NLP
Introduction to NLP

Probabilistic Parsing (2/2)
Main Tasks with PCFGs

• Given a grammar G and a sentence s, let T(s) be all parse trees that correspond to s

• Task 1
  – find which tree t among T(s) maximizes the probability p(t)

• Task 2
  – find the probability of the sentence p(s) as the sum of all possible tree probabilities p(t)
Probabilistic Parsing Methods

• Probabilistic Earley algorithm
  – Top-down parser with a dynamic programming table

• Probabilistic Cocke–Kasami–Younger (CKY) algorithm
  – Bottom-up parser with a dynamic programming table
Probabilistic Grammars

• Probabilities can be learned from a training corpus
  – Treebank

• Intuitive meaning
  – Parse #1 is twice as probable as parse #2

• Possible to do reranking

• Possible to combine with other stages
  – E.g., speech recognition, translation
Maximum Likelihood Estimates

• Use the parsed training set for getting the counts
  \[ P_{ML}(\alpha \rightarrow \beta) = \frac{\text{Count}(\alpha \rightarrow \beta)}{\text{Count}(\alpha)} \]

• Example:
  \[ P_{ML}(S \rightarrow \text{NP VP}) = \frac{\text{Count}(S \rightarrow \text{NP VP})}{\text{Count}(S)} \]
<table>
<thead>
<tr>
<th>Grammar</th>
<th>Lexicon</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S \rightarrow NP \ VP$</td>
<td>$Det \rightarrow that$ [.10] $</td>
</tr>
<tr>
<td>$S \rightarrow Aux \ NP \ VP$</td>
<td>$Noun \rightarrow book$ [.10] $</td>
</tr>
<tr>
<td>$S \rightarrow VP$</td>
<td></td>
</tr>
<tr>
<td>$NP \rightarrow Pronoun$</td>
<td>$meal$ [.15] $</td>
</tr>
<tr>
<td>$NP \rightarrow Proper-Noun$</td>
<td>$flights$ [.40] $</td>
</tr>
<tr>
<td>$NP \rightarrow Det Nominal$</td>
<td>$Verb \rightarrow book$ [.30] $</td>
</tr>
<tr>
<td>$NP \rightarrow Nominal$</td>
<td>$prefer$ [.40]</td>
</tr>
<tr>
<td>Nominal $\rightarrow Noun$</td>
<td>$Pronoun \rightarrow I$ [.40] $</td>
</tr>
<tr>
<td>Nominal $\rightarrow Nominal$</td>
<td>$Verb \rightarrow book$ [.30] $</td>
</tr>
<tr>
<td>Nominal $\rightarrow Nominal\ Noun$</td>
<td>$prefer$ [.40]</td>
</tr>
<tr>
<td>Nominal $\rightarrow Nominal\ PP$</td>
<td>$Pronoun \rightarrow I$ [.40] $</td>
</tr>
<tr>
<td>VP $\rightarrow Verb$</td>
<td>$NWA$ [.40]</td>
</tr>
<tr>
<td>VP $\rightarrow Verb\ NP$</td>
<td>$Aux \rightarrow does$ [.60] $</td>
</tr>
<tr>
<td>VP $\rightarrow Verb\ NP\ PP$</td>
<td>$Preposition \rightarrow from$ [.30] $</td>
</tr>
<tr>
<td>VP $\rightarrow Verb\ PP$</td>
<td>$on$ [.20] $</td>
</tr>
<tr>
<td>VP $\rightarrow VP\ PP$</td>
<td>$through$ [.05]</td>
</tr>
<tr>
<td>PP $\rightarrow Preposition\ NP$</td>
<td></td>
</tr>
</tbody>
</table>
Example

S -> NP VP  \[p0=1\]
NP -> DT N  \[p1=.8\]
NP -> NP PP  \[p2=.2\]
PP -> PRP NP  \[p3=1\]
VP -> V NP  \[p4=.7\]
VP -> VP PP  \[p5=.3\]
DT -> 'a'  \[p6=.25\]
DT -> 'the'  \[p7=.75\]
N -> 'child'  \[p8=.5\]
N -> 'cake'  \[p9=.3\]
N -> 'fork'  \[p10=.2\]
PRP -> 'with'  \[p11=.1\]
PRP -> 'to'  \[p12=.9\]
V -> 'saw'  \[p13=.4\]
V -> 'ate'  \[p14=.6\]
the child ate the cake with the fork
The child ate the cake with the fork.
the child ate the cake with the fork
The child ate the cake with the fork.
The child ate the cake with the fork.
the child ate the cake with the fork.
Question

• Now, on your own, compute the probability of the entire sentence using Probabilistic CKY.
• Don’t forget that there may be multiple parses, so you will need to add the corresponding probabilities.
Notes

• Stanford Demo

• PTB statistics
  – 50,000 sentences (40,000 training; 2,400 testing)

• PTB peculiarities
  – includes traces and other null elements
  – Flat NP structure (e.g., NP -> DT JJ JJ NNP NNS)

• Parent transformation
  – Subject NPs are more likely to be modified than object NPs
  – E.g., replace NP with NP^S
NLP