Natural Language Processing

Lecture 13:
More on CFG Parsing
Probabilistic/Weighted Parsing
Example: ambiguous parse

```
S
| VP
  | NP
  | Verb
  | Book
  | Det
  | the
  | Nominal
  | Nominal
  | Noun
  | flight
  | Noun
  | dinner

S
| VP
  | NP
  | Verb
  | Book
  | Det
  | the
  | Nominal
  | Noun
  | dinner
  | Nominal
  | flight
```
# Probabilistic CFG

<table>
<thead>
<tr>
<th>Grammar</th>
<th>Probability</th>
<th>Lexicon</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S \rightarrow NP \ VP$</td>
<td>.80</td>
<td>$Det \rightarrow that [.10] \mid a [.30] \mid the [.60]$</td>
</tr>
<tr>
<td>$S \rightarrow Aux \ NP \ VP$</td>
<td>.15</td>
<td>$Noun \rightarrow book [.10] \mid flight [.30]$</td>
</tr>
<tr>
<td>$S \rightarrow VP$</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>$NP \rightarrow Pronoun$</td>
<td>.35</td>
<td>$Verb \rightarrow book [.30] \mid include [.30]$</td>
</tr>
<tr>
<td>$NP \rightarrow Proper-Noun$</td>
<td>.30</td>
<td></td>
</tr>
<tr>
<td>$NP \rightarrow Det Nominal$</td>
<td>.20</td>
<td>$Pronoun \rightarrow I [.40] \mid she [.05]$</td>
</tr>
<tr>
<td>$NP \rightarrow Nominal$</td>
<td>.15</td>
<td></td>
</tr>
<tr>
<td>$Nominal \rightarrow Noun$</td>
<td>.75</td>
<td>$Proper-Noun \rightarrow Houston [.60]$</td>
</tr>
<tr>
<td>$Nominal \rightarrow Nominal Noun$</td>
<td>.20</td>
<td></td>
</tr>
<tr>
<td>$Nominal \rightarrow Nominal PP$</td>
<td>.05</td>
<td>$Aux \rightarrow does [.60] \mid can [40]$</td>
</tr>
<tr>
<td>$VP \rightarrow Verb$</td>
<td>.35</td>
<td>$Preposition \rightarrow from [.30] \mid to [.30]$</td>
</tr>
<tr>
<td>$VP \rightarrow Verb NP$</td>
<td>.20</td>
<td></td>
</tr>
<tr>
<td>$VP \rightarrow Verb NP PP$</td>
<td>.10</td>
<td>$\quad \mid on [.20] \mid near [.15]$</td>
</tr>
<tr>
<td>$VP \rightarrow Verb PP$</td>
<td>.15</td>
<td>$\quad \mid through [.05]$</td>
</tr>
<tr>
<td>$VP \rightarrow VP PP$</td>
<td>.15</td>
<td></td>
</tr>
<tr>
<td>$PP \rightarrow Preposition NP$</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>
Ambiguous parse w/probabilities

P(left) = 2.2 \times 10^{-6} 

P(right) = 6.1 \times 10^{-7}
Review: Context-Free Grammars

• Vocabulary of terminal symbols, $\Sigma$
• Set of nonterminal symbols (a.k.a. variables), $N$
• Special start symbol $S \in N$
• Production rules of the form $X \rightarrow \alpha$

where

$X \in N$
$\alpha \in (N \cup \Sigma)^*$  \hspace{1cm} (in CNF: $\alpha \in N^2 \cup \Sigma$)
Probabilistic Context-Free Grammars

• Vocabulary of terminal symbols, $\Sigma$
• Set of nonterminal symbols (a.k.a. variables), $N$
• Special start symbol $S \in N$
• Production rules of the form $X \rightarrow \alpha$, each with a positive weight $p(X \rightarrow \alpha)$, where
  
  - $X \in N$
  - $\alpha \in (N \cup \Sigma)^*$ (in CNF: $\alpha \in N^2 \cup \Sigma$)
  - $\forall X \in N$, $\sum_\alpha p(X \rightarrow \alpha) = 1$
CKY Algorithm: Review

for $i = 1 \ldots n$

\[ C[i-1, i] = \{ V \mid V \rightarrow w_i \} \]

for $\ell = 2 \ldots n$ // width

for $i = 0 \ldots n - \ell$ // left boundary

$k = i + \ell$ // right boundary

for $j = i + 1 \ldots k - 1$ // midpoint

\[ C[i, k] = C[i, k] \cup \{ V \mid V \rightarrow YZ, Y \in C[i, j], Z \in C[j, k] \} \]

return true if $S \in C[0, n]$
Weighted CKY Algorithm

for $i = 1 \ldots n$, $V \in N$
\[ C[V, i-1, i] = p(V \rightarrow w_i) \]

for $\ell = 2 \ldots n$ // width of span
  for $i = 0 \ldots n - \ell$ // left boundary
    $k = i + \ell$ // right boundary
    for $j = i + 1 \ldots k - 1$ // midpoint
      for each binary rule $V \rightarrow YZ$
      \[ C[V, i, k] = \max \{ C[V, i, k], \]
      \[ C[Y, i, j] \times C[Z, j, k] \]
      \[ \times p(V \rightarrow YZ) \} \]

return true if $S \in C[\cdot, 0, n]$
CKY Algorithm: Review

\[ C[V, i - 1, i] = \begin{cases} 
  \text{TRUE} & \text{if } V \rightarrow w_i \\
  \text{FALSE} & \text{otherwise}
\end{cases} \]

\[ C[V, i, j] = \begin{cases} 
  \text{TRUE} & \text{if } \exists j, Y, Z \text{ such that } V \rightarrow YZ \\
  & \text{and } C[Y, i, k] \\
  & \text{and } C[Z, k, j] \\
  & \text{and } i < k < j \\
  \text{FALSE} & \text{otherwise}
\end{cases} \]

goal = C[S, 0, n]
Weighted CKY Algorithm

base case:
\[ C[X, i-1, i] = p(X \rightarrow w_i) \]

induction:
\[ C[X, i, k] = \max_{j, Y, Z} p(X \rightarrow Y \ Z) \times C[Y, i, j] \times C(Z, j, k) \]

goal:
\[ C[S, 0, n] \text{ where } n = |w| \]
\[ p(\tau^*, w_1, w_2, \ldots, w_n) = C[S, 0, n] \]
P-CKY algorithm from book

function PROBABILISTIC-CKY(words, grammar) returns most probable parse and its probability

for $j \leftarrow$ from 1 to LENGTH(words) do
  for all $\{ A \mid A \rightarrow \text{words}[j] \in \text{grammar} \}$
    $\text{table}[j - 1, j, A] \leftarrow P(A \rightarrow \text{words}[j])$
  for $i \leftarrow$ from $j - 2$ downto 0 do
    for $k \leftarrow i + 1$ to $j - 1$ do
      for all $\{ A \mid A \rightarrow BC \in \text{grammar},$
        and $\text{table}[i, k, B] > 0$ and $\text{table}[k, j, C] > 0 \}$
        if $(\text{table}[i, j, A] < P(A \rightarrow BC) \times \text{table}[i, k, B] \times \text{table}[k, j, C])$ then
          $\text{table}[i, j, A] \leftarrow P(A \rightarrow BC) \times \text{table}[i, k, B] \times \text{table}[k, j, C]$
          $\text{back}[i, j, A] \leftarrow \{k, B, C\}$
      return BUILD_TREE($\text{back}[1, \text{LENGTH}(\text{words}), S]), \text{table}[1, \text{LENGTH}(\text{words}), S]$
<table>
<thead>
<tr>
<th>Det: 0.40</th>
<th>NP: 0.30 * 0.40 * 0.02 = 0.0024</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0,1]</td>
<td>[0,2]</td>
</tr>
<tr>
<td>N: 0.02</td>
<td>[1,2]</td>
</tr>
<tr>
<td>[1,3]</td>
<td>[1,4]</td>
</tr>
<tr>
<td>V: 0.05</td>
<td>[2,3]</td>
</tr>
<tr>
<td>[2,4]</td>
<td>[3,4]</td>
</tr>
<tr>
<td>[3,5]</td>
<td>[4,5]</td>
</tr>
</tbody>
</table>
Parsing as (Weighted) Deduction

Antecedents

\[ \frac{A : u \quad B : v}{C : w} \]

Side conditions

\[ \phi \]

Consequent

“If \( A \) and \( B \) are true with weights \( u \) and \( v \), and phi is also true, then \( C \) is true with weight \( w \).”
Earley’s Algorithm
Jay Earley, PhD

My Self-Therapy Journey Story

Self-Therapy Journey (STJ) is my brain-child and my life purpose. It brings together the two professional sides of me that have been separate up till now—the computer scientist and the psychologist. It is very exciting to integrate these two disparate sides of me in this amazing project.
Example Grammar (same for CKY)

<table>
<thead>
<tr>
<th>Grammar</th>
<th>Lexicon</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S \rightarrow NP\ VP$</td>
<td>$Det \rightarrow that</td>
</tr>
<tr>
<td>$S \rightarrow Aux\ NP\ VP$</td>
<td>$Noun \rightarrow book</td>
</tr>
<tr>
<td>$S \rightarrow VP$</td>
<td>$Verb \rightarrow book</td>
</tr>
<tr>
<td>$NP \rightarrow Pronoun$</td>
<td>$Pronoun \rightarrow I</td>
</tr>
<tr>
<td>$NP \rightarrow Proper-Noun$</td>
<td>$Proper-Noun \rightarrow Houston</td>
</tr>
<tr>
<td>$NP \rightarrow Det\ Nominal$</td>
<td>$Aux \rightarrow does$</td>
</tr>
<tr>
<td>$Nominal \rightarrow Noun$</td>
<td>$Preposition \rightarrow from</td>
</tr>
<tr>
<td>$Nominal \rightarrow Nominal\ Noun$</td>
<td>$Nominal \rightarrow Nominal PP$</td>
</tr>
<tr>
<td>$Nominal \rightarrow Nominal\ PP$</td>
<td>$VP \rightarrow Verb$</td>
</tr>
<tr>
<td>$VP \rightarrow Verb\ NP$</td>
<td>$VP \rightarrow Verb\ NP\ PP$</td>
</tr>
<tr>
<td>$VP \rightarrow Verb\ PP$</td>
<td>$VP \rightarrow VP\ PP$</td>
</tr>
<tr>
<td>$PP \rightarrow Preposition\ NP$</td>
<td>$</td>
</tr>
</tbody>
</table>
Earley Parsing

• Allows arbitrary CFGs
• Top-down control
• Fills a table (or chart) in a single sweep over the input
  – Table is length N+1; N is number of words
  – Table entries represent
    • Completed constituents and their locations
    • In-progress constituents
    • Predicted constituents
States

• The table-entries are called states and are represented with **dotted-rules**.

\[
\begin{align*}
S & \rightarrow \text{VP} & \text{A VP is predicted} \\
\text{NP} & \rightarrow \text{Det} \cdot \text{Nominal} & \text{An NP is in progress} \\
\text{VP} & \rightarrow \text{V NP} & \text{A VP has been found}
\end{align*}
\]
States/Locations

• S . VP [0,0]  
A VP is predicted at the start of the sentence

• NP Det . Nominal [1,2]  
An NP is in progress; the Det goes from 1 to 2

• VP V NP . [0,3]  
A VP has been found starting at 0 and ending at 3
Earley top-level

• As with most dynamic programming approaches, the answer is found by looking in the table in the right place.
• In this case, there should be an $S$ state in the final column that spans from 0 to $N$ and is complete. That is,

$$S \xrightarrow{\alpha} [0,N]$$

• If that’s the case, you’re done.
Earley top-level (2)

• So sweep through the table from 0 to N...
  – New predicted states are created by starting top-down from $S$
  – New incomplete states are created by advancing existing states as new constituents are discovered
  – New complete states are created in the same way.
Earley top-level (3)

• More specifically...
  1. *Predict* all the states you can upfront
  2. Read a word
     1. Extend states based on matches
     2. Generate new predictions
     3. Go to step 2
  3. When you’re out of words, look at the chart to see if you have a winner
function EARLEY-PARSE(words, grammar) returns chart

ENQUEUE((γ → • S, [0, 0]), chart[0])
for i ← from 0 to LENGTH(words) do
  for each state in chart[i] do
    if INCOMPLETE?(state) and
      NEXT-CAT(state) is not a part of speech then
        PREDICTOR(state)
    elseif INCOMPLETE?(state) and
      NEXT-CAT(state) is a part of speech then
        SCANNER(state)
    else
      COMPLETER(state)
  end
end
return(chart)
Earley code: 3 main functions

procedure PREDICTOR((A → α • B β, [i, j]))
    for each (B → γ) in GRAMMAR-RULES-FOR(B, grammar) do
        ENQUEUE((B → • γ, [j, j]), chart[j])
    end

procedure SCANNER((A → α • B β, [i, j]))
    if B ⊂ PARTS-OF-SPEECH(word[j]) then
        ENQUEUE((B → word[j], [j, j + 1]), chart[j+1])
    end

procedure COMPLETER((B → γ •, [j, k]))
    for each (A → α • B β, [i, j]) in chart[j] do
        ENQUEUE((A → α B • β, [i, k]), chart[k])
    end

procedure ENQUEUE(state, chart-entry)
    if state is not already in chart-entry then
        PUSH(state, chart-entry)
    end
Extended Earley Example

• *Book that flight*

• We should find:
  an $S$ from 0 to 3 that is a completed state
<table>
<thead>
<tr>
<th>Chart[0]</th>
<th>S0</th>
<th>$\gamma \rightarrow \bullet S$</th>
<th>[0,0]</th>
<th>Dummy start state</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>$S \rightarrow \bullet NP \ VP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>$S \rightarrow \bullet Aux \ NP \ VP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>$S \rightarrow \bullet VP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>$NP \rightarrow \bullet Pronoun$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>$NP \rightarrow \bullet Proper-Noun$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S6</td>
<td>$NP \rightarrow \bullet Det \ Nominal$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>$VP \rightarrow \bullet Verb$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td>$VP \rightarrow \bullet Verb \ NP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S9</td>
<td>$VP \rightarrow \bullet Verb \ NP \ PP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S10</td>
<td>$VP \rightarrow \bullet Verb \ PP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S11</td>
<td>$VP \rightarrow \bullet VP \ PP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>Chart[1]</td>
<td>S12</td>
<td>$\text{Verb} \rightarrow \text{book} \bullet$</td>
<td>[0,1]</td>
<td>Scanner</td>
</tr>
<tr>
<td>S13</td>
<td>$VP \rightarrow \text{Verb} \bullet$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S14</td>
<td>$VP \rightarrow \text{Verb} \bullet NP$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S15</td>
<td>$VP \rightarrow \text{Verb} \bullet NP \ PP$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S16</td>
<td>$VP \rightarrow \text{Verb} \bullet PP$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S17</td>
<td>$S \rightarrow VP \bullet$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S18</td>
<td>$VP \rightarrow VP \bullet PP$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S19</td>
<td>$NP \rightarrow \bullet Pronoun$</td>
<td>[1,1]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S20</td>
<td>$NP \rightarrow \bullet Proper-Noun$</td>
<td>[1,1]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S21</td>
<td>$NP \rightarrow \bullet Det \ Nominal$</td>
<td>[1,1]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S22</td>
<td>$PP \rightarrow \bullet Prep \ NP$</td>
<td>[1,1]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>Chart[0]</td>
<td>S0</td>
<td>$\gamma \rightarrow \bullet S$</td>
<td>[0,0]</td>
<td>Dummy start state</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>----------------------------------</td>
<td>-------</td>
<td>-------------------</td>
</tr>
<tr>
<td>S1</td>
<td>$S \rightarrow \bullet NP \ VP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>$S \rightarrow \bullet Aux NP \ VP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>$S \rightarrow \bullet VP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>$NP \rightarrow \bullet Pronoun$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>$NP \rightarrow \bullet Proper-Noun$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S6</td>
<td>$NP \rightarrow \bullet Det Nominal$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>$VP \rightarrow \bullet Verb$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td>$VP \rightarrow \bullet Verb NP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S9</td>
<td>$VP \rightarrow \bullet Verb NP \ PP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S10</td>
<td>$VP \rightarrow \bullet Verb PP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S11</td>
<td>$VP \rightarrow \bullet VP PP$</td>
<td>[0,0]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>Chart[1]</td>
<td>S12</td>
<td>$\text{Verb} \rightarrow book \bullet$</td>
<td>[0,1]</td>
<td>Scanner</td>
</tr>
<tr>
<td>S13</td>
<td>$VP \rightarrow \text{Verb} \bullet$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S14</td>
<td>$VP \rightarrow \text{Verb} \bullet NP$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S15</td>
<td>$VP \rightarrow \text{Verb} \bullet NP \ PP$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S16</td>
<td>$VP \rightarrow \text{Verb} \bullet PP$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S17</td>
<td>$S \rightarrow VP \bullet$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S18</td>
<td>$VP \rightarrow VP \bullet PP$</td>
<td>[0,1]</td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S19</td>
<td>$NP \rightarrow \bullet Pronoun$</td>
<td>[1,1]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S20</td>
<td>$NP \rightarrow \bullet Proper-Noun$</td>
<td>[1,1]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S21</td>
<td>$NP \rightarrow \bullet Det Nominal$</td>
<td>[1,1]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S22</td>
<td>$PP \rightarrow \bullet Prep NP$</td>
<td>[1,1]</td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>Chart[2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>S23</td>
<td>$Det \rightarrow \text{that}$</td>
<td>$\bullet$</td>
<td>[1,2]</td>
<td></td>
</tr>
<tr>
<td>S24</td>
<td>$NP \rightarrow Det \bullet \text{Nominal}$</td>
<td>[1,2]</td>
<td></td>
<td>Completer</td>
</tr>
<tr>
<td>S25</td>
<td>$\text{Nominal} \rightarrow \bullet \text{Noun}$</td>
<td>[2,2]</td>
<td></td>
<td>Predictor</td>
</tr>
<tr>
<td>S26</td>
<td>$\text{Nominal} \rightarrow \bullet \text{Nominal Noun}$</td>
<td>[2,2]</td>
<td></td>
<td>Predictor</td>
</tr>
<tr>
<td>S27</td>
<td>$\text{Nominal} \rightarrow \bullet \text{Nominal PP}$</td>
<td>[2,2]</td>
<td></td>
<td>Predictor</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chart[3]</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S28</td>
<td>$\text{Noun} \rightarrow \text{flight}$</td>
<td>$\bullet$</td>
<td>[2,3]</td>
<td></td>
<td>Scanner</td>
</tr>
<tr>
<td>S29</td>
<td>$\text{Nominal} \rightarrow \text{Noun}$</td>
<td>$\bullet$</td>
<td>[2,3]</td>
<td></td>
<td>Completer</td>
</tr>
<tr>
<td>S30</td>
<td>$NP \rightarrow Det \text{Nominal}$</td>
<td>$\bullet$</td>
<td>[1,3]</td>
<td></td>
<td>Completer</td>
</tr>
<tr>
<td>S31</td>
<td>$\text{Nominal} \rightarrow \text{Nominal}$</td>
<td>$\bullet \text{Noun}$</td>
<td>[2,3]</td>
<td></td>
<td>Completer</td>
</tr>
<tr>
<td>S32</td>
<td>$\text{Nominal} \rightarrow \text{Nominal}$</td>
<td>$\bullet \text{PP}$</td>
<td>[2,3]</td>
<td></td>
<td>Completer</td>
</tr>
<tr>
<td>S33</td>
<td>$VP \rightarrow \text{Verb} NP$</td>
<td>$\bullet$</td>
<td>[0,3]</td>
<td></td>
<td>Completer</td>
</tr>
<tr>
<td>S34</td>
<td>$VP \rightarrow \text{Verb} NP \bullet \text{PP}$</td>
<td>[0,3]</td>
<td></td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>S35</td>
<td>$PP \rightarrow \bullet \text{Prep NP}$</td>
<td>[3,3]</td>
<td></td>
<td>Predictor</td>
<td></td>
</tr>
<tr>
<td>S36</td>
<td>$S \rightarrow VP$</td>
<td>$\bullet$</td>
<td>[0,3]</td>
<td></td>
<td>Completer</td>
</tr>
<tr>
<td>S37</td>
<td>$VP \rightarrow VP \bullet \text{PP}$</td>
<td>[0,3]</td>
<td></td>
<td>Completer</td>
<td></td>
</tr>
<tr>
<td>Chart[2]</td>
<td>S23</td>
<td>Det → that ⬤</td>
<td>[1,2]</td>
<td>Scanner</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----</td>
<td>---------------</td>
<td>-------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>S24</td>
<td>NP  → Det ⬤ Nominal</td>
<td>[1,2]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S25</td>
<td>Nominal → ⬤ Noun</td>
<td>[2,2]</td>
<td>Predictor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S26</td>
<td>Nominal → ⬤ Nominal Noun</td>
<td>[2,2]</td>
<td>Predictor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S27</td>
<td>Nominal → ⬤ Nominal PP</td>
<td>[2,2]</td>
<td>Predictor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S29</td>
<td>Nominal → Noun ⬤</td>
<td>[2,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S30</td>
<td>NP  → Det Nominal ⬤</td>
<td>[1,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S31</td>
<td>Nominal → Nominal ⬤ Noun</td>
<td>[2,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S32</td>
<td>Nominal → Nominal ⬤ PP</td>
<td>[2,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S33</td>
<td>VP  → Verb NP ⬤</td>
<td>[0,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S34</td>
<td>VP  → Verb NP ⬤ PP</td>
<td>[0,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S35</td>
<td>PP  → ⬤ Prep NP</td>
<td>[3,3]</td>
<td>Predictor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S36</td>
<td>S → VP ⬤</td>
<td>[0,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S37</td>
<td>VP  → VP ⬤ PP</td>
<td>[0,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S23</td>
<td>Det $\rightarrow$ that $\cdot$</td>
<td>[1,2]</td>
<td>Scanner</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>----------------------------------</td>
<td>------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S24</td>
<td>NP $\rightarrow$ Det $\cdot$ Nominal</td>
<td>[1,2]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S25</td>
<td>Nominal $\rightarrow$ $\cdot$ Noun</td>
<td>[2,2]</td>
<td>Predictor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S26</td>
<td>Nominal $\rightarrow$ $\cdot$ Nominal Noun</td>
<td>[2,2]</td>
<td>Predictor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S27</td>
<td>Nominal $\rightarrow$ $\cdot$ Nominal PP</td>
<td>[2,2]</td>
<td>Predictor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S28</td>
<td>Noun $\rightarrow$ flight $\cdot$</td>
<td>[2,3]</td>
<td>Scanner</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S29</td>
<td>Nominal $\rightarrow$ Noun $\cdot$</td>
<td>[2,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S30</td>
<td>NP $\rightarrow$ Det Nominal $\cdot$</td>
<td>[1,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S31</td>
<td>Nominal $\rightarrow$ Nominal $\cdot$ Noun</td>
<td>[2,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S32</td>
<td>Nominal $\rightarrow$ Nominal $\cdot$ PP</td>
<td>[2,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S33</td>
<td>VP $\rightarrow$ Verb NP $\cdot$</td>
<td>[0,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S34</td>
<td>VP $\rightarrow$ Verb NP $\cdot$ PP</td>
<td>[0,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S35</td>
<td>PP $\rightarrow$ $\cdot$ Prep NP</td>
<td>[3,3]</td>
<td>Predictor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S36</td>
<td>S $\rightarrow$ VP $\cdot$</td>
<td>[0,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S37</td>
<td>VP $\rightarrow$ VP $\cdot$ PP</td>
<td>[0,3]</td>
<td>Completer</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Earley’s Algorithm in equations

• We can look at this from the declarative programming point of view too.

<table>
<thead>
<tr>
<th>ROOT → • S [0,0]</th>
<th>book</th>
<th>the</th>
<th>flight</th>
<th>through</th>
<th>Chicago</th>
</tr>
</thead>
</table>

goal: 
ROOT → S• [0,n]
Earley’s Algorithm:  PREDICT

Given \( V \to \alpha \bullet X \beta \ [i, j] \)
and the rule \( X \to \gamma \),
create \( X \to \bullet \gamma \ [j, j] \)

<table>
<thead>
<tr>
<th>book</th>
<th>the</th>
<th>flight</th>
<th>through</th>
<th>Chicago</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROOT ( \to \bullet S \ [0,0] )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S ( \to \bullet VP \ [0,0] )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S ( \to \bullet NP \ VP \ [0,0] )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VP ( \to \bullet V \ NP \ [0,0] )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP ( \to \bullet DT \ N \ [0,0] )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Earley’s Algorithm: SCAN

Given $V \rightarrow \alpha \bullet T \beta [i, j]$ and the rule $T \rightarrow w_{j+1}$, create $T \rightarrow w_{j+1} \bullet [j, j+1]$

<table>
<thead>
<tr>
<th>ROOT $\rightarrow \bullet S [0,0]$</th>
<th>V $\rightarrow$ book $\bullet$ [0,1]</th>
<th>book</th>
<th>the</th>
<th>flight</th>
<th>through</th>
<th>Chicago</th>
</tr>
</thead>
<tbody>
<tr>
<td>S $\rightarrow \bullet VP [0,0]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S $\rightarrow \bullet NP VP [0,0]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VP $\rightarrow \bullet V NP [0,0]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VP $\rightarrow \bullet V NP [0,0]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP $\rightarrow \bullet DT N [0,0]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP $\rightarrow \bullet DT N [0,0]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$VP \rightarrow \bullet V NP [0,0]$  
$V \rightarrow$ book  
$V \rightarrow$ book $\bullet$ [0,1]
Earley’s Algorithm: COMPLETE

Given $V \rightarrow \alpha \bullet X \beta$ \([i, j]\) and $X \rightarrow \gamma \bullet$ \([j, k]\), create $V \rightarrow \alpha X \bullet \beta$ \([i, k]\).

| ROOT $\rightarrow$ $\bullet$ S [0,0] |
|-----|----------------------------------|
| S $\rightarrow$ $\bullet$ VP [0,0] |
| S $\rightarrow$ $\bullet$ NP VP [0,0] |
| ... |
| VP $\rightarrow$ $\bullet$ V NP [0,0] |
| ... |
| NP $\rightarrow$ $\bullet$ DT N [0,0] |
| ... |
| V $\rightarrow$ book [0,1] |
| VP $\rightarrow$ V $\bullet$ NP [0,1] |

book the flight through Chicago
Earley’s Algorithm

Goal: \([\text{ROOT} \rightarrow S\bullet, 0, n]\)  
Items: \([X \rightarrow \alpha \bullet \beta, i, j]\)

Axiom: \([\text{ROOT} \rightarrow \bullet S, 0, 0]\)

Inference rules:

\[
\begin{align*}
\text{PREDICT} & \quad \frac{[V \rightarrow \alpha \bullet X\beta, i, j]}{[X \rightarrow \bullet \gamma, j, j] : p} \quad X \xrightarrow{p} \gamma \in R \\
\text{SCAN} & \quad \frac{[V \rightarrow \alpha \bullet w_{j+1}\beta, i, j] : p}{[V \rightarrow \alpha w_{j+1} \bullet \beta, i, j] : p} \\
\text{COMPLETE} & \quad \frac{[X \rightarrow \gamma\bullet, j, k] : p \quad [V \rightarrow \alpha \bullet X\beta, i, j] : q}{[V \rightarrow \alpha X \bullet \beta, i, k] : pq}
\end{align*}
\]
Thought Questions

• Runtime?
  – $O(n^3)$
• Memory?
  – $O(n^2)$
• Can we make it faster?
• Recovering trees?
Make it an Earley Parser

- Record which sub rules we used to complete edges
Heads in CFGs
The luxury auto maker last year sold 1,214 cars in the U.S.
The luxury auto maker last year sold 1,214 cars in the U.S.
The luxury auto maker last year sold 1,214 cars in the U.S.
The luxury auto maker last year sold 1,214 cars in the U.S.
it can remember one million truly inspiring teachers from Rainbow Technologies.
I have been able to force to be more receptive to therapy, and to keep the committee informed, usually in advance, of covert actions; the victims are large and costly machines.
As their varied strategies suggest, Another suggestion would predict they will pay off.
The two-day trip reportedly has said it would be done.
Others have soared to the car market well.
A spokesman for paying the bill declined to pay taxes, but the fact that adjusted payouts on behalf of preventative medicine in terms of 29 years could be distributed.
P&G, in the space of Orrick, Herrington & Sutcliffe, rarely rolls forward on a modest 1.1 million shares on the block.
In the eight months last Friday, bond prices closed yesterday at $ 30.2 million, down 25 cents.
Still, Honda says is calling for slight declines when there was posted within its pre-1967 borders.
Moreover, Allianz's Mr. Jarrett also sees only a  ``` internal erosion ``` of about 35 of St. Petersburg, Fla. due 1994.
it is predicting negative third: and fourth-quarter growth.
Grace said luxury-car sales increased 1.4 % to 221.61 billion yen -LRB- $ 188.2 -RRB- , from $ 234.4 million a share, or $ 9.6 million, a year earlier.
But AGIP already has been group vice president for such a gizmo at Texas Air.
And when other rules are safeguarded by the Appropriations Committee, the White House passed a $ 1.5765 billion loan market-revision bill providing the first construction funds for the economy's ambitious radio station in fiscal 1990 and incorporating far-reaching provisions affecting the erratic copper market.
The urging also has yet opened in September in September.
But Mr. Lorenzo is to elaborate on the latest reports of the line.
Some Related Rules

- \( \text{NAC} \rightarrow \text{NNP}, \text{NNP} \quad 0.002463 \)
- \( \text{NAC} \rightarrow \text{JJ NNP}, \text{NNP} \quad 0.002463 \)
- \( \text{NAC} \rightarrow \text{NNP}, \text{NNP} \quad 0.002463 \)
- \( \text{NAC} \rightarrow \text{NNP CD}, \text{CD} \quad 0.002463 \)
- \( \text{NAC} \rightarrow \text{NNP NNP NNP}, \text{NNP} \quad 0.002463 \)
- \( \text{NAC} \rightarrow \text{NNP NNP}, \text{NNP} \quad 0.004926 \)
- \( \text{NAC} \rightarrow \text{NNP NNPS}, \text{NNP} \quad 0.007389 \)
- \( \text{NAC} \rightarrow \text{NNP NNP}, \text{NNP} \quad 0.019704 \)
- \( \text{NAC} \rightarrow \text{NNP NNP CD}, \text{CD} \quad 0.024631 \)
- \( \text{NAC} \rightarrow \text{NNP NNP}, \text{NNP} \quad 0.125616 \)
- \( \text{NAC} \rightarrow \text{NNP NNP}, \text{NNP} \quad 0.374384 \)
Bigram Model for NAC
Lexicalized Rules
Markovizing Lexicalized Rules

\[ VP+dumped+VBD \rightarrow VBD+dumped+VBD \]
\[ p(\text{Heir} = VBD+dumped+VBD \mid \text{Parent} = VP+dumped+VBD) \]

\[ VP+dumped+VBD \rightarrow ^{\uparrow} VBD+dumped+VBD \]
\[ p(\text{left-stop} \mid \text{Parent} = VP+dumped+VBD, \text{Heir} = VBD+dumped+VBD) \]

\[ VP+dumped+VBD \rightarrow ^{\uparrow} VBD+dumped+VBD \ np+sacks+NNS \]
\[ p(\text{RightChild} = \text{NP+sacks+NNS} \mid \text{Parent} = VP+dumped+VBD, \text{Heir} = VBD+dumped+VBD) \]

\[ VP+dumped+VBD \rightarrow ^{\uparrow} VBD+dumped+VBD \ np+sacks+NNS \ pp+into+P \]
\[ p(\text{RightChild} = \text{PP+into+P} \mid \text{Parent} = VP+dumped+VBD, \text{Heir} = VBD+dumped+VBD) \]

\[ VP+dumped+VBD \rightarrow ^{\uparrow} VBD+dumped+VBD \ np+sacks+NNS \ pp+into+P \$ \]
\[ p(\text{right-stop} \mid \text{Parent} = VP+dumped+VBD, \text{Heir} = VBD+dumped+VBD) \]
The luxury auto maker last year sold 1,214 cars in the U.S.
The luxury auto maker last year sold 1,214 cars in the U.S.
Dependency vs Constituent

Dependency:
```plaintext
We are trying to understand the difference.
```

Constituency (BPS):
```plaintext
We are trying to understand the difference.
```
Dependency Trees

• Links between heads and their dependents
  – Head is a Linguistic notion
  – Sort of “most important part”

• Only one head, acyclic

• Why?
  – Can be simpler to parse
  – Can be simpler for later ML processes
Dependency Trees

• Links between heads and their dependents
  – Head is a Linguistic notion
  – Sort of “most important part”
• Only one head, acyclic
• Why?
  – Can be simpler to parse
  – Can be simpler for later ML processes
• CT -> DT easier than DT -> CT
What is the head?

• Auxiliaries or main verbs?
  – I have written a letter.

• Prepositions or nouns?
  – A picture of my son

• Clause-initial elements? (Complementizers)
  – Who yawned?
  – I wonder which people yawned.
  – The student who yawned.
  – I think that the student yawned.

• Parts, kinds, and quantities?
  – I drank a cup of tea.
  – I drank a kind of tea.
  – I talked to a number of people.
Which word is the head?

- Lexical words
  - the book
  - at school
  - has yawned

- Function words
  - the book
  - at school
  - has yawned

- Open class: you can make up new nouns and verbs

- Closed class: you cannot make up new determiners, prepositions, or auxiliary verbs (although new ones can develop over time)

Stanford Dependency Parser provides two versions: lexical heads or functional heads
What you see most often in dependency treebanks

- the *book*
- *at* school
- The student *has yawned*
- The student *has yawned*
- very *tall*
- *that* the student yawned
- *that* the student *yawned*
  - As in “I think that the student yawned”
So what is the definition of “head”? 

• The word that provides the main meaning:  
  – “this smart student of linguistics with long hair” is a student, not a smart or a hair or a long, etc. So “student” is the head.

• The word that provides the most important inflectional features  
  – Inflection includes things like tense, number, and gender
Which noun phrases are plural?

**Singular**
- The teacher
- The short teacher
- The teacher of the class
- The teacher of the classes
- The children’s teacher
- The child’s teacher

**Plural**
- The teachers
- The short teachers
- The teachers of the class
- The teachers of the classes
- The children’s teachers
- The child’s teachers

Only the head “teacher/teachers” determines whether the noun phrase is singular or plural. The other nouns “class/classes” and “child/children” do not make the noun phrase singular or plural.
Dependency Parsing

• Standard CFG (with Heads) plus CKY
  – But more computationally expensive
• Graph Algorithms
  – e.g. McDonald’s MSTParse (Maximum Spanning Tree)
• Constraint satisfaction
  – Create all links and remove them (Karlsson 1990)
• Or actual parse the dependencies
  – Nivre et al 2008: MaltParser
• Neural dependency parses (Chen & Manning 2014)
Dependency Parsing

• Parse left to right
  – Make decisions about linking and shifting

• Use ML classifier to decide what to do
  – Condition on
  – Some lexical word links are more common [ chair -> the]
  – Dependency distance: mostly short links
  – Intervening material: rarely span over verbs, punc
  – Valency of heads: number of expect dependents of a head
ROOT Discussion of the outstanding issues was completed.
Ze hadden languit naast elkaar op de strandstoelen kunnen gaan liggen.
Other Grammar Formalisms
Unification-Based Grammars

• $S \rightarrow NP \ VP$
  \[ [NP \ NUMBER] = [VP \ NUMBER] \]

• Det $\rightarrow$ these
  \[ [Det \ NUMBER] = \text{plural} \]

• MD $\rightarrow$ does
  \[ [MD \ NUMBER] = \text{singular} \]
  \[ [MD \ PERSON] = \text{third} \]
Categorial Grammar (CCG)

• 5 rules
• $A/B + B = A$
• $B + A \backslash B = A$
• $A/B + B/C = A/C$
• $A \text{ CONJ } A' = A$
• $A = X/(X \backslash A)$
• But the lexical items become more complex
Categorial Grammar (CCG)

John = np  
Mary = np  
likes = (s\np)/np

Forward application  
X/Y Y => X  
Backward application  
Y X\Y => X

Thus  
John   likes   Mary
np   (s\np)/np   np

---------- Forward
s\np

---------- Backward
s
Categorial Grammar (CCG)

a, the np/n
old n/n
in (np\np)/np
man, ball, park n
kicked (s\np)/np

the old man kicked a ball in the park
np/n n/n n (s\np)/np np/n n (np\np)/np np np/n n

n np

np np

np np

np np

s np

s
Advanced Grammars

- Standard CFG
- CKY vs Earley
- Lexicalized Grammars
- Other formalisms
  - TAG, HPSG, LFG
  - Unification Grammars
  - Categorial Grammars