Derivation Trees of a Tree Adjoining Grammar

Sylvain Schmitz

LORIA, INRIA Nancy - Grand Est, Nancy, France

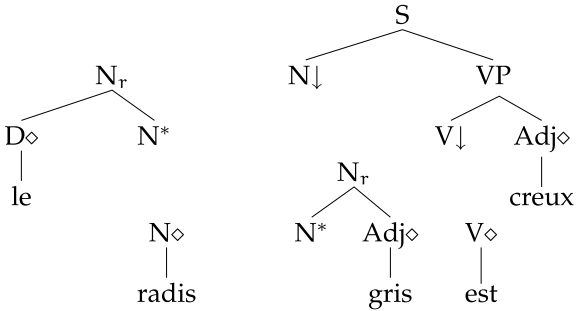
MOSTRARE Seminars, April 18, 2008

Contents

Tree Adjoining Grammars

Joshi et al. [1975], Joshi and Schabes [1997]

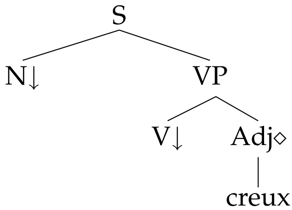
Elementary trees:



Tree Adjoining Grammars

Joshi et al. [1975], Joshi and Schabes [1997]

Two operations:

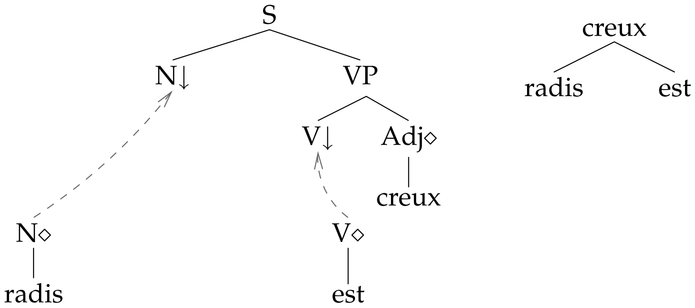


creux

Tree Adjoining Grammars

Joshi et al. [1975], Joshi and Schabes [1997]

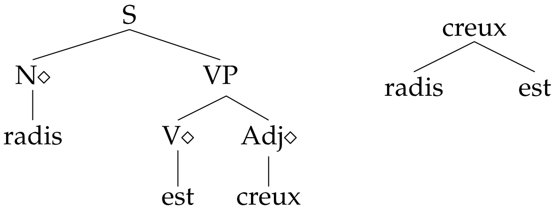
Two operations: substitution:



Tree Adjoining Grammars

Joshi et al. [1975], Joshi and Schabes [1997]

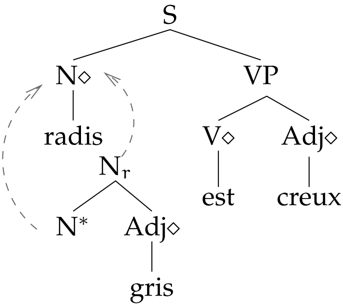
Two operations: substitution:

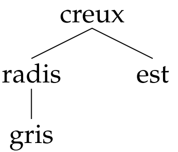


Tree Adjoining Grammars

Joshi et al. [1975], Joshi and Schabes [1997]

Two operations: adjunction:

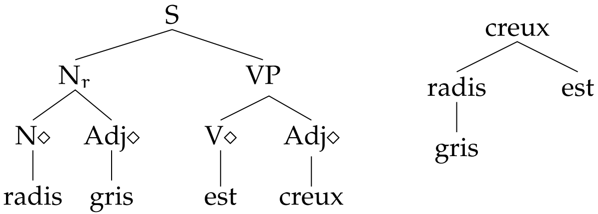




Tree Adjoining Grammars

Joshi et al. [1975], Joshi and Schabes [1997]

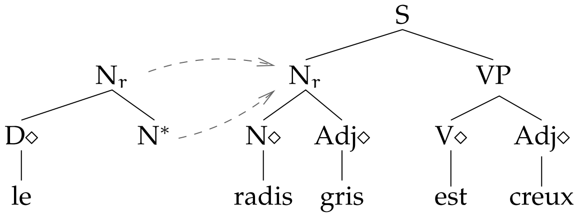
Two operations: adjunction:

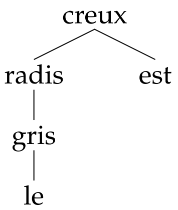


Tree Adjoining Grammars

Joshi et al. [1975], Joshi and Schabes [1997]

Two operations: adjunction:

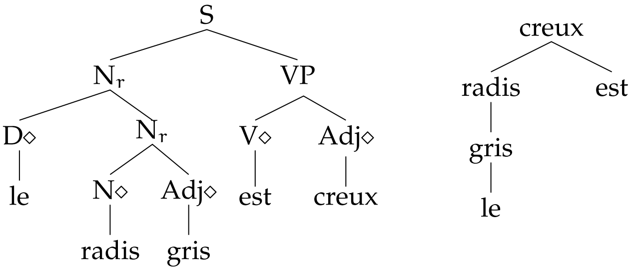




Tree Adjoining Grammars

Joshi et al. [1975], Joshi and Schabes [1997]

Derived and derivation trees:



Formal Properties

Strings

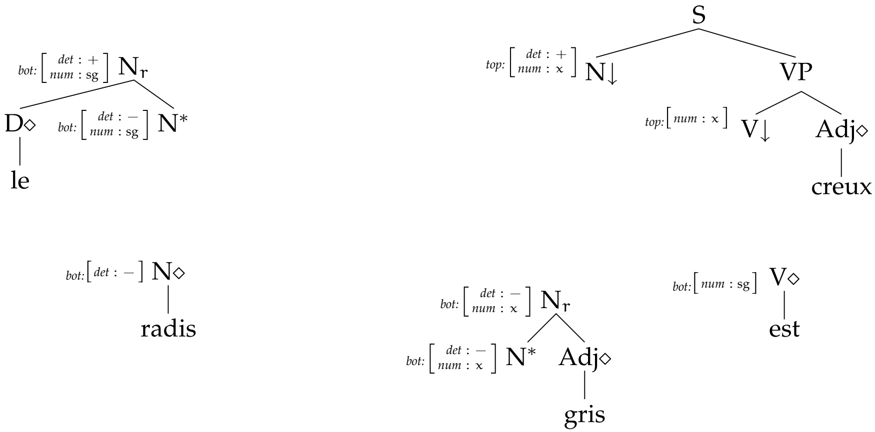
- beyond context-free: *mildly context-sensitive*
- polynomial time parsing: $\mathcal{O}(|G|\, n^6)$

Trees

- beyond regular
- a restricted class of context-free tree languages: *linear* and *monadic* [Kepster and Rogers, 2007]

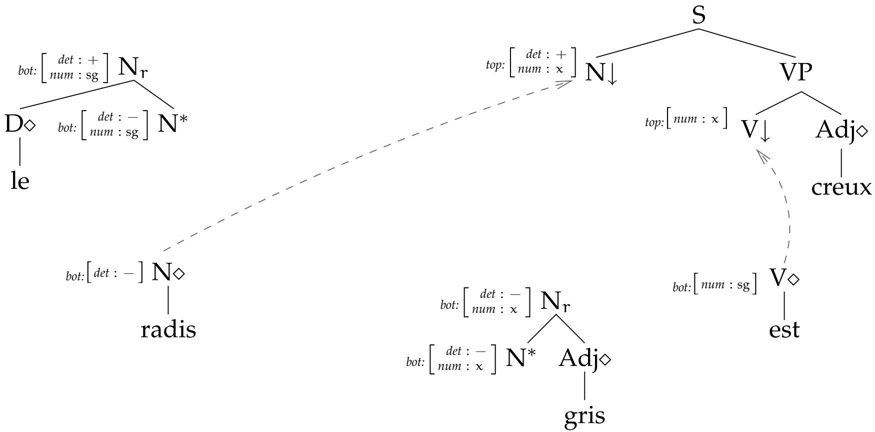
Feature-Based TAG

Vijay-Shanker and Joshi [1988]



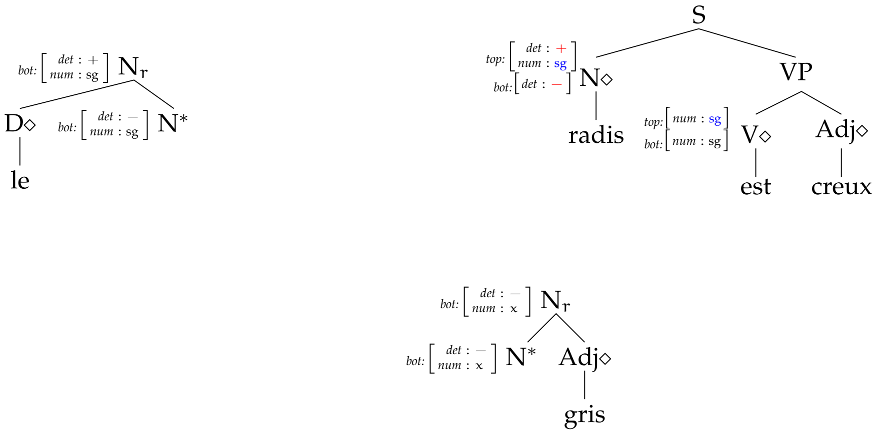
Feature-Based TAG

Vijay-Shanker and Joshi [1988]



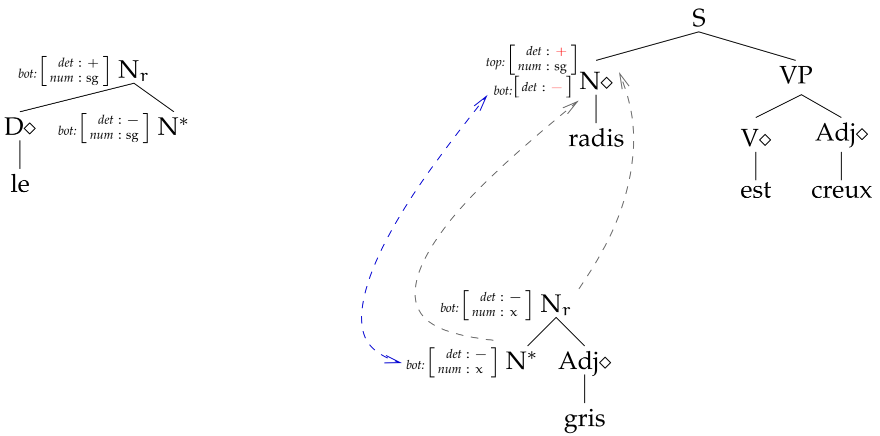
Feature-Based TAG

Vijay-Shanker and Joshi [1988]



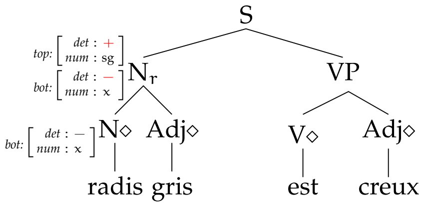
Feature-Based TAG

Vijay-Shanker and Joshi [1988]



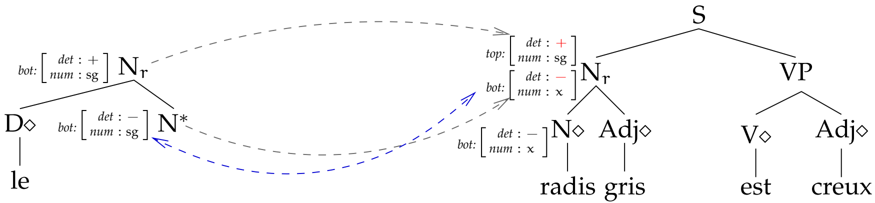
Feature-Based TAG

Vijay-Shanker and Joshi [1988]



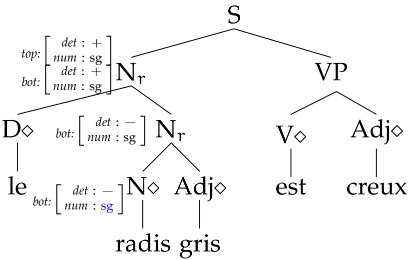
Feature-Based TAG

Vijay-Shanker and Joshi [1988]



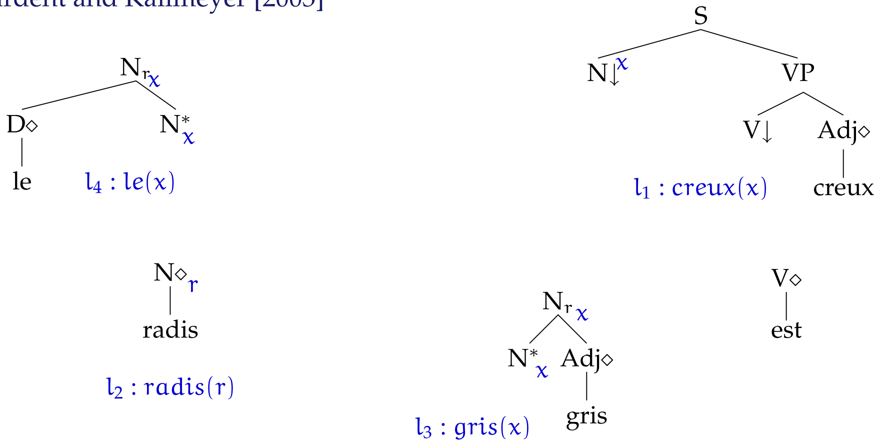
Feature-Based TAG

Vijay-Shanker and Joshi [1988]



Semantics Through Unification

Gardent and Kallmeyer [2003]

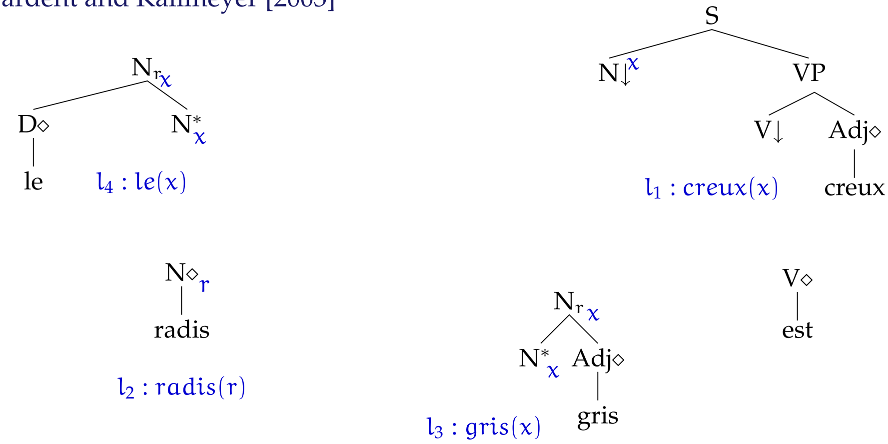


Yields flat semantics:

$$l_1 : \mathrm{creux}(r) \wedge l_2 : \mathrm{radis}(r) \wedge l_3 : \mathrm{gris}(r) \wedge l_4 : \mathrm{le}(r)$$

Semantics Through Unification

Gardent and Kallmeyer [2003]



Yields flat semantics:

$$l_1 : creux(r) \wedge l_2 : radis(r) \wedge l_3 : gris(r) \wedge l_4 : le(r)$$

Two-level Syntax

Pogodalla [2004], Shieber [2006], and many others

- derivation trees as intermediate representations
- morphisms / transductions into
    1. derived trees (or strings)
    2. semantics
- inverse morphisms / transductions for parsing and generation

Derivation Tree Language (1)

(reduced) Regular tree grammar

$S_I \to \text{creux}(S_A, N_I, VP_A, V_I, Adj_A)$

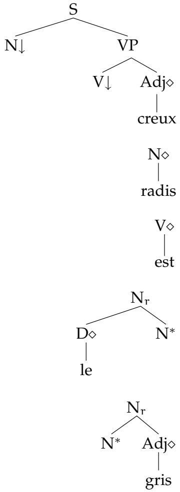
$N_I \to \text{radis}(N_A)$

$V_I \to \text{est}(V_A)$

$N_A \to \text{le}(N_A, D_A)$

$N_A \to \text{gris}(N_A, Adj_A)$

$Adj_A \to \varepsilon(),\ D_A \to \varepsilon(),\ N_A \to \varepsilon(),\ S_A \to \varepsilon(),\ V_A \to \varepsilon(),\ VP_A \to \varepsilon()$



Derivation Tree Language (1)

(reduced) Regular tree grammar

$S_I \rightarrow \mathrm{creux}(S_A, N_I, VP_A, V_I, Adj_A)$

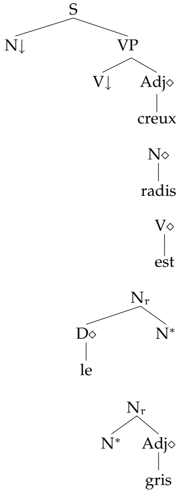
$N_I \rightarrow \mathrm{radis}(N_A)$

$V_I \rightarrow \mathrm{est}(V_A)$

$N_A \rightarrow \mathrm{le}(N_A, D_A)$

$N_A \rightarrow \mathrm{gris}(N_A, Adj_A)$

$Adj_A \rightarrow \varepsilon()$, $D_A \rightarrow \varepsilon()$, $N_A \rightarrow \varepsilon()$, $S_A \rightarrow \varepsilon()$, $V_A \rightarrow \varepsilon()$, $VP_A \rightarrow \varepsilon()$



Derivation Tree Language (1)

(reduced) Regular tree grammar

$$S_I \rightarrow \text{creux}(S_A, N_I, VP_A, V_I, \text{Adj}_A)$$

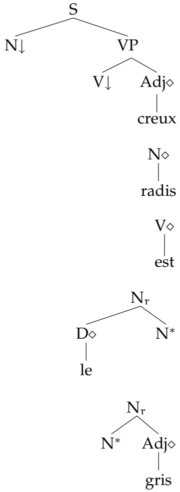
$$N_I \rightarrow \text{radis}(N_A)$$

$$V_I \rightarrow \text{est}(V_A)$$

$$N_A \rightarrow \text{le}(N_A, D_A)$$

$$N_A \rightarrow \text{gris}(N_A, \text{Adj}_A)$$

$$\text{Adj}_A \rightarrow \varepsilon(),\ D_A \rightarrow \varepsilon(),\ N_A \rightarrow \varepsilon(),\ S_A \rightarrow \varepsilon(),\ V_A \rightarrow \varepsilon(),\ VP_A \rightarrow \varepsilon()$$



Derivation Tree Language (1)

(reduced) Regular tree grammar

$$S_I \rightarrow \text{creux}(S_A, N_I, VP_A, V_I, Adj_A)$$

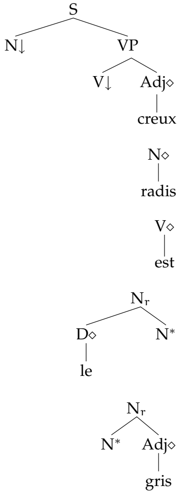
$$N_I \rightarrow \text{radis}(N_A)$$

$$V_I \rightarrow \text{est}(V_A)$$

$$N_A \rightarrow \text{le}(N_A, D_A)$$

$$N_A \rightarrow \text{gris}(N_A, Adj_A)$$

$$Adj_A \rightarrow \varepsilon(),\ D_A \rightarrow \varepsilon(),\ N_A \rightarrow \varepsilon(),\ S_A \rightarrow \varepsilon(),\ V_A \rightarrow \varepsilon(),\ VP_A \rightarrow \varepsilon()$$



Derivation Tree Language (1)

(reduced) Regular tree grammar

$$S_I \rightarrow \text{creux}(S_A, N_I, VP_A, V_I, Adj_A)$$

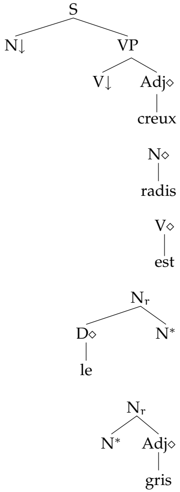
$$N_I \rightarrow \text{radis}(N_A)$$

$$V_I \rightarrow \text{est}(V_A)$$

$$N_A \rightarrow \text{le}(N_A, D_A)$$

$$N_A \rightarrow \text{gris}(N_A, Adj_A)$$

$$Adj_A \rightarrow \varepsilon(),\ D_A \rightarrow \varepsilon(),\ N_A \rightarrow \varepsilon(),\ S_A \rightarrow \varepsilon(),\ V_A \rightarrow \varepsilon(),\ VP_A \rightarrow \varepsilon()$$



Derivation Tree Language (1)

(reduced) Regular tree grammar

$$S_I \rightarrow \text{creux}(S_A, N_I, VP_A, V_I, \text{Adj}_A)$$

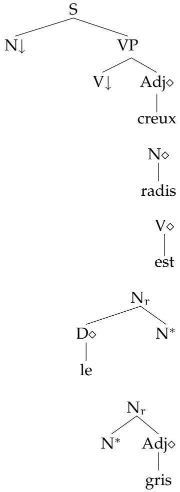
$$N_I \rightarrow \text{radis}(N_A)$$

$$V_I \rightarrow \text{est}(V_A)$$

$$N_A \rightarrow \text{le}(N_A, D_A)$$

$$N_A \rightarrow \text{gris}(N_A, \text{Adj}_A)$$

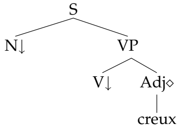
$$\text{Adj}_A \rightarrow \varepsilon(),\ D_A \rightarrow \varepsilon(),\ N_A \rightarrow \varepsilon(),\ S_A \rightarrow \varepsilon(),\ V_A \rightarrow \varepsilon(),\ VP_A \rightarrow \varepsilon()$$



Derivation Tree Language (1)

(reduced) Regular tree grammar

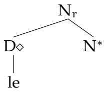
$$S_I \rightarrow \text{creux}(S_A, N_I, VP_A, V_I, Adj_A)$$



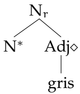
$$N_I \rightarrow \text{radis}(N_A)$$

$$V_I \rightarrow \text{est}(V_A)$$

$$N_A \rightarrow \text{le}(N_A, D_A)$$



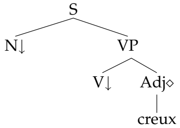
$$N_A \rightarrow \text{gris}(N_A, Adj_A)$$



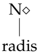
$$Adj_A \rightarrow \varepsilon(),\ D_A \rightarrow \varepsilon(),\ N_A \rightarrow \varepsilon(),\ S_A \rightarrow \varepsilon(),\ V_A \rightarrow \varepsilon(),\ VP_A \rightarrow \varepsilon()$$

Derivation Tree Language (1)

(reduced) Regular tree grammar

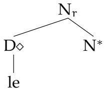


$$S_I \rightarrow \text{creux}(S_A, N_I, VP_A, V_I, Adj_A)$$

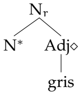


$$N_I \rightarrow \text{radis}(N_A)$$

$$V_I \rightarrow \text{est}(V_A)$$



$$N_A \rightarrow \text{le}(N_A, D_A)$$



$$N_A \rightarrow \text{gris}(N_A, Adj_A)$$

$$Adj_A \rightarrow \varepsilon(),\ D_A \rightarrow \varepsilon(),\ N_A \rightarrow \varepsilon(),\ S_A \rightarrow \varepsilon(),\ V_A \rightarrow \varepsilon(),\ VP_A \rightarrow \varepsilon()$$

From Derivations to Derived Trees

Macro tree transducer [Shieber, 2006]

$$q_I(\mathrm{creux}(x_1, x_2)) \to S(q_I(x_1),\ VP(q_I(x_2),\ Adj(\mathrm{creux})))$$

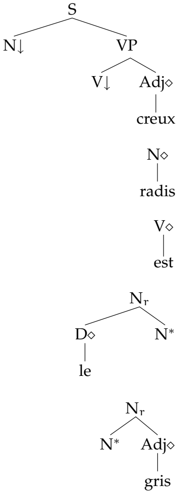
$$q_I(\mathrm{radis}(x)) \to q_A(x,\ N(\mathrm{radis}))$$

$$q_I(\mathrm{est}()) \to V(\mathrm{est})$$

$$q_A(\mathrm{le}(x),\ y) \to q_A(x,\ N(D(\mathrm{le}),\ y))$$

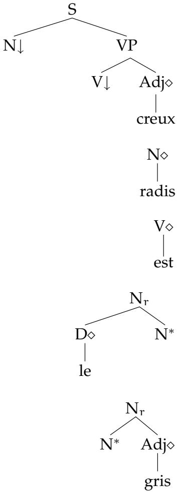
$$q_A(\mathrm{gris}(x),\ y) \to q_A(x,\ N(y,\ Adj(\mathrm{gris})))$$

$$q_A(\varepsilon(),\ y) \to y$$



From Derivations to Derived Trees

Macro tree transducer [Shieber, 2006]



$$q_I(\mathrm{creux}(x_1, x_2)) \to S(q_I(x_1),\ VP(q_I(x_2),\ \mathrm{Adj}(\mathrm{creux})))$$

$$q_I(\mathrm{radis}(x)) \to q_A(x,\ N(\mathrm{radis}))$$

$$q_I(\mathrm{est}()) \to V(\mathrm{est})$$

$$q_A(\mathrm{le}(x),\ y) \to q_A(x,\ N(D(\mathrm{le}),\ y))$$

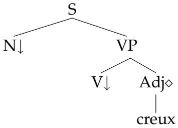
$$q_A(\mathrm{gris}(x),\ y) \to q_A(x,\ N(y,\ \mathrm{Adj}(\mathrm{gris})))$$

$$q_A(\varepsilon(),\ y) \to y$$

From Derivations to Derived Trees

Macro tree transducer [Shieber, 2006]

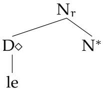
$$q_I(\mathrm{creux}(x_1, x_2)) \to S(q_I(x_1),\ VP(q_I(x_2),\ \mathrm{Adj}(\mathrm{creux})))$$



$$q_I(\mathrm{radis}(x)) \to q_A(x,\ N(\mathrm{radis}))$$

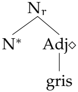
$$q_I(\mathrm{est}()) \to V(\mathrm{est})$$

$$q_A(\mathrm{le}(x),\ y) \to q_A(x,\ N(D(\mathrm{le}),\ y))$$



$$q_A(\mathrm{gris}(x),\ y) \to q_A(x,\ N(y,\ \mathrm{Adj}(\mathrm{gris})))$$

$$q_A(\varepsilon(),\ y) \to y$$



From Derivations to Semantics?

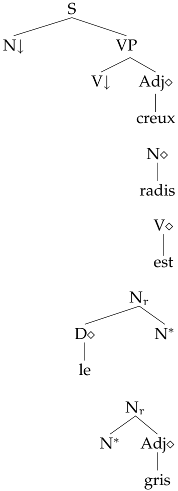
$$q(\mathrm{creux}(x_1, x_2),\ r) \to \wedge(\mathrm{creux}(r),\ q(x_1,\ r))$$

$$q(\mathrm{radis}(x),\ r) \to \wedge(\mathrm{radis}(r),\ q(x,\ r))$$

$$q(\mathrm{le}(x),\ r) \to \wedge(\mathrm{le}(r),\ q(x,\ r))$$

$$q(\mathrm{gris}(x),\ r) \to \wedge(\mathrm{gris}(r),\ q(x,\ r))$$

$$q(\varepsilon(),\ r) \to \mathrm{true}$$



From Derivations to Semantics?

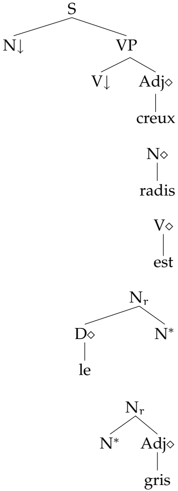
$$q(\text{creux}(x_1, x_2),\ r) \rightarrow \bigwedge(\text{creux}(r),\ q(x_1,\ r))$$

$$q(\text{radis}(x),\ r) \rightarrow \bigwedge(\text{radis}(r),\ q(x,\ r))$$

$$q(\text{le}(x),\ r) \rightarrow \bigwedge(\text{le}(r),\ q(x,\ r))$$

$$q(\text{gris}(x),\ r) \rightarrow \bigwedge(\text{gris}(r),\ q(x,\ r))$$

$$q(\varepsilon(),\ r) \rightarrow \text{true}$$



From Derivations to Semantics?

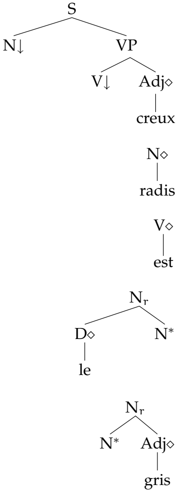
$$q(\text{creux}(x_1, x_2),\ r) \rightarrow \wedge(\text{creux}(r),\ q(x_1,\ r))$$

$$q(\text{radis}(x),\ r) \rightarrow \wedge(\text{radis}(r),\ q(x,\ r))$$

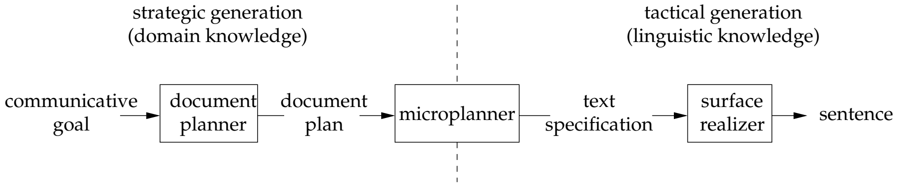
$$q(\text{le}(x),\ r) \rightarrow \wedge(\text{le}(r),\ q(x,\ r))$$

$$q(\text{gris}(x),\ r) \rightarrow \wedge(\text{gris}(r),\ q(x,\ r))$$

$$q(\varepsilon(),\ r) \rightarrow \text{true}$$



Sentence Generation Pipeline



Surface Realization Algorithms

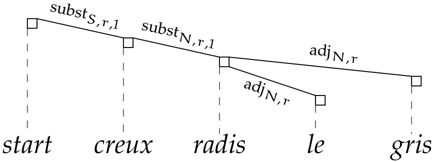
E.g. Carroll et al. [1999], Gardent and Kow [2007], ...

1. select all elementary trees that subsume parts of the input semantics
2. compute all combinations—tabulate intermediate results
3. extract sentences

NP-complete! [Koller and Striegnitz, 2002]

Generation as Dependency Parsing

Koller and Striegnitz [2002]: Topological Dependency Grammar



atom	labels	valency
start	$\emptyset$	$\{subst_{S,r,1}\}$
creux	$\{subst_{S,r,1}\}$	$\{subst_{N,r,1}\}$
radis	$\{subst_{N,r,1}\}$	$\{adj^{*}_{N,r}\}$
le	$\{adj_{N,r}\}$	$\emptyset$
gris	$\{adj_{N,r}\}$	$\emptyset$

Generation as Planning

Koller and Stone [2007]: Planning Domain Definition Language

Action start().

Effect: subst(S,r), sem(creux,r), sem(radis,r), sem(le,r), sem(gris,r)

Action creux(r). Precond: sem(creux,r), subst(S,r)

Effect: ¬sem(creux,r), ¬subst(S,r), subst(N,r), subst(V)

Action radis(r). Precond: sem(radis,r), subst(N,r)

Effect: ¬sem(radis,r), ¬subst(N,r), canadjoin(N,r)

Action est(). Precond: subst(V)

Effect: ¬subst(V)

Action le(r). Precond: sem(le,r), canadjoin(N,r)

Effect: ¬sem(le,r), ¬mustadjoin(N,r)

Action gris(r). Precond: sem(gris,r), canadjoin(N,r)

Effect: ¬sem(gris,r), ¬mustadjoin(N,r)

Action end(). Precond: $\forall$A,u. ¬sem(A,u) $\wedge$¬subst(A,u) $\wedge$¬mustadjoin(A,u)

Features in RTG

- $\mathcal{D}$: set of feature structures
- work on terms over $N \times \mathcal{D}$
- rules of form $(A, d) \rightarrow \alpha((B_1, d'_1), \ldots, (B_n, d'_n))$
- rewrites $(s, e) \Rightarrow (t, e')$ maintain a global environment for substitutions in feature structures

$$s = C[(A, d')],\ t = C[\alpha((B_1, \sigma(d'_1)), \ldots, (B_n, \sigma(d'_n)))]$$
$$\sigma = \mathsf{mgu}(d, e(d')),\ e' = \sigma \circ e$$

- language

$$L(G) = \{t \in T(\mathcal{F}) \mid \exists e, ((S, \top), id) \Rightarrow^* (t, e)\}$$

Derivation Tree Language (2)

$$(S_I, \top) \rightarrow \text{creux}\left(N_I \left[top : \begin{bmatrix} det : + \\ num : x \end{bmatrix} \right], V_I \left[top : [\, num : x \,] \right]\right)$$

$$N_I \left[top : t \right] \rightarrow \text{radis}\left(N_A \begin{bmatrix} top : t \\ bot : [\, det : - \,] \end{bmatrix}\right)$$

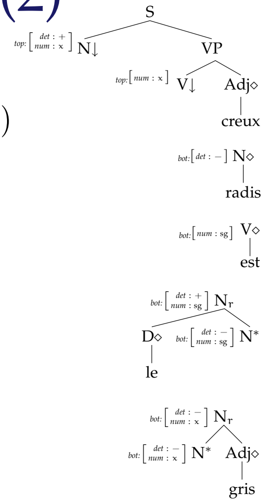
$$V_I \left[top : t \right] \rightarrow \text{est}\left(V_A \begin{bmatrix} top : t \\ bot : [\, num : sg \,] \end{bmatrix}\right)$$

$$N_A \begin{bmatrix} top : t \\ bot : \begin{bmatrix} det : - \\ num : sg \end{bmatrix} \end{bmatrix} \rightarrow \text{le}\left(N_A \begin{bmatrix} top : t \\ bot : \begin{bmatrix} det : + \\ num : sg \end{bmatrix} \end{bmatrix}\right)$$

$$N_A \begin{bmatrix} top : t \\ bot : \begin{bmatrix} det : - \\ num : x \end{bmatrix} \end{bmatrix} \rightarrow \text{gris}\left(N_A \begin{bmatrix} top : t \\ bot : \begin{bmatrix} det : - \\ num : x \end{bmatrix} \end{bmatrix}\right)$$

$$N_A \begin{bmatrix} top : x \\ bot : x \end{bmatrix} \rightarrow \varepsilon()$$

$$V_A \begin{bmatrix} top : x \\ bot : x \end{bmatrix} \rightarrow \varepsilon()$$



Example

$$((\mathrm{S},\top),\mathit{id})$$

$$\left(\mathrm{creux}\left(\mathrm{N}\begin{bmatrix}\mathtt{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0\mathtt{x}\end{bmatrix}\end{bmatrix},\mathrm{V}\begin{bmatrix}\mathtt{top}:\begin{bmatrix}\mathit{num}:0\mathtt{x}\end{bmatrix}\end{bmatrix}\right),\mathit{id}\right)$$

$$\left(\mathrm{radis}\left(\mathrm{N}_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0\mathtt{x}\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{det}:-\end{bmatrix}\end{bmatrix}\right),\left\{1\mathtt{t}\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0\mathtt{x}\end{bmatrix}\right\}\right)$$

$$\left(\mathrm{gris}\left(\mathrm{N}_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0\mathtt{x}\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{det}:-\\ \mathit{num}:11\mathtt{x}\end{bmatrix}\end{bmatrix}\right),\left\{11\mathtt{t}\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0\mathtt{x}\end{bmatrix}\right\}\right)$$

$$\left(\mathrm{le}\left(\mathrm{N}_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0\mathtt{x}\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:\mathrm{sg}\end{bmatrix}\end{bmatrix}\right),\left\{111\mathtt{t}\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0\mathtt{x}\end{bmatrix},11\mathtt{x}\mapsto\mathrm{sg}\right\}\right)$$

$$\left(\varepsilon,\left\{1111\mathtt{x}\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:\mathrm{sg}\end{bmatrix},0\mathtt{x}\mapsto\mathrm{sg}\right\}\right)$$

$$\left(\mathrm{est}\left(\mathrm{V}_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{num}:\mathrm{sg}\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{num}:\mathrm{sg}\end{bmatrix}\end{bmatrix}\right),\left\{2\mathtt{t}\mapsto\begin{bmatrix}\mathit{num}:\mathrm{sg}\end{bmatrix}\right\}\right)$$

$$\left(\varepsilon,\left\{21\mathtt{x}\mapsto\begin{bmatrix}\mathtt{num}:\mathrm{sg}\end{bmatrix}\right\}\right)$$

Example

- $((S,\top),id)$
  - $\left(\text{creux}\left(N\left[\text{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\right], V\left[\text{top}:\left[\mathit{num}:0x\right]\right]\right), id\right)$
    - $\left(\text{radis}\left(N_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\\ \mathit{bot}:\left[\mathit{det}:-\right]\end{bmatrix}\right), \left\{1t\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\right\}\right)$
      - $\left(\text{gris}\left(N_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{det}:-\\ \mathit{num}:11x\end{bmatrix}\end{bmatrix}\right), \left\{11t\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\right\}\right)$
        - $\left(\text{le}\left(N_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:\text{sg}\end{bmatrix}\end{bmatrix}\right), \left\{111t\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}, 11x\mapsto\text{sg}\right\}\right)$
          - $\left(\varepsilon, \left\{1111x\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:\text{sg}\end{bmatrix}, 0x\mapsto\text{sg}\right\}\right)$
    - $\left(\text{est}\left(V_A\begin{bmatrix}\mathit{top}:\left[\mathit{num}:\text{sg}\right]\\ \mathit{bot}:\left[\mathit{num}:\text{sg}\right]\end{bmatrix}\right), \left\{2t\mapsto\left[\mathit{num}:\text{sg}\right]\right\}\right)$
      - $\left(\varepsilon, \left\{21x\mapsto\left[\text{num}:\text{sg}\right]\right\}\right)$

Example

$$((S,\top),id)$$

$$\left(\text{creux}\left(N\begin{bmatrix}\text{top}:\begin{bmatrix}det:+\\num:0x\end{bmatrix}\end{bmatrix},V\begin{bmatrix}\text{top}:\begin{bmatrix}num:0x\end{bmatrix}\end{bmatrix}\right),id\right)$$

$$\left(\text{radis}\left(N_A\begin{bmatrix}top:\begin{bmatrix}det:+\\num:0x\end{bmatrix}\\bot:\begin{bmatrix}det:-\end{bmatrix}\end{bmatrix}\right),\left\{1t\mapsto\begin{bmatrix}det:+\\num:0x\end{bmatrix}\right\}\right)$$

$$\left(\text{est}\left(V_A\begin{bmatrix}top:\begin{bmatrix}num:sg\end{bmatrix}\\bot:\begin{bmatrix}num:sg\end{bmatrix}\end{bmatrix}\right),\left\{2t\mapsto\begin{bmatrix}num:sg\end{bmatrix}\right\}\right)$$

$$\left(\text{gris}\left(N_A\begin{bmatrix}top:\begin{bmatrix}det:+\\num:0x\end{bmatrix}\\bot:\begin{bmatrix}det:-\\num:11x\end{bmatrix}\end{bmatrix}\right),\left\{11t\mapsto\begin{bmatrix}det:+\\num:0x\end{bmatrix}\right\}\right)$$

$$\left(\varepsilon,\left\{21x\mapsto\begin{bmatrix}num:sg\end{bmatrix}\right\}\right)$$

$$\left(\text{le}\left(N_A\begin{bmatrix}top:\begin{bmatrix}det:+\\num:0x\end{bmatrix}\\bot:\begin{bmatrix}det:+\\num:sg\end{bmatrix}\end{bmatrix}\right),\left\{111t\mapsto\begin{bmatrix}det:+\\num:0x\end{bmatrix},11x\mapsto sg\right\}\right)$$

$$\left(\varepsilon,\left\{1111x\mapsto\begin{bmatrix}det:+\\num:sg\end{bmatrix},0x\mapsto sg\right\}\right)$$

Example

- $((S, \top), id)$
  - $\left(\text{creux}\left(N \left[\text{top} : \begin{bmatrix} det : + \\ num : 0x \end{bmatrix}\right], V \left[\text{top} : \left[num : 0x\right]\right]\right), id\right)$
    - $\left(\text{radis}\left(N_A \begin{bmatrix} top : \begin{bmatrix} det : + \\ num : 0x \end{bmatrix} \\ bot : \left[det : -\right] \end{bmatrix}\right), \left\{1t \mapsto \begin{bmatrix} det : + \\ num : 0x \end{bmatrix}\right\}\right)$
      - $\left(\text{gris}\left(N_A \begin{bmatrix} top : \begin{bmatrix} det : + \\ num : 0x \end{bmatrix} \\ bot : \begin{bmatrix} det : - \\ num : 11x \end{bmatrix} \end{bmatrix}\right), \left\{11t \mapsto \begin{bmatrix} det : + \\ num : 0x \end{bmatrix}\right\}\right)$
        - $\left(\text{le}\left(N_A \begin{bmatrix} top : \begin{bmatrix} det : + \\ num : 0x \end{bmatrix} \\ bot : \begin{bmatrix} det : + \\ num : sg \end{bmatrix} \end{bmatrix}\right), \left\{111t \mapsto \begin{bmatrix} det : + \\ num : 0x \end{bmatrix}, 11x \mapsto sg\right\}\right)$
          - $\left(\varepsilon, \left\{1111x \mapsto \begin{bmatrix} det : + \\ num : sg \end{bmatrix}, 0x \mapsto sg\right\}\right)$
    - $\left(\text{est}\left(V_A \begin{bmatrix} top : \left[num : sg\right] \\ bot : \left[num : sg\right] \end{bmatrix}\right), \left\{2t \mapsto \left[num : sg\right]\right\}\right)$
      - $\left(\varepsilon, \left\{21x \mapsto \left[num : sg\right]\right\}\right)$

Example

- $((S, \top), id)$
  - $\left(\mathrm{creux}\left(N\begin{bmatrix} top : \begin{bmatrix} det : + \\ num : 0x \end{bmatrix}\end{bmatrix}, V\begin{bmatrix} top : \begin{bmatrix} num : 0x \end{bmatrix}\end{bmatrix}\right), id\right)$
    - $\left(\mathrm{radis}\left(N_A\begin{bmatrix} top : \begin{bmatrix} det : + \\ num : 0x \end{bmatrix} \\ bot : \begin{bmatrix} det : - \end{bmatrix}\end{bmatrix}\right), \left\{1t \mapsto \begin{bmatrix} det : + \\ num : 0x \end{bmatrix}\right\}\right)$
      - $\left(\mathrm{gris}\left(N_A\begin{bmatrix} top : \begin{bmatrix} det : + \\ num : 0x \end{bmatrix} \\ bot : \begin{bmatrix} det : - \\ num : 11x \end{bmatrix}\end{bmatrix}\right), \left\{11t \mapsto \begin{bmatrix} det : + \\ num : 0x \end{bmatrix}\right\}\right)$
        - $\left(\mathrm{le}\left(N_A\begin{bmatrix} top : \begin{bmatrix} det : + \\ num : 0x \end{bmatrix} \\ bot : \begin{bmatrix} det : + \\ num : sg \end{bmatrix}\end{bmatrix}\right), \left\{111t \mapsto \begin{bmatrix} det : + \\ num : 0x \end{bmatrix}, 11x \mapsto sg\right\}\right)$
          - $\left(\varepsilon, \left\{1111x \mapsto \begin{bmatrix} det : + \\ num : sg \end{bmatrix}, 0x \mapsto sg\right\}\right)$
    - $\left(\mathrm{est}\left(V_A\begin{bmatrix} top : \begin{bmatrix} num : sg \end{bmatrix} \\ bot : \begin{bmatrix} num : sg \end{bmatrix}\end{bmatrix}\right), \left\{2t \mapsto \begin{bmatrix} num : sg \end{bmatrix}\right\}\right)$
      - $\left(\varepsilon, \left\{21x \mapsto \begin{bmatrix} num : sg \end{bmatrix}\right\}\right)$

Example

$$((S,\top),id)$$

$$\left(\mathrm{creux}\left(N\left[\mathrm{top}:\begin{bmatrix}det:+\\num:0x\end{bmatrix}\right],V\left[\mathrm{top}:\left[num:0x\right]\right]\right),id\right)$$

$$\left(\mathrm{radis}\left(N_A\begin{bmatrix}top:\begin{bmatrix}det:+\\num:0x\end{bmatrix}\\bot:\left[det:-\right]\end{bmatrix}\right),\left\{1t\mapsto\begin{bmatrix}det:+\\num:0x\end{bmatrix}\right\}\right)$$

$$\left(\mathrm{est}\left(V_A\begin{bmatrix}top:\left[num:sg\right]\\bot:\left[num:sg\right]\end{bmatrix}\right),\left\{2t\mapsto\left[num:sg\right]\right\}\right)$$

$$\left(\mathrm{gris}\left(N_A\begin{bmatrix}top:\begin{bmatrix}det:+\\num:0x\end{bmatrix}\\bot:\begin{bmatrix}det:-\\num:11x\end{bmatrix}\end{bmatrix}\right),\left\{11t\mapsto\begin{bmatrix}det:+\\num:0x\end{bmatrix}\right\}\right)$$

$$\left(\varepsilon,\left\{21x\mapsto\left[num:sg\right]\right\}\right)$$

$$\left(\mathrm{le}\left(N_A\begin{bmatrix}top:\begin{bmatrix}det:+\\num:0x\end{bmatrix}\\bot:\begin{bmatrix}det:+\\num:sg\end{bmatrix}\end{bmatrix}\right),\left\{111t\mapsto\begin{bmatrix}det:+\\num:0x\end{bmatrix},11x\mapsto sg\right\}\right)$$

$$\left(\varepsilon,\left\{1111x\mapsto\begin{bmatrix}det:+\\num:sg\end{bmatrix},0x\mapsto sg\right\}\right)$$

Example

$$((S, \top), \mathit{id})$$

$$\left(\mathrm{creux}\left(N \begin{bmatrix} \mathrm{top} : \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : 0x \end{bmatrix} \end{bmatrix}, V \begin{bmatrix} \mathrm{top} : \begin{bmatrix} \mathit{num} : 0x \end{bmatrix} \end{bmatrix}\right), \mathit{id}\right)$$

$$\left(\mathrm{radis}\left(N_A \begin{bmatrix} \mathit{top} : \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : 0x \end{bmatrix} \\ \mathit{bot} : \begin{bmatrix} \mathit{det} : - \end{bmatrix} \end{bmatrix}\right), \left\{1t \mapsto \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : 0x \end{bmatrix}\right\}\right)$$

$$\left(\mathrm{est}\left(V_A \begin{bmatrix} \mathit{top} : \begin{bmatrix} \mathit{num} : \mathrm{sg} \end{bmatrix} \\ \mathit{bot} : \begin{bmatrix} \mathit{num} : \mathrm{sg} \end{bmatrix} \end{bmatrix}\right), \left\{2t \mapsto \begin{bmatrix} \mathit{num} : \mathrm{sg} \end{bmatrix}\right\}\right)$$

$$\left(\mathrm{gris}\left(N_A \begin{bmatrix} \mathit{top} : \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : 0x \end{bmatrix} \\ \mathit{bot} : \begin{bmatrix} \mathit{det} : - \\ \mathit{num} : 11x \end{bmatrix} \end{bmatrix}\right), \left\{11t \mapsto \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : 0x \end{bmatrix}\right\}\right)$$

$$\left(\varepsilon, \left\{21x \mapsto \begin{bmatrix} \mathrm{num} : \mathrm{sg} \end{bmatrix}\right\}\right)$$

$$\left(\mathrm{le}\left(N_A \begin{bmatrix} \mathit{top} : \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : 0x \end{bmatrix} \\ \mathit{bot} : \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : \mathrm{sg} \end{bmatrix} \end{bmatrix}\right), \left\{111t \mapsto \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : 0x \end{bmatrix}, 11x \mapsto \mathrm{sg}\right\}\right)$$

$$\left(\varepsilon, \left\{1111x \mapsto \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : \mathrm{sg} \end{bmatrix}, 0x \mapsto \mathrm{sg}\right\}\right)$$

Example

$$((S,\top),\mathit{id})$$

$$\left(\mathrm{creux}\left(N\begin{bmatrix}\mathrm{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\end{bmatrix}, V\begin{bmatrix}\mathrm{top}:\begin{bmatrix}\mathit{num}:0x\end{bmatrix}\end{bmatrix}\right),\mathit{id}\right)$$

$$\left(\mathrm{radis}\left(N_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{det}:-\end{bmatrix}\end{bmatrix}\right),\left\{1t\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\right\}\right)$$

$$\left(\mathrm{gris}\left(N_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{det}:-\\ \mathit{num}:11x\end{bmatrix}\end{bmatrix}\right),\left\{11t\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\right\}\right)$$

$$\left(\mathrm{le}\left(N_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:\mathrm{sg}\end{bmatrix}\end{bmatrix}\right),\left\{111t\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:0x\end{bmatrix}, 11x\mapsto\mathrm{sg}\right\}\right)$$

$$\left(\varepsilon,\left\{1111x\mapsto\begin{bmatrix}\mathit{det}:+\\ \mathit{num}:\mathrm{sg}\end{bmatrix}, 0x\mapsto\mathrm{sg}\right\}\right)$$

$$\left(\mathrm{est}\left(V_A\begin{bmatrix}\mathit{top}:\begin{bmatrix}\mathit{num}:\mathrm{sg}\end{bmatrix}\\ \mathit{bot}:\begin{bmatrix}\mathit{num}:\mathrm{sg}\end{bmatrix}\end{bmatrix}\right),\left\{2t\mapsto\begin{bmatrix}\mathit{num}:\mathrm{sg}\end{bmatrix}\right\}\right)$$

$$\left(\varepsilon,\left\{21x\mapsto\begin{bmatrix}\mathrm{num}:\mathrm{sg}\end{bmatrix}\right\}\right)$$

Left Corner Transform

Rosenkrantz and Lewis II [1970]

- top feature structures of root nodes usually empty
- reverse order of root adjunctions on initial trees

Derivation Tree Language (3)

$$(S, \top) \longrightarrow \text{creux}\left(N_I \left[\mathit{top} : \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : x \end{bmatrix} \right], V_I \left[\mathit{top} : \left[\mathit{num} : x \right] \right] \right)$$

$$N \left[\mathit{bot} : \left[\mathit{det} : - \right] \right] \longrightarrow \text{radis}$$

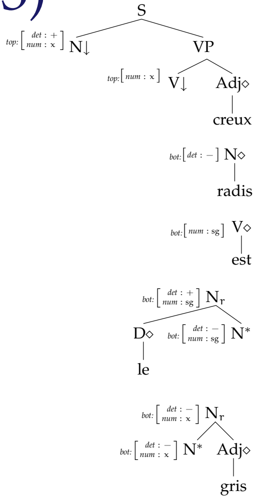
$$V \left[\mathit{bot} : \left[\mathit{num} : \mathrm{sg} \right] \right] \longrightarrow \text{est}$$

$$N \begin{bmatrix} \mathit{top} : t \\ \mathit{bot} : \begin{bmatrix} \mathit{det} : + \\ \mathit{num} : \mathrm{sg} \end{bmatrix} \end{bmatrix} \longrightarrow \text{le}\left(N \begin{bmatrix} \mathit{top} : t \\ \mathit{bot} : \begin{bmatrix} \mathit{det} : - \\ \mathit{num} : \mathrm{sg} \end{bmatrix} \end{bmatrix} \right)$$

$$N \begin{bmatrix} \mathit{top} : t \\ \mathit{bot} : \begin{bmatrix} \mathit{det} : - \\ \mathit{num} : x \end{bmatrix} \end{bmatrix} \longrightarrow \text{gris}\left(N \begin{bmatrix} \mathit{top} : t \\ \mathit{bot} : \begin{bmatrix} \mathit{det} : - \\ \mathit{num} : x \end{bmatrix} \end{bmatrix} \right)$$

$$N_I \left[\mathit{top} : t \right] \longrightarrow \varepsilon \left(N \begin{bmatrix} \mathit{top} : t \\ \mathit{bot} : t \end{bmatrix} \right)$$

$$V_I \left[\mathit{top} : t \right] \longrightarrow \varepsilon \left(V \begin{bmatrix} \mathit{top} : t \\ \mathit{bot} : t \end{bmatrix} \right)$$



Overgeneration

Definition

Equivalently, if a grammar

- assigns incorrect structure to grammatical sentences
- accepts agrammatical sentences
- generates agrammatical sentences

precision score in "precision and recall" metrics.

Error Mining

van Noord [2004], Sagot and Éric de la Clergerie [2006]

Applied to grammatical coverage (“recall”)

1. parse a large corpus of correct sentences
2. failures indicate coverage issues
3. statistical analysis identifies a probable culprit for each failure
4. might attempt to provide corrections

Error Mining

van Noord [2004], Sagot and Éric de la Clergerie [2006]

Applied to grammatical coverage (“recall”)

1. parse a large corpus of correct sentences
2. failures indicate coverage issues
3. statistical analysis identifies a probable culprit for each failure
4. might attempt to provide corrections

For Overgeneration?

Which test suite for pass/failure?

- a TreeBank
- a corpus of incorrect sentences
- sentences generated from the grammar
    - which input?
    - generation from logic formulae is computationally expensive

For Overgeneration?

Which test suite for pass/failure?

- a TreeBank
- a corpus of incorrect sentences
- sentences generated from the grammar
  - which input?
  - generation from logic formulae is computationally expensive

For Overgeneration?

Which test suite for pass/failure?

- a TreeBank
- a corpus of incorrect sentences
- sentences generated from the grammar
  - which input?
  - generation from logic formulae is computationally expensive

For Overgeneration?

Which test suite for pass/failure?

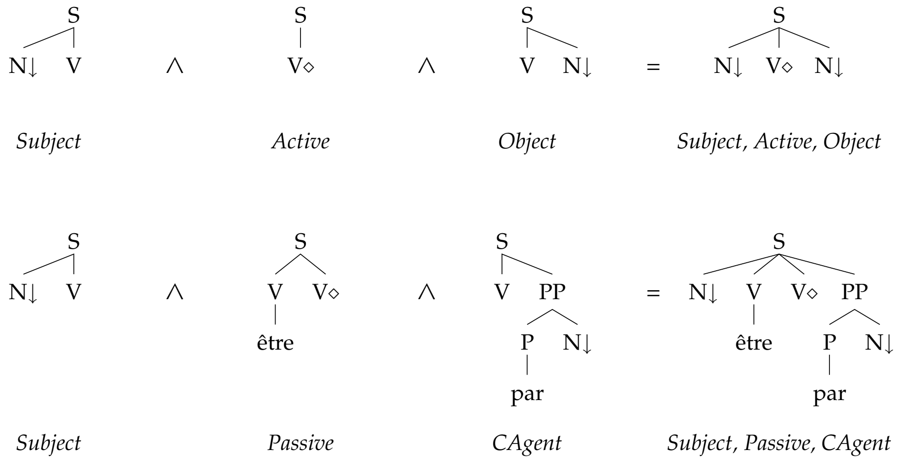
- a TreeBank
- a corpus of incorrect sentences
- sentences generated from the grammar
  - which input?
  - generation from logic formulae is computationally expensive

“Exhaustive” Generation

- not in terms of elementary trees (about 6,000)
- in terms of *linguistic phenomena*
  - grammar compiled from a meta grammar
  - compilation traces
  - 87 classes match linguistic phenomena

Meta Grammar

[Crabbé, 2005]



Guided Generation

- use distances from elementary trees to classes
- generate derivation trees that subsume a target multiset of classes
- avoid the already generated classes
- ... in progress

Guided Generation

- use distances from elementary trees to classes
- generate derivation trees that subsume a target multiset of classes
- avoid the already generated classes
- ... in progress

Concluding Remarks

- derivation trees capture syntactic dependencies
- appropriate intermediate level between syntax and semantics
- not just for TAG

References

W. S. Brainerd. Tree generating regular systems. *Information and Control*, 14 (2):217–231, 1969. doi: 10.1016/S0019-9958(69)90065-5.

J. Carroll, A. Copestake, D. Flickinger, and V. Poznanski. An efficient chart generator for (semi-)lexicalist grammars. In *ENLG'99*, pages 86–95, 1999. URL `http://www.cogs.susx.ac.uk/lab/nlp/carroll/papers/ewnlg99.pdf`.

B. Crabbé. Grammatical development with XMG. In P. Blache, E. Stabler, J. Busquets, and R. Moot, editors, *LACL'05*, volume 3492 of *Lecture Notes in Computer Science*, pages 84–100. Springer, 2005. ISBN 978-3-540-25783-7. doi: 10.1007/11422532_6.

J. Engelfriet and H. Vogler. Macro tree transducers. *J. Comput. Syst. Sci.*, 31 (1):71–146, 1985. doi: 10.1016/0022-0000(85)90066-2.

C. Gardent and L. Kallmeyer. Semantic construction in feature-based TAG. In *EACL'03*, pages 123–130. ACL Press, 2003. ISBN 1-333-56789-0. doi: 10.3115/1067807.1067825.

C. Gardent and E. Kow. A symbolic approach to near-deterministic surface realisation using Tree Adjoining Grammar. In *ACL'07*, pages 328–335. ACL Press, 2007. URL `http://www.aclweb.org/anthology/P07-1042`.

A. K. Joshi and Y. Schabes. Tree-adjoining grammars. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3: Beyond Words, chapter 2, pages 69–124. Springer, 1997. ISBN 3-540-60649-1. URL `http://citeseer.ist.psu.edu/joshi97treeadjoining.html`.

A. K. Joshi, L. S. Levy, and M. Takahashi. Tree adjunct grammars. *J. Comput. Syst. Sci.*, 10(1):136–163, 1975. doi: 10.1016/S0022-0000(75)80019-5.

S. Kepster and J. Rogers. The equivalence of tree adjoining grammars and monadic linear context-free tree grammars. In M. Kracht, editor, *MoL 10*, 2007. URL `http://kracht.humnet.ucla.edu/marcus/mol10/abstracts/kepser-rogers-mol.pdf`.

A. Koller and M. Stone. Sentence generation as a planning problem. In *ACL'07*, pages 336–343. ACL Press, 2007. URL `http://www.aclweb.org/anthology/P07-1043`.

A. Koller and K. Striegnitz. Generation as dependency parsing. In *ACL'02*, pages 17–24. ACL Press, 2002. doi: 10.3115/1073083.1073088.

S. Pogodalla. Computing semantic representation: Towards ACG abstract terms as derivation trees. In O. Rambow and M. Stone, editors, *TAG+7*, pages 64–71, 2004. URL `http://www.cs.rutgers.edu/TAG+7/papers/pogodalla.pdf`.

D. J. Rosenkrantz and P. M. Lewis II. Deterministic left corner parsing. In *11th Annual Symposium on Switching and Automata Theory*, pages 139–152. IEEE Computer Society, 1970.

B. Sagot and Éric de la Clergerie. Error mining in parsing results. In *ACL'06*, pages 329–336. ACL Press, 2006. doi: 10.3115/1220175.1220217. URL `http://www.aclweb.org/anthology/P06-1042`.

S. Schmitz and J. Le Roux. Feature unification in TAG derivation trees. In C. Gardent and A. Sarkar, editors, *TAG+9*, 2008. URL `http://www.loria.fr/~schmitsy/pub/tagunif.en.pdf`.

S. M. Shieber. Unifying synchronous tree-adjoining grammars and tree transducers via bimorphisms. In *EACL'06*. ACL Press, 2006. ISBN 1-932432-59-0. URL `http://www.aclweb.org/anthology/E06-1048`.

G. van Noord. Error mining for wide-coverage grammar engineering. In *ACL'04*, pages 446–453. ACL Press, 2004. doi: 10.3115/1218955.1219012. URL `http://www.aclweb.org/anthology/P04-1057`.

K. Vijay-Shanker and A. K. Joshi. Feature structures based tree adjoining grammars. In *COLING'88*, pages 714–719. ACL Press, 1988. ISBN 963-8431-56-3. doi: 10.3115/991719.991783.