

CS 378 Lecture 13

CKY, Dependency

Announcements

- A2 back, solutions posted soon
- A3 due Tuesday
- Midterm: review next class

Recap PCFGs

N, T S R
symbols words) start / rules, probs

| | | | |
|--------------|--|--------------|-----|
| DT → the | | S → NP VP | |
| NN → child | | NP → DT NN | 1/2 |
| NNS → raises | | NP → NN NNS | 1/2 |
| VBZ → raises | | VP → VBZ PRP | |
| PRP → it | | | |

$$\log(1/2) = -1$$

P(rule | parent)

CKY

Input: PCFG
sentence \bar{x}

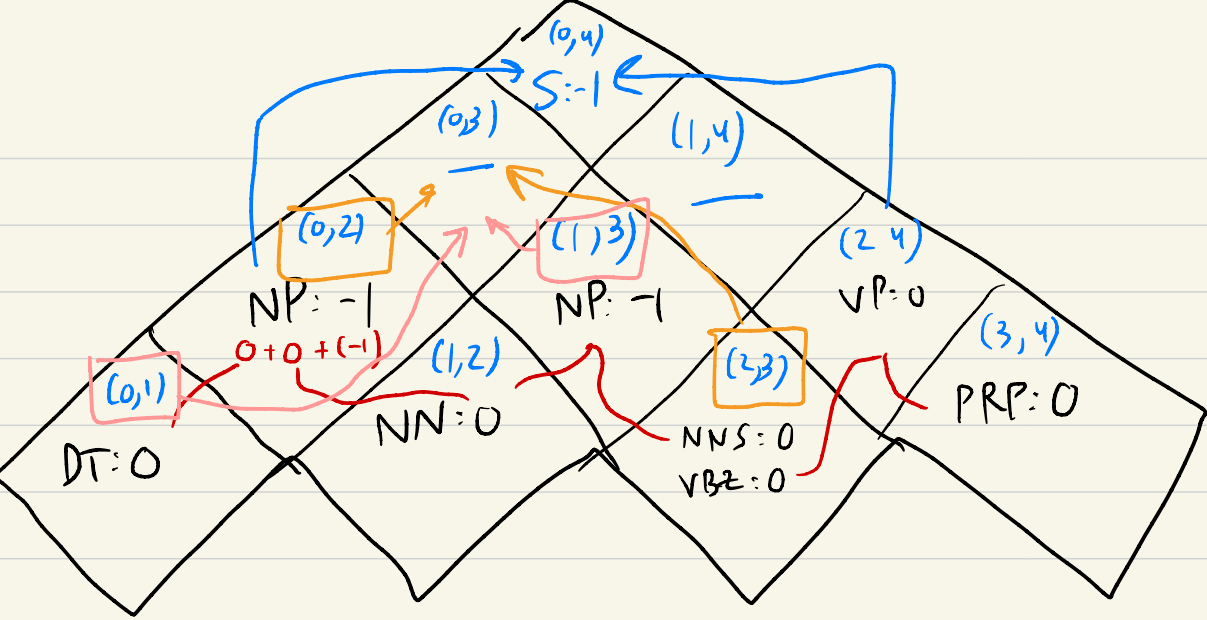
Output: $\underset{T}{\operatorname{argmax}} P(T|\bar{x})$
most likely tree for \bar{x}

base: $T(i, j+1, X) = \log P(w_i | X)$

recursive:

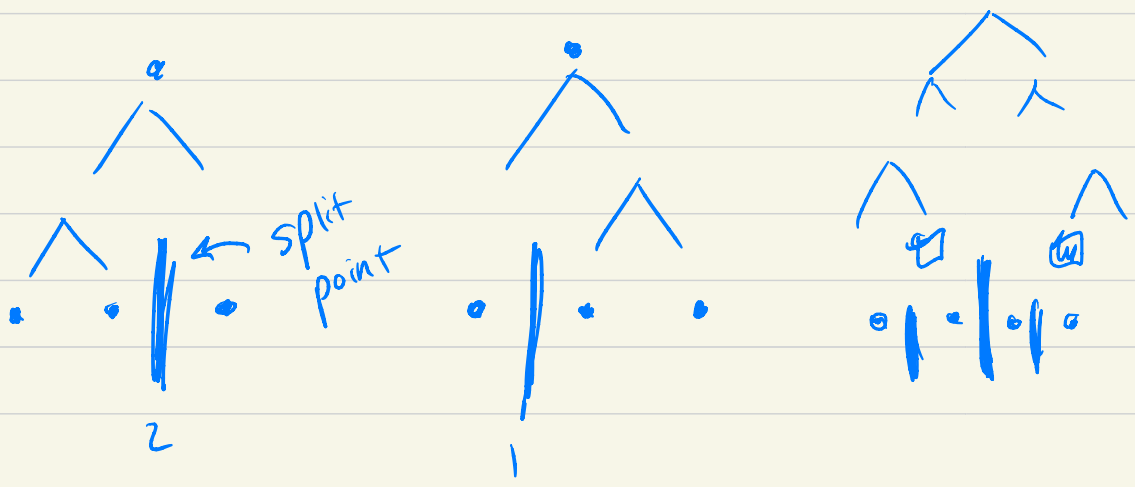
$$T(i, j, X) = \max_{\substack{k: i < k < j \\ \text{split point}}} \max_{\substack{X \rightarrow X_1 X_2 \\ \text{rule}}} \left[\log P(X \rightarrow X_1 X_2) \right. \\ \left. + T[i, k, X_1] + T[k, j, X_2] \right]$$

T: len \times len \times num symbols



the child raises it
 0 1 2 3 4

each cell: log prob of best way of building symbol. No symbol \rightarrow impossible ($-\infty$)



Poll

S: -1

①

NP: -1

VP: 0

NN: 0

NNS: 0
VBZ: 0

NNS: 0
VBZ: 0

PRP: 0

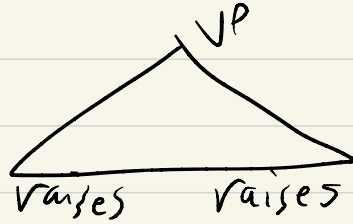
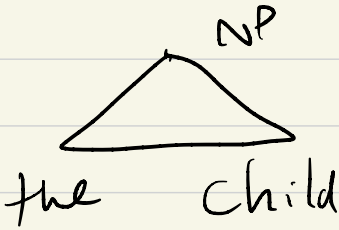
Child

raises

raises

it

② Add VP → VBZ NNS
decrease VP → VBZ PRP



PRP → raises

Time complexity: sentence length n
 $O(n^3 G)$ CKY + unaries

S
|
NP