Parsing Methods as Deductive Systems

from [Shieber, Schabes, and Pereira. 1993]

Algorithm = Logic + Control

Robert Kowalski

Basic Notions

Inference Rule: $\frac{A_1,\ldots,A_k}{B}$ $\quad < $ ***side conditions on*** $A_1, \ldots, A_k, B >$

Derivation of a formula B from assumption $A_1, \ldots, A_m$:

a sequence of formulas $S_1, \ldots, S_n$ such that

$B = S_n$, and

for each S_i

either S_i is one of the A_j
[or S_i is an instance of an axiom A_j]

or there is a rule of inference R and the formulas $S_{i_1}, \ldots, S_{i_k}$

with $i_1, \ldots, i_k < i$ such that
[for appropriate substitutions of terms for the meta-variables in R]

$S_{i_1}, \ldots, S_{i_k}$ match the antecedents of the rule,
S_i matches the consequent,
and the rule's side conditions are satisfied.

Notation: **$A_1, \ldots, A_m \models B$.**

The CYK deductive parsing system

- **Item form:** $[A, i, j]$
- **Axioms:** $[A, i, i+1]$ **if** $A \longrightarrow w_i$
- **Goals:** $[S, 0, n]$
- **Inference rules:** $\frac{[B,i,j] \quad [C,j,k]}{[A,i,k]}$ **if** $A \longrightarrow B\ C$

An example CFG

$$
\begin{aligned}
\mathbf{S} &\rightarrow \mathbf{NP}\ \mathbf{VP} \\
\mathbf{NP} &\rightarrow \mathbf{Det}\ \mathbf{N}\ \mathbf{OptRel} \\
\mathbf{NP} &\rightarrow \mathbf{PN} \\
\mathbf{VP} &\rightarrow \mathbf{TV}\ \mathbf{NP} \\
\mathbf{VP} &\rightarrow \mathbf{IV} \\
\mathbf{OptRel} &\rightarrow \mathbf{RelPro}\ \mathbf{VP} \\
\mathbf{OptRel} &\rightarrow \epsilon
\end{aligned}
\qquad
\begin{aligned}
\mathbf{Det} &\rightarrow \text{a} \\
\mathbf{N} &\rightarrow \text{lindy} \\
\mathbf{PN} &\rightarrow \text{Trip} \\
\mathbf{IV} &\rightarrow \text{swings} \\
\mathbf{TV} &\rightarrow \text{dances} \\
\mathbf{RelPro} &\rightarrow \text{that}
\end{aligned}
$$

Pure Bottom-up (Shift-Reduce) Parsing

- **Item form:** $[\alpha\bullet, j]$
- **Axioms:** $[\bullet, 0]$
- **Goals:** $[S\bullet, n]$
- **Inference rules:**

  Shift: $\frac{[\alpha\bullet,j]}{[\alpha w_{j+1}\bullet,j+1]}$

  Reduce: $\frac{[\alpha\gamma\bullet,j]}{[\alpha B\bullet,j]}$ **if** $B \rightarrow \gamma$
- **Invariant:** $[\alpha\bullet, j]$**:** $\alpha w_{j+1} \ldots w_n \Rightarrow^{\star} w_1 \ldots w_n$.

Example: bottom-up parsing

a lindy swings

1	$[\bullet, 0]$	AXIOM
2	$[\textbf{a}\bullet, 1]$	SHIFT **from 1**
3	$[\boldsymbol{Det}\bullet, 1]$	REDUCE **from 2**
4	$[\boldsymbol{Det}\ \textbf{lindy}\bullet, 2]$	SHIFT **from 3**
5	$[\boldsymbol{Det}\ \boldsymbol{N}\bullet, 2]$	REDUCE **from 4**
6	$[\boldsymbol{Det}\ \boldsymbol{N}\ \boldsymbol{OptRel}\bullet, 2]$	REDUCE **from 5**
7	$[\boldsymbol{NP}\bullet, 2]$	REDUCE **from 6**
8	$[\boldsymbol{NP}\ \textbf{swings}\bullet, 3]$	SHIFT **from 7**
9	$[\boldsymbol{NP}\ \boldsymbol{IV}\bullet, 3]$	REDUCE **from 8**
10	$[\boldsymbol{NP}\ \boldsymbol{VP}\bullet, 3]$	REDUCE **from 9**
11	$[\boldsymbol{S}\bullet, 3]$	REDUCE **from 10**

Pure Top-down (Recursive Descent) Parsing

- **Item form:** $[\bullet\beta, j]$
- **Axioms:** $[\bullet S, 0]$
- **Goals:** $[\bullet, n]$
- **Inference rules:**

  Scanning: $\frac{[\bullet w_{j+1}\beta, j]}{[\bullet\beta, j+1]}$

  Prediction: $\frac{[\bullet B\beta, j]}{[\bullet\gamma\beta, j]}$ **if** $B \rightarrow \gamma$
- **Invariant:** $[\bullet\beta, j]$**:** $S \Rightarrow^\star w_1 \ldots w_j\beta$**.**

Example: top-down parsing

a lindy swings

1	$[\bullet \boldsymbol{S}, 0]$	AXIOM
2	$[\bullet \boldsymbol{NP}\ \boldsymbol{VP}, 0]$	PREDICT **from 1**
3	$[\bullet \boldsymbol{Det}\ \boldsymbol{N}\ \boldsymbol{OptRel}\ \boldsymbol{VP}, 0]$	PREDICT **from 2**
4	$[\bullet \textbf{a}\ \boldsymbol{N}\ \boldsymbol{OptRel}\ \boldsymbol{VP}, 0]$	PREDICT **from 3**
5	$[\bullet \boldsymbol{N}\ \boldsymbol{OptRel}\ \boldsymbol{VP}, 1]$	SCAN **from 4**
6	$[\bullet \textbf{lindy}\ \boldsymbol{OptRel}\ \boldsymbol{VP}, 1]$	PREDICT **from 5**
7	$[\bullet \boldsymbol{OptRel}\ \boldsymbol{VP}, 2]$	SCAN **from 6**
8	$[\bullet \boldsymbol{VP}, 2]$	PREDICT **from 7**
9	$[\bullet \boldsymbol{IV}, 2]$	PREDICT **from 8**
10	$[\bullet \textbf{swings}, 2]$	PREDICT **from 9**
11	$[\bullet, 3]$	SCAN **from 10**

Earley Parsing [1970] as Deduction

- **Item form:** $[i, A \rightarrow \alpha \bullet \beta, j]$
- **Axioms:** $[0, S' \rightarrow \bullet S, 0]$ **(where** S' **is a new nonterminal)**
- **Goals:** $[0, S' \rightarrow S \bullet, n]$
- **Inference rules:**

  Scanning: $\frac{[i, A \rightarrow \alpha \bullet w_{j+1} \beta, j]}{[i, A \rightarrow \alpha w_{j+1} \bullet \beta, j+1]}$

  Prediction: $\frac{[i, A \rightarrow \alpha \bullet B \beta, j]}{[j, B \rightarrow \bullet \gamma, j]}$ **if** $B \rightarrow \gamma$

  Completion: $\frac{[i, A \rightarrow \alpha \bullet B \beta, k] [k, B \rightarrow \gamma \bullet, j]}{[i, A \rightarrow \alpha B \bullet \beta, j]}$

- **Invariant:** $[i, A \rightarrow \alpha \bullet \beta, j]$**:**

  $S \Rightarrow^\star w_1 \ldots w_i A \gamma$**,**

  $\alpha w_{j+1} \ldots w_n \Rightarrow^\star w_{i+1} \ldots w_n$**.**

Example: Earley parsing

a lindy swings

1	$[0, S' \rightarrow S\bullet, 0]$	AXIOM
2	$[0, S' \rightarrow \bullet \boldsymbol{NP}\ \boldsymbol{VP}, 0]$	PREDICT **from 1**
3	$[0, \boldsymbol{NP} \rightarrow \bullet \boldsymbol{Det}\ \boldsymbol{N}\ \boldsymbol{OptRel}, 0]$	PREDICT **from 2**
4	$[0, \boldsymbol{Det} \rightarrow \bullet\mathbf{a}, 0]$	PREDICT **from 3**
5	$[0, \boldsymbol{Det} \rightarrow \mathbf{a}\bullet, 1]$	SCAN **from 4**
6	$[0, \boldsymbol{NP} \rightarrow \boldsymbol{Det} \bullet \boldsymbol{N}\ \boldsymbol{OptRel}, 1]$	COMPLETE **from 3 and 5**
7	$[1, N \rightarrow \bullet\mathbf{lindy}, 1]$	PREDICT **from 6**
8	$[1, N \rightarrow \mathbf{lindy}\bullet, 2]$	SCAN **from 7**
9	$[0, \boldsymbol{NP} \rightarrow \boldsymbol{Det}\ \boldsymbol{N} \bullet \boldsymbol{OptRel}, 2]$	COMPLETE **from 6 and 8**

Example: Earley parsing (cont'd)

10	$[2, \boldsymbol{OptRel} \rightarrow \bullet, 2]$	PREDICT **from 9**
11	$[0, \boldsymbol{NP} \rightarrow \boldsymbol{Det}\ \boldsymbol{N}\ \boldsymbol{OptRel}\bullet, 2]$	COMPLETE **from 9 and 10**
12	$[0, \boldsymbol{S} \rightarrow \boldsymbol{NP} \bullet \boldsymbol{VP}, 2]$	COMPLETE **from 2 and 11**
13	$[2, \boldsymbol{VP} \rightarrow \bullet\boldsymbol{IV}, 2]$	PREDICT **from 12**
14	$[2, \boldsymbol{IV} \rightarrow \bullet\textbf{swings}, 2]$	PREDICT **from 13**
15	$[2, \boldsymbol{IV} \rightarrow \textbf{swings}\bullet, 3]$	SCAN **from 14**
16	$[2, \boldsymbol{VP} \rightarrow \boldsymbol{IV}\bullet, 3]$	COMPLETE **from 13 and 15**
17	$[0, S \rightarrow \boldsymbol{NP}\ \boldsymbol{VP}\bullet, 3]$	COMPLETE **from 12 and 16**
18	$[0, S' \rightarrow S\bullet, 3]$	COMPLETE **from 1 and 17**

Head-Corner Parsing

Sample Grammar:

$$S \rightarrow \boldsymbol{NP} \; *\boldsymbol{VP}$$
$$\boldsymbol{VP} \rightarrow *v \; \boldsymbol{NP}$$
$$\boldsymbol{NP} \rightarrow \boldsymbol{det} \; *\boldsymbol{noun}$$

Items:

$[l, r, A]$: **predict items, or goals**

$[l, r, A;\ B \rightarrow \alpha.\beta.\gamma, i, j]$: **head-corner (HC) items**

$[a, j-1, j]$: **terminal items**

Head-corner relation:

$>_h$ **on** $N \times (V \cup \{\epsilon\})$ **defined by:**

$A >_h U$ **if there is** $p = A \rightarrow \alpha \in P$ **with** U **the head of** p**.**

$>_h^*$ **is the reflexive and transitive closure of** $>_h$**.**

Deductive Head-Corner Parsing

$$
\begin{aligned}
D^{\textit{Init}} &= \{[\$, n, n+1] \vdash [0, n, S]\} \\
D^{\textit{HC(a)}} &= \{[l, r, A], [b, j-1, j] \vdash [l, r, A;\ B \rightarrow \alpha.b.\gamma, j-1, j]\} \\
D^{\textit{HC(A)}} &= \{[l, r, A;\ C \rightarrow .\delta., i, j] \vdash [l, r, A;\ B \rightarrow \alpha.C.\gamma, i, j]\} \\
D^{\textit{HC}(\epsilon)} &= \{[l, r, A] \vdash [l, r, A;\ B \rightarrow .., j, j]\} \\
D^{\textit{lPred}} &= \{[l, r, A;\ B \rightarrow \alpha C.\beta.\gamma, i, j] \vdash [l, i, C]\} \\
D^{\textit{rPred}} &= \{[l, r, A;\ B \rightarrow \alpha.\beta.C\gamma, i, j] \vdash [j, r, C]\} \\
D^{\textit{lScan}} &= \{[a, j-1, j], [l, r, A;\ B \rightarrow \alpha a.\beta.\gamma, j, k] \vdash \\
&\qquad [l, r, A; B \rightarrow \alpha.a\beta.\gamma, j-1, k]\} \\
D^{\textit{rScan}} &= \{[l, r, A;\ B \rightarrow \alpha.\beta.a\gamma, i, j], [a, j, j+1] \vdash \\
&\qquad [l, r, A;\ B \rightarrow \alpha.\beta a.\gamma, i, j+1]\} \\
D^{\textit{lCompl}} &= \{[l, j, C;\ C \rightarrow .\delta., i, j], [l, r, A;\ B \rightarrow \alpha C.\beta.\gamma, j, k] \vdash \\
&\qquad [l, r, A;\ B \rightarrow \alpha.C\beta.\gamma, i, k]\} \\
D^{\textit{rCompl}} &= \{[l, r, A;\ B \rightarrow \alpha.\beta.C\gamma, i, j], [j, r, C;\ C \rightarrow .\delta., j, k] \vdash \\
&\qquad [l, r, A;\ B \rightarrow \alpha.\beta C.\gamma, i, k]\}
\end{aligned}
$$