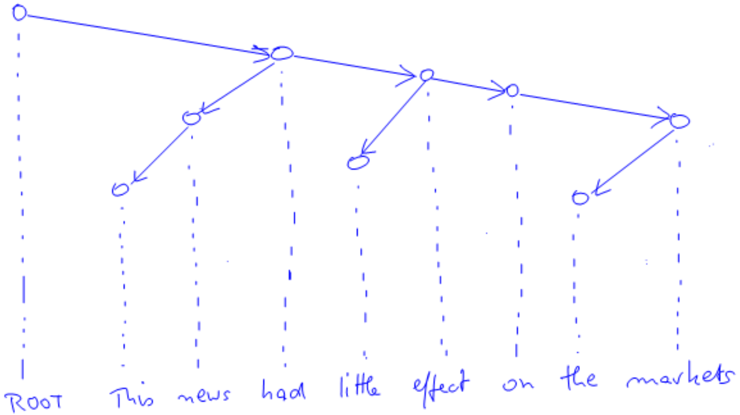
Dynamic programming algorithms for transition-based dependency parsing

M. Kuhlmann, C. Gomez-Rodriguez, G. Satta

Proc. of 49th Annual Meeting of ACL, 2011

by Heiko Vogler

dependency structure (projective) [Tesnière 55]



$w = w_0 \ldots w_{n-1}$ w_i word

$V_w = \{0, \ldots, n-1\}$

transition-based dependency parsing [Nivre 2008, 2005]
(arc-standard)

transition system $(C, \vdash)$

- C set (configurations)

  $c = (\sigma, \beta, A)$

  A: partial dependency graph over V_w

  buffer $\beta \in V_w^*$, top: left

  stack $\sigma \in V_w^*$, top: right

  given $c \in C$:

  - $\sigma(c)$: stack of c
  - $\beta(c)$: buffer of c

- $\vdash \subseteq C \times C$ (computation relation) transition system for transition-based dep. parsing

for every $\sigma, \beta \in V_w^*$ and $i, j \in V_w$:

$(\sigma, i\beta, A) \vdash (\sigma i, \beta, A)$ (shift)

$(\sigma i j, \beta, A) \vdash (\sigma j, \beta, A \cup \{j \to i\})$ (left-arc)

$(\sigma i j, \beta, A) \vdash (\sigma i, \beta, A \cup \{i \to j\})$ (right-arc)

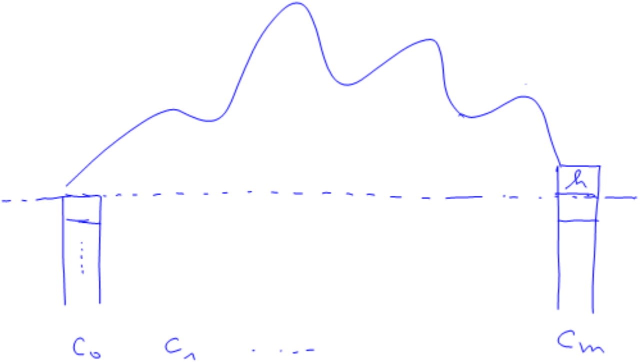
initial configuration: $(\varepsilon, 0 \ldots n-1, \emptyset)$

final configuration: $(0, \varepsilon, \underbrace{A}_{\text{result of dep. parsing}})$

push computation : $\gamma = c_0 \dots c_m$ $m \geq 1$

(P1) $(\forall i : 1 \leq i < m)(\exists \sigma_i \in V_w^*) : \sigma(c_i) = \sigma(c_0)\,\sigma_i$

(P2) $(\exists h \in V_w) : \sigma(c_m) = \sigma(c_0)\,h$.



deduction system: tabulates the computations of the arc-standard parsing for a given input word $w = w_0 \dots w_{n-1}$

(abbreviation: $i\,(i+1) \dots (n-1) \rightsquigarrow \beta_i$)

item: $[i, h, j]$ $\quad 0 \leq i \leq h < j \leq n$

$\forall c_0 \in C$ with $\beta(c_0) = \beta_i$ $\exists$ push computation $\gamma = c_0 \dots c_m$:

$\beta(c_m) = \beta_j$ and $\sigma(c_m) = \sigma(c_0)\,h$

(illustration:

$$(\sigma,\ i\,(i+1) \dots (n-1),\ A) \vdash^* (\sigma h,\ j \dots (n-1),\ A')$$

)

item: $[i, h, j]$ $\quad 0 \le i \le h < j \le n$

$\forall c_0 \in C$ with $\beta(c_0) = \beta_i$ $\exists$ push computation $\gamma = c_0 \dots c_m$:
$\beta(c_m) = \beta_j$ and $\sigma(c_m) = \sigma(c_0)\, h$

goal: $[0, 0, n]$

axioms: $[i, i, i+1]$ $\quad (\forall i : 0 \le i < n)$

rules:

$$\frac{[i, h_1, k]\,[k, h_2, j]}{[i, h_1, j]}$$

$(la : h_1 \to h_2)$

$$\left(\begin{array}{c} \vdots \\ \hline \end{array}, \; i \dots k \dots j \dots (n-1), \; A \right)$$

$$\vdash^{*} \left(\begin{array}{c} \boxed{h_1} \\ \vdots \\ \hline \end{array}, \; k \dots j \dots (n-1), \; A' \right)$$

$$\vdash^{*} \left(\begin{array}{c} \boxed{h_2} \\ \boxed{h_1} \\ \vdots \\ \hline \end{array}, \; j \dots (n-1), \; A'' \right)$$

$$\vdash \left(\begin{array}{c} \boxed{h_1} \\ \vdots \\ \hline \end{array}, \; j \dots (n-1), \; A'' \cup \{h_1 \to h_2\} \right)$$

item: $[i, h, j]$ $\quad 0 \leq i \leq h < j \leq n$

$\forall c_0 \in C$ with $\beta(c_0) = \beta_i$ $\exists$ push computation $\gamma = c_0 \dots c_m$:

$\beta(c_m) = \beta_j$ and $\sigma(c_m) = \sigma(c_0)\, h$

goal: $[0, 0, n]$

axioms: $[i, i, i+1]$ $\quad (\forall i: \; 0 \leq i < n)$

rules:

$$\frac{[i, h_1, k][k, h_2, j]}{[i, \cancel{h_1}\, h_2, j]}$$

$(\cancel{la: h_1 \to h_2})$

$ra: h_2 \to h_1$

$$(\boxed{\vdots}\,,\; i \dots k \dots j \dots (n-1),\; A)$$

$$\vdash^* \; (\boxed{h_1}\,\boxed{\vdots}\,,\; k \dots j \dots (n-1),\; A')$$

$$\vdash^* \; (\boxed{h_2}\,\boxed{h_1}\,\boxed{\vdots}\,,\; j \dots (n-1),\; A'')$$

$$\vdash \; (\boxed{\cancel{h_1}}\; h_2\,\boxed{\vdots}\,,\; j \dots (n-1),\; A'' \cup \{\cancel{h_1 \to h_2}\} \; h_2 \to h_1)$$

item: $[i, h, j]$ $\quad 0 \le i \le h < j \le n$

$\forall c_0 \in C$ with $\beta(c_0) = \beta_i$ $\;\exists$ push computation $\gamma = c_0 \dots c_m$:
$\beta(c_m) = \beta_j$ and $\sigma(c_m) = \sigma(c_0)\, h$

goal: $[0, 0, n]$

axioms: $[i, i, i+1]$ $\quad (\forall i:\ 0 \le i < n)$

rules:

$$\frac{[i, h_1, k]\,[k, h_2, j]}{[i, h_1, j]}$$

$(la: h_1 \to h_2)$

$$\left(\begin{array}{c}\vdots\end{array},\ i \dots h \dots j \dots (n-1),\ A\right)$$

$$\vdash^{*} \left(\begin{array}{c}\boxed{h_1}\\ \vdots\end{array},\ k \dots j \dots (n-1),\ A'\right)$$

$$\vdash^{*} \left(\begin{array}{c}\boxed{h_2}\\ \boxed{h_1}\\ \vdots\end{array},\ j \dots (n-1),\ A''\right)$$

$$\vdash \left(\begin{array}{c}\boxed{h_1}\\ \vdots\end{array},\ j \dots (n-1),\ A'' \cup \{h_1 \to h_2\}\right)$$

Space: $\mathcal{O}(|w|^3)$

time: $\mathcal{O}(|w|^5)$