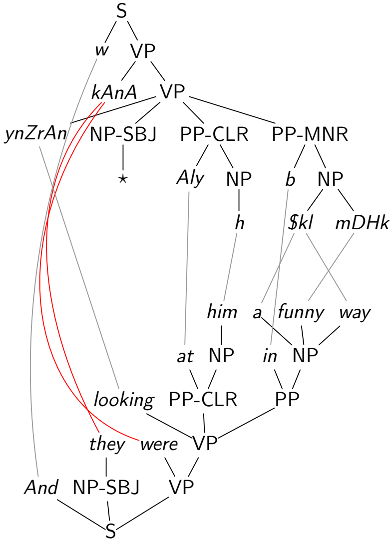
maletti@ims.uni-stuttgart.de

Gießen — August 2, 2014

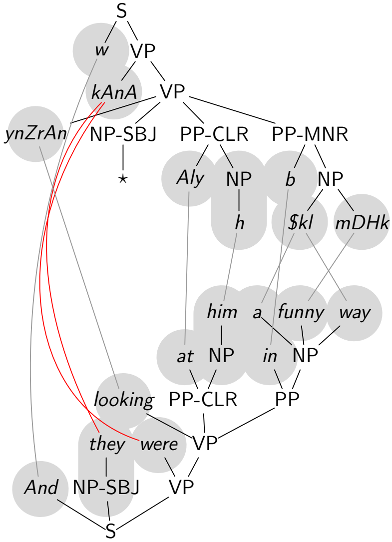
Syntax-based Statistical Machine Translation

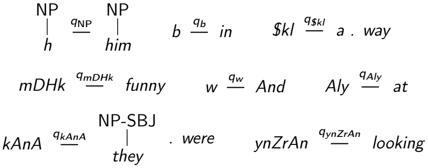


Syntax-based Statistical Machine Translation

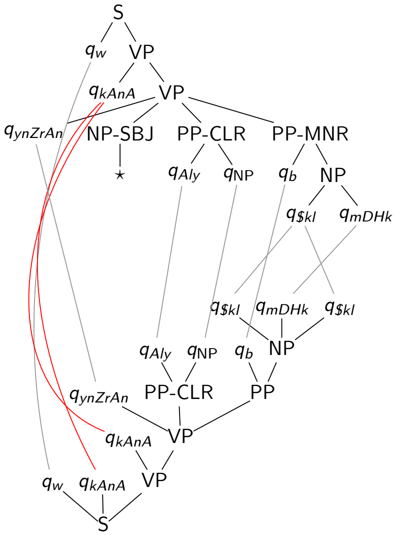


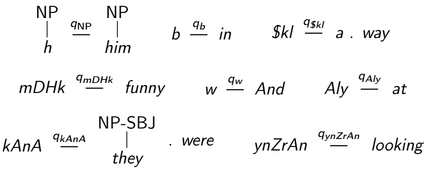
Syntax-based Statistical Machine Translation



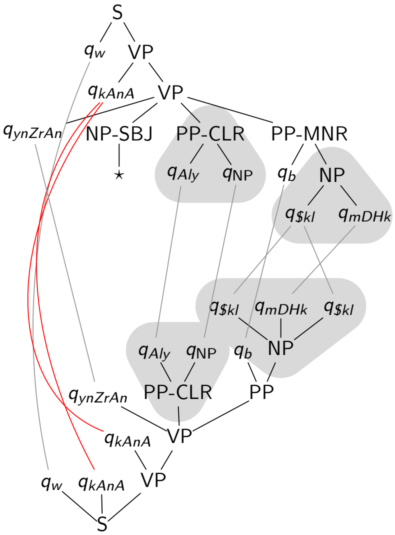


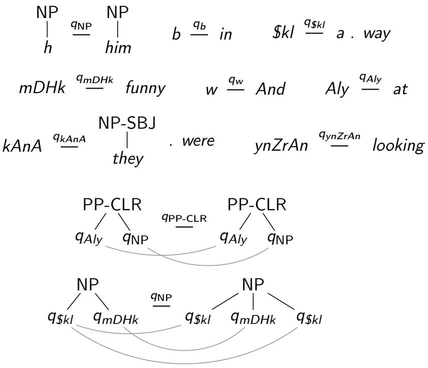
Syntax-based Statistical Machine Translation



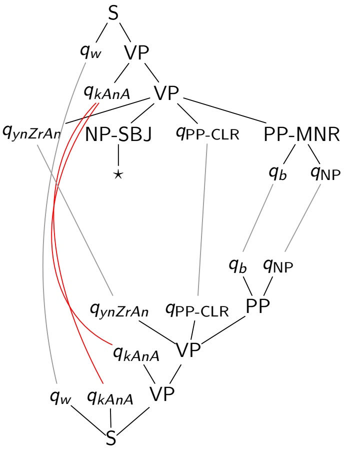


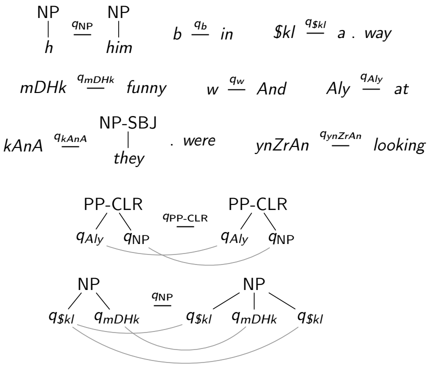
Syntax-based Statistical Machine Translation



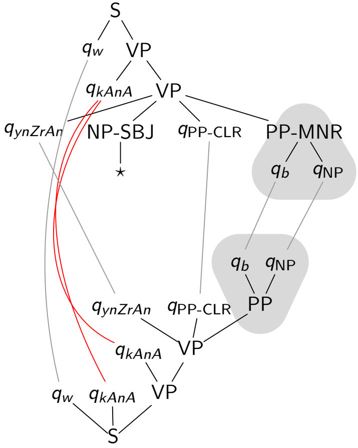


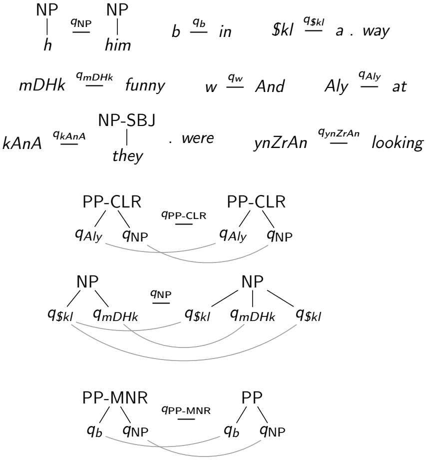
Syntax-based Statistical Machine Translation





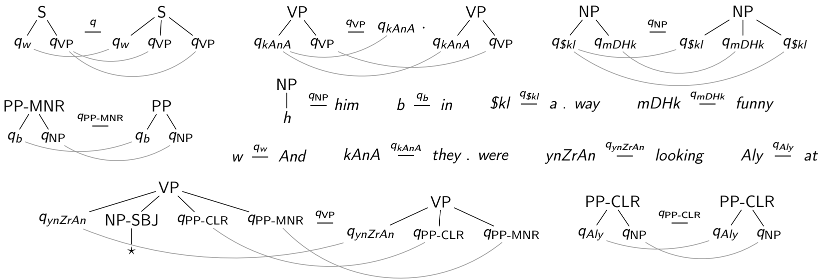
Syntax-based Statistical Machine Translation





Syntax-based Statistical Machine Translation

Extracted rules



Linear Extended Multi Bottom-up Tree Transducer

Definition (MBOT)

linear extended multi bottom-up tree transducer (Q, Σ, I, R)

- finite set Q — states
- alphabet Σ — input and output symbols
- $I \subseteq Q$ — initial states
- finite set $R \subseteq T_\Sigma(Q) \times Q \times T_\Sigma(Q)^*$ — rules
  - each $q \in Q$ occurs at most once in ℓ — $(\ell, q, \vec{r}) \in R$
  - each $q \in Q$ that occurs in $\vec{r}$ also occurs in ℓ — $(\ell, q, \vec{r}) \in R$

Linear Extended Multi Bottom-up Tree Transducer

Definition (Syntactic properties)

MBOT (Q, Σ, I, R) is

- linear extended top-down tree transducer with regular look-ahead (XTOP^R) if $|\vec{r}| \leq 1$ $\quad \forall(\ell, q, \vec{r}) \in R$
- linear extended top-down tree transducer (XTOP) if $|\vec{r}| = 1$ $\quad \forall(\ell, q, \vec{r}) \in R$

Linear Extended Multi Bottom-up Tree Transducer

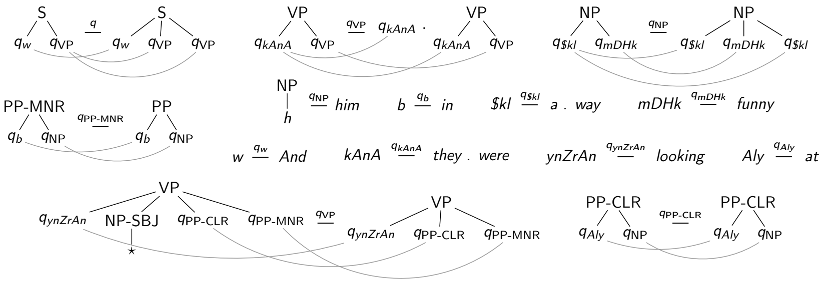
Definition (Syntactic properties)

MBOT (Q, Σ, I, R) is

- linear extended top-down tree transducer with regular look-ahead (XTOP^R) if $|\vec{r}| \leq 1$ $\quad \forall(\ell, q, \vec{r}) \in R$
- linear extended top-down tree transducer (XTOP) if $|\vec{r}| = 1$ $\quad \forall(\ell, q, \vec{r}) \in R$
- ε-free if $\ell \notin Q$ $\quad \forall(\ell, q, \vec{r}) \in R$

Linear Extended Multi Bottom-up Tree Transducer

Extracted rules



Properties

XTOP^R: ✗ XTOP: ✗ ε-free: ✓

Another Example

Example (textual)

MBOT $M = (Q, \Sigma, \{\star\}, R)$

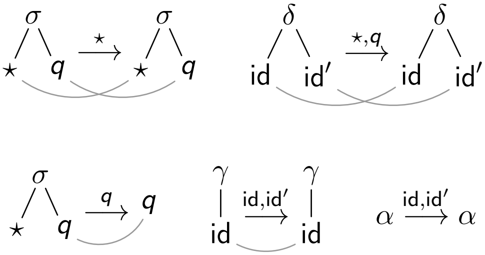
- $Q = \{\star, q, \mathrm{id}, \mathrm{id}'\}$
- $\Sigma = \{\sigma, \delta, \gamma, \alpha\}$
- the following rules in R:

$$\sigma(\star, q) \xrightarrow{\star} \sigma(\star, q) \qquad \sigma(\star, q) \xrightarrow{q} q$$

$$\delta(\mathrm{id}, \mathrm{id}') \xrightarrow{\star, q} \delta(\mathrm{id}, \mathrm{id}') \qquad \gamma(\mathrm{id}) \xrightarrow{\mathrm{id}, \mathrm{id}'} \gamma(\mathrm{id}) \qquad \alpha \xrightarrow{\mathrm{id}, \mathrm{id}'} \alpha$$

Another Example

Graphical representation

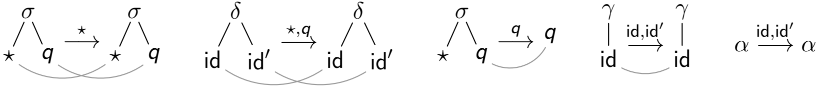


Properties

XTOP^R: ✓ XTOP: ✓ ε-free: ✓

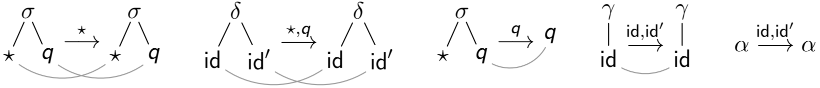
Semantics — Synchronous Generation

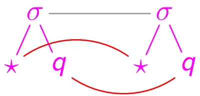
Rules



Semantics — Synchronous Generation

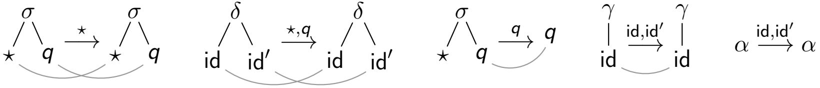
Rules

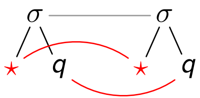




Semantics — Synchronous Generation

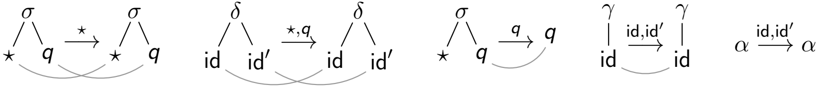
Rules

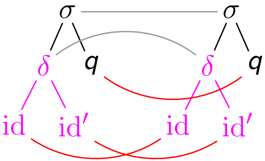




Semantics — Synchronous Generation

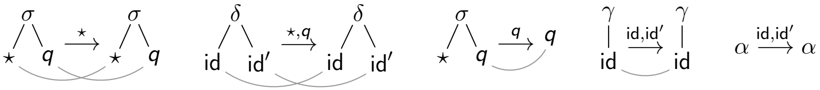
Rules

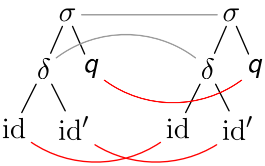




Semantics — Synchronous Generation

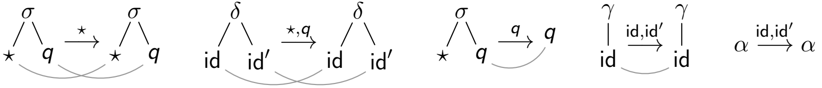
Rules

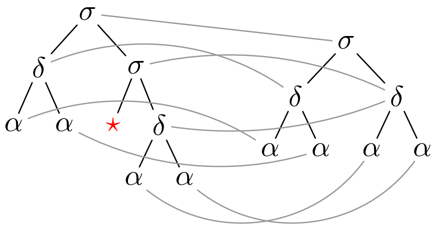




Semantics — Synchronous Generation

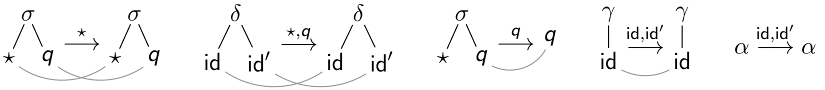
Rules

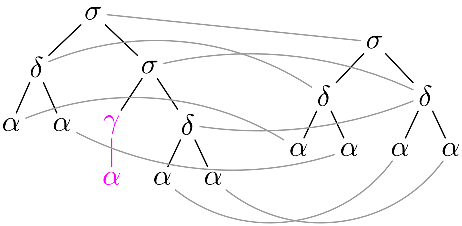




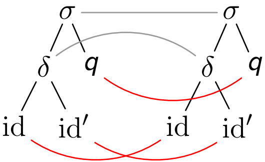
Semantics — Synchronous Generation

Rules





Semantics — Synchronous Generation



Definition (sentential forms)

$$\langle t, A, D, u \rangle$$

- $t \in T_\Sigma(Q)$ — input tree
- $A \subseteq \mathbb{N}^* \times \mathbb{N}^*$ — active links (red)
- $D \subseteq \mathbb{N}^* \times \mathbb{N}^*$ — disabled links (gray)
- $u \in T_\Sigma(Q)$ — output tree

Semantics — Synchronous Generation

Definition (Generation step)

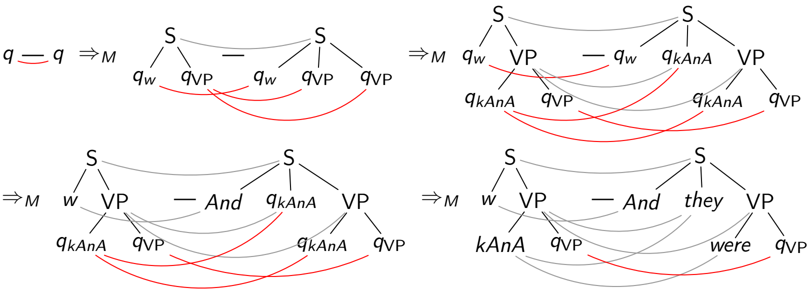
$$\langle t, A, D, u\rangle \Rightarrow_M \langle t', A', D', u'\rangle$$

if and only if $\exists q \in Q$, $\exists v \in \mathrm{pos}(t)$ labeled by q, and $\exists \ell \xrightarrow{q} \vec{r} \in P$

- $|\vec{r}| = |A(v)|$ and $\vec{w} = \overrightarrow{A(v)}$
- $t' = t[\ell]_v$ and $u' = u[\vec{r}]_{\vec{w}}$
- $A' = (A \setminus L) \cup \mathrm{links}_{v,\vec{w}}(\ell \xrightarrow{q} \vec{r})$ and $D' = D \cup L$ with

$$L = \{(v, w) \mid w \in A(v)\}$$

Semantics — Synchronous Generation



Semantics — Synchronous Generation

Definition

- state-computed dependencies:

$$M_q = \{\langle t, D, u\rangle \mid t, u \in T_\Sigma,\ \langle q, \{(\varepsilon, \varepsilon)\}, \emptyset, q\rangle \Rightarrow_M^* \langle t, \emptyset, D, u\rangle\}$$

- computed dependencies:

$$\mathrm{dep}(M) = \bigcup_{q \in I} M_q$$

Semantics — Synchronous Generation

Definition

- state-computed dependencies:

$$M_q = \{\langle t, D, u\rangle \mid t, u \in T_\Sigma,\ \langle q, \{(\varepsilon, \varepsilon)\}, \emptyset, q\rangle \Rightarrow_M^* \langle t, \emptyset, D, u\rangle\}$$

- computed dependencies:

$$\mathrm{dep}(M) = \bigcup_{q \in I} M_q$$

- computed transformation:

$$\tau_M = \{(t, u) \mid \langle t, D, u\rangle \in \mathrm{dep}(M)\}$$

The Problem

Definition (Regularity-preserving)

transformation $\tau \subseteq T_\Sigma \times T_\Sigma$ preserves regularity
if $\tau(L) = \{u \mid (t, u) \in \tau,\ t \in L\}$ is regular for all regular $L \subseteq T_\Sigma$

rp-MBOT = class of all regularity preserving transformations computable by MBOT

Compositions

- $\tau_1 \,;\, \tau_2 = \{(s, u) \mid \exists t \colon (s, t) \in \tau_1, (t, u) \in \tau_2\}$
- support modular development
- allow integration of external knowledge sources
- occur naturally in query rewriting

The Problem

Known:

- $XTOP^R \subsetneq MBOT$
- MBOT is closed under composition
- all $\tau \in XTOP^R$ preserve regularity

The Problem

Known:

- $\text{XTOP}^{\text{R}} \subsetneq \text{MBOT}$
- MBOT is closed under composition
- all $\tau \in \text{XTOP}^{\text{R}}$ preserve regularity

Question:

Is $(\text{XTOP}^{\text{R}})^* \subsetneq \text{rp-MBOT}$? $(\text{XTOP}^{\text{R}})^* \subseteq \text{rp-MBOT}$ is true

$$(\text{XTOP}^{\text{R}})^* = \bigcup_{k \geq 1} \underbrace{\text{XTOP}^{\text{R}} ; \cdots ; \text{XTOP}^{\text{R}}}_{k \text{ factors}}$$

The Problem

Motivation

- we use regularity-preserving MBOT for efficiency
  in our translation systems
- their power is currently not well understood

The Problem

Motivation

- we use regularity-preserving MBOT for efficiency
  in our translation systems
- their power is currently not well understood
- e.g. general MBOT can handle discontinuities
- Is this still possible with regularity-preserving MBOT?
  or do they have the same power as compositions of XTOP

Contents

Properties of Dependencies

Definition (Hierarchical properties)

A dependency $\langle t, D, u \rangle$ is

- input hierarchical if
  1. $w_2 \not< w_1$
  2. $\exists (v_1, w_1') \in D$ with $w_1' \leq w_2$

  for all $(v_1, w_1), (v_2, w_2) \in D$ with $v_1 < v_2$

Properties of Dependencies

Definition (Hierarchical properties)

A dependency $\langle t, D, u \rangle$ is

- input hierarchical if
  1. $w_2 \not< w_1$
  2. $\exists (v_1, w_1') \in D$ with $w_1' \leq w_2$

  for all $(v_1, w_1), (v_2, w_2) \in D$ with $v_1 < v_2$

Properties of Dependencies

Definition (Hierarchical properties)

A dependency $\langle t, D, u \rangle$ is

- input hierarchical if
  1. $w_2 \not< w_1$
  2. $\exists (v_1, w_1') \in D$ with $w_1' \leq w_2$

  for all $(v_1, w_1), (v_2, w_2) \in D$ with $v_1 < v_2$
- strictly input hierarchical if
  1. $v_1 < v_2$ implies $w_1 \leq w_2$
  2. $v_1 = v_2$ implies $w_1 \leq w_2$ or $w_2 \leq w_1$

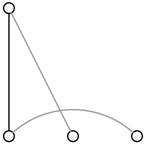
  for all $(v_1, w_1), (v_2, w_2) \in D$

Properties of Dependencies

Definition (Distance properties)

A dependency $\langle t, D, u \rangle$ is

- input link-distance bounded by $b \in \mathbb{N}$ if
  for all $(v_1, w_1), (v_1 v', w_2) \in D$ with $|v'| > b$
  $\exists (v_1 v, w_3) \in D$ such that $v < v'$ and $1 \leq |v| \leq b$

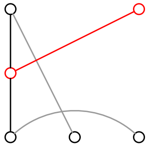


Properties of Dependencies

Definition (Distance properties)

A dependency $\langle t, D, u \rangle$ is

- input link-distance bounded by $b \in \mathbb{N}$ if
  for all $(v_1, w_1), (v_1 v', w_2) \in D$ with $|v'| > b$
  $\exists (v_1 v, w_3) \in D$ such that $v < v'$ and $1 \leq |v| \leq b$



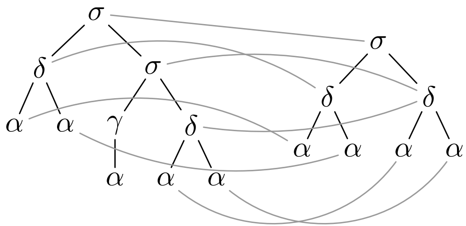
Properties of Dependencies

Definition (Distance properties)

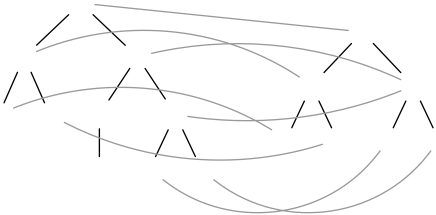
A dependency $\langle t, D, u\rangle$ is

- input link-distance bounded by $b \in \mathbb{N}$ if
  for all $(v_1, w_1), (v_1 v', w_2) \in D$ with $|v'| > b$
  $\exists (v_1 v, w_3) \in D$ such that $v < v'$ and $1 \leq |v| \leq b$
- strict input link-distance bounded by b if for all
  $v_1, v_1 v' \in \text{pos}(t)$ with $|v'| > b$
  $\exists (v_1 v, w_3) \in D$ such that $v < v'$ and $1 \leq |v| \leq b$

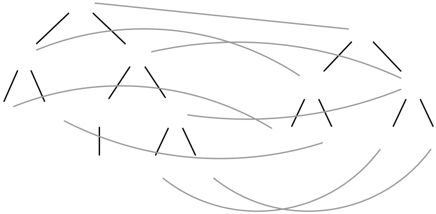
Properties of Dependencies



Properties of Dependencies

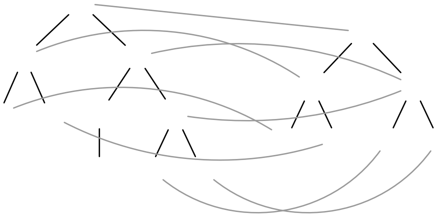


Properties of Dependencies



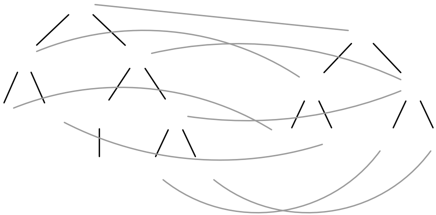
strictly input hierarchical

Properties of Dependencies



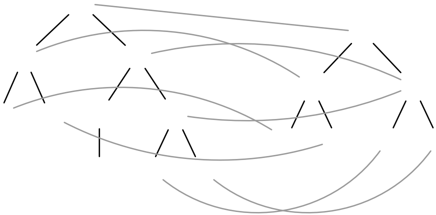
strictly input hierarchical and strictly output hierarchical

Properties of Dependencies



strictly input hierarchical and strictly output hierarchical
with strict input link-distance 2

Properties of Dependencies



strictly input hierarchical and strictly output hierarchical
with strict input link-distance 2 and strict output link-distance 1

Properties of Dependencies

	hierarchical		link-distance bounded	
Model \ Property	input	output	input	output
XTOP^R	strictly	strictly	✓	strictly
MBOT	✓	strictly	✓	strictly

Linking Theorem for ε-free XTOP^R

Theorem

Let $M_1, \dots, M_k$ be ε-free XTOP^R over Σ such that

$$\{(c[t_1, \dots, t_n]\,,\, c'[t_1, \dots, t_n]) \mid t_1, \dots, t_n \in T\} \subseteq \tau_{M_1} \,;\, \cdots \,;\, \tau_{M_k}$$

for some contexts $c, c' \in C_\Sigma(X_n)$ and special $T \subseteq T_\Sigma$.

$\forall 1 \leq i \leq k$, $\forall 1 \leq j \leq n$

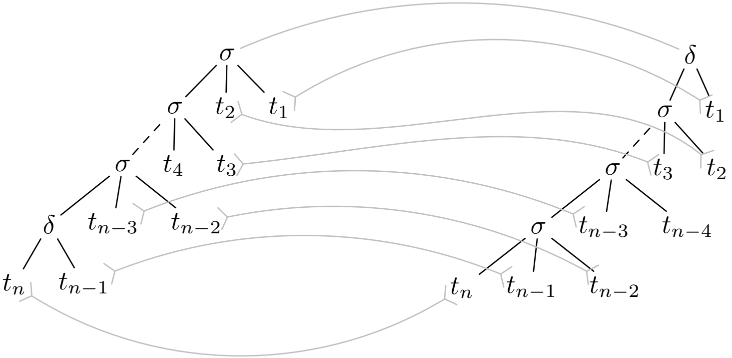
$\exists t_j \in T$, $\exists \langle u_{i-1}, D_i, u_i \rangle \in \mathrm{dep}(M_i)$, $\exists (v_{ji}, w_{ji}) \in D_i$ *such that*

- $u_0 = c[t_1, \dots, t_n]$ *and* $u_k = c'[t_1, \dots, t_n]$
- $\mathrm{pos}_{x_j}(c') \leq w_{jk}$
- $v_{ji} \leq w_{j(i-1)}$ *if* $i \geq 2$
- $\mathrm{pos}_{x_j}(c) \leq v_{j1}$

Linking Theorem for ε-free XTOP$^{\mathsf{R}}$

Corollary [see Sect. 3.4 in [Arnold, Dauchet 1982]]

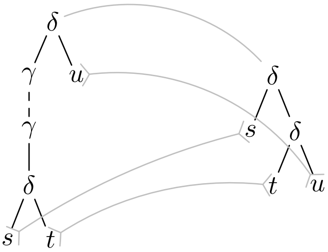
The illustrated tree transformation τ cannot be computed by any ε-free XTOP$^{\mathsf{R}}$



Linking Theorem for ε-free XTOP^{R}

Corollary [see Thm. 5.2 in [Maletti et al. 2009]]

The illustrated tree transformation τ cannot be computed by any ε-free XTOP^{R}



Contents

Topicalization

Example

- *It rained yesterday night.*

  Topicalized: *Yesterday night, it rained.*

Topicalization

Example

- *It rained yesterday night.*

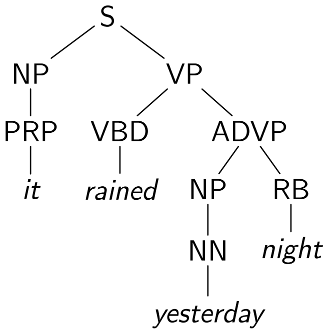
  Topicalized: *Yesterday night, it rained.*

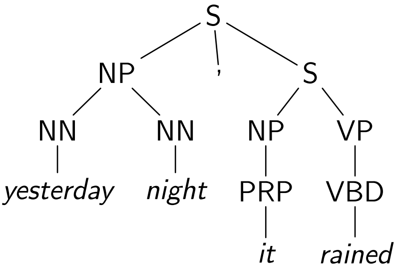
- *We toiled all day yesterday at the restaurant that charges extra for clean plates.*

  Topicalized: *At the restaurant that charges extra for clean plates, we toiled all day yesterday.*

Topicalization

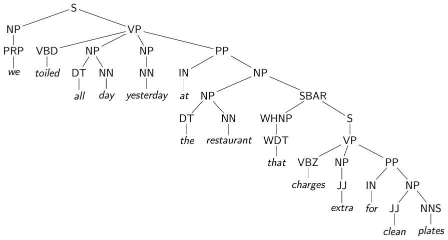
Example (on the tree level)

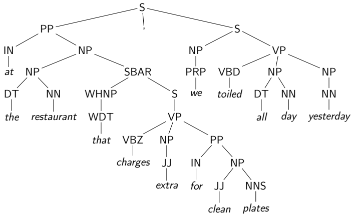




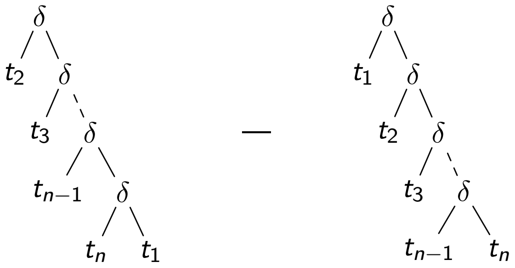
Topicalization

Example (on the tree level)





Abstract Topicalization



Abstract Topicalization

Theorem

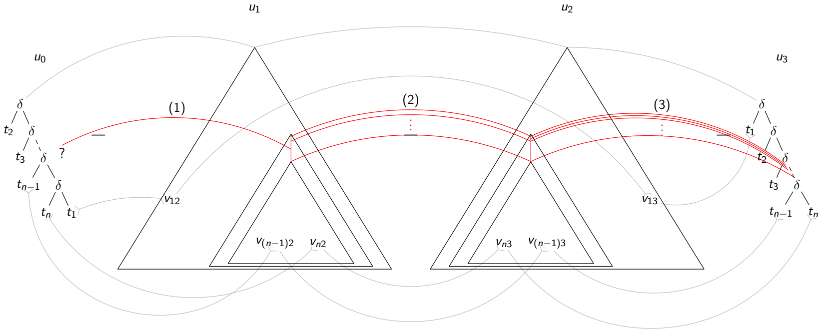
Abstract topicalization

- *preserves regularity and*
- *can be computed by an MBOT*

Abstract Topicalization

Theorem

Abstract topicalization cannot be computed by any composition chain of ε-free XTOP^R



It is known that 3 ε-free XTOP^R suffice

Consequence

Corollary

$$(\mathsf{XTOP}^{\mathsf{R}})^* \subsetneq \text{rp-MBOT}$$

Linking Theorem for ε-free MBOT

Theorem

Let $M = (Q, \Sigma, I, R)$ *be an* ε*-free MBOT such that*

$$\{(c[t_1, \ldots, t_n]\,,\, c'[t_1, \ldots, t_n]) \mid t_1, \ldots, t_n \in T\} \subseteq \tau_M$$

for some contexts $c, c' \in C_\Sigma(X_n)$ *and special* $T \subseteq T_\Sigma$*.*

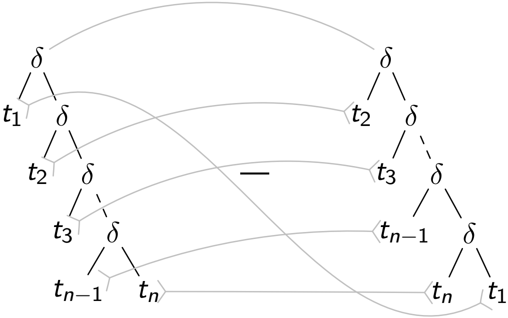
$\forall 1 \leq j \leq n$, $\exists t_j \in T$, $\exists \langle u, D, u' \rangle \in \mathrm{dep}(M)$, $\exists (v_j, w_j) \in D$ *with*

- $u = c[t_1, \ldots, t_n]$ *and* $u' = c'[t_1, \ldots, t_n]$
- $\mathrm{pos}_{x_j}(c) \leq v_j$
- $\mathrm{pos}_{x_j}(c') \leq w_j$

Linking Theorem for ε-free MBOT

Corollary

Inverse of topicalization cannot be computed by any ε-free MBOT



Summary & References

Summary

1. $(\text{XTOP}^{\text{R}})^* \subsetneq$ rp-MBOT
2. rp-MBOT not closed under inverses
3. What happens to invertable MBOT?

Summary & References

Summary

1. $(\mathrm{XTOP}^{\mathrm{R}})^* \subsetneq$ rp-MBOT
2. rp-MBOT not closed under inverses
3. What happens to invertable MBOT?

References

- J. Engelfriet, E. Lilin, ∼: Extended multi bottom-up tree transducers — Composition and decomposition. Acta Inf., 2009
- Z. Fülöp, ∼: Composition closure of ε-free linear extended top-down tree transducers. Proc. 17th DLT, LNCS 7907, 2013
- P. Koehn: Statistical machine translation. Cambridge Univ. Press, 2009
- ∼, J. Graehl, M. Hopkins, K. Knight: The power of extended top-down tree transducers. SIAM J. Comput., 2009