Natural Language Processing
Syntactic parsing

Sofia Serrano
sofias6@cs.washington.edu

Credit to Yulia Tsvetkov for slides

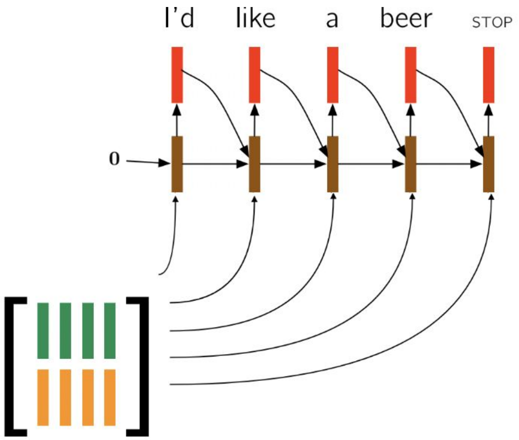
Announcements

- A3 is out on gitlab!
- Quiz 6 goes out on Canvas today at the end of lecture
  - Available until Friday 2/24 at 2:20pm; you'll have 15 minutes to complete it once you start.
  - Remember that you can use your notes during the quiz
  - Will cover material from (last) Wednesday's lecture and (last) Friday's lecture (so, transformers and machine translation)

Wrapping up machine translation

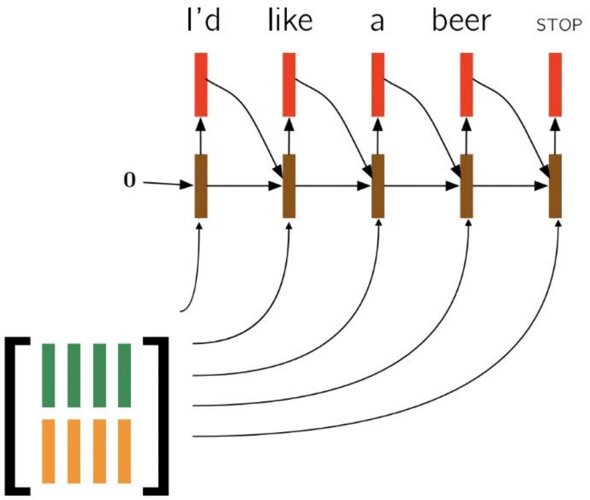
What's our loss function?

Just cross-entropy loss at each output timestep for correct output word (either pretending we've gotten all previous tokens in the output correct, or not– see this pytorch NMT tutorial for details).



Where might we find this kind of parallel data?

Note that this seq2seq approach can apply to more tasks than NMT!



Evaluating an MT system: Adequacy, fluency

In the past, these were usually separately computed elements of a score.

They’re still relevant today, but these days, they’re sort of both rolled into…

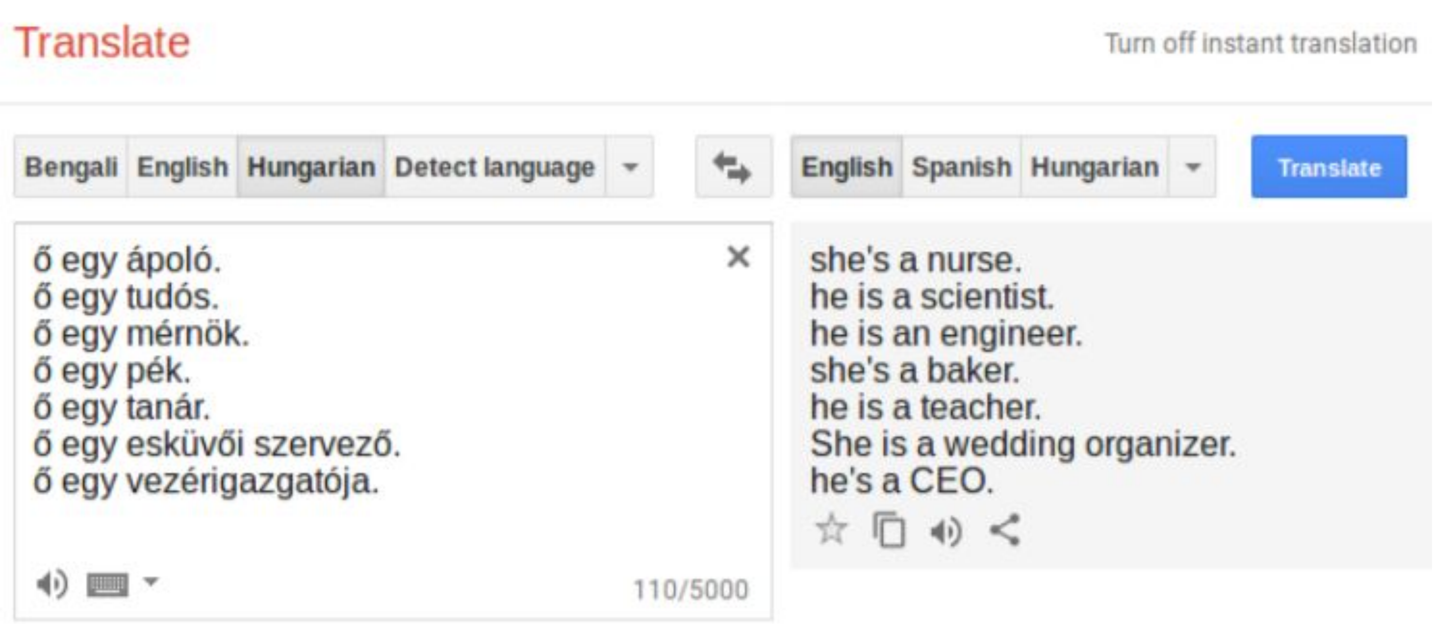
These days: BLEU score, mostly.

For each instance in your test set, have a set of human-written reference translations.

$$p_n = \frac{\text{number of } n\text{-grams appearing in both reference and hypothesis translations}}{\text{number of } n\text{-grams appearing in the hypothesis translation}}$$

Once you've got reference translations for a test set, pretty cheap to compute.

Evaluating bias in MT systems

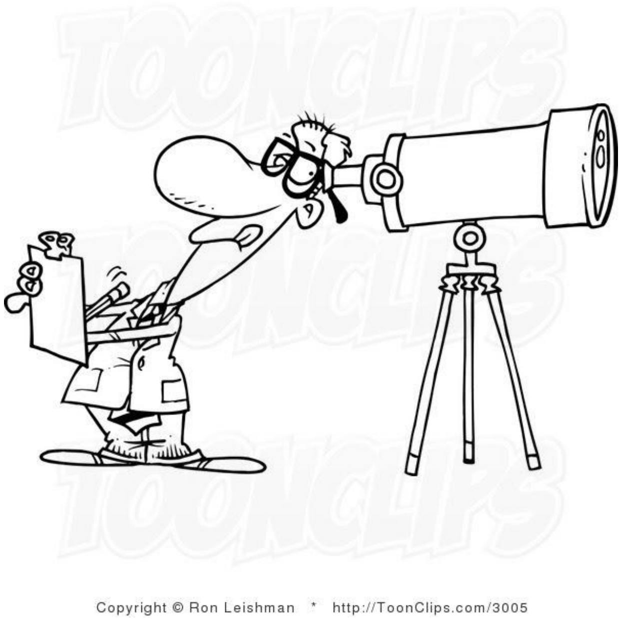


See, for example, [Stanovsky et al. 2019]

Syntax

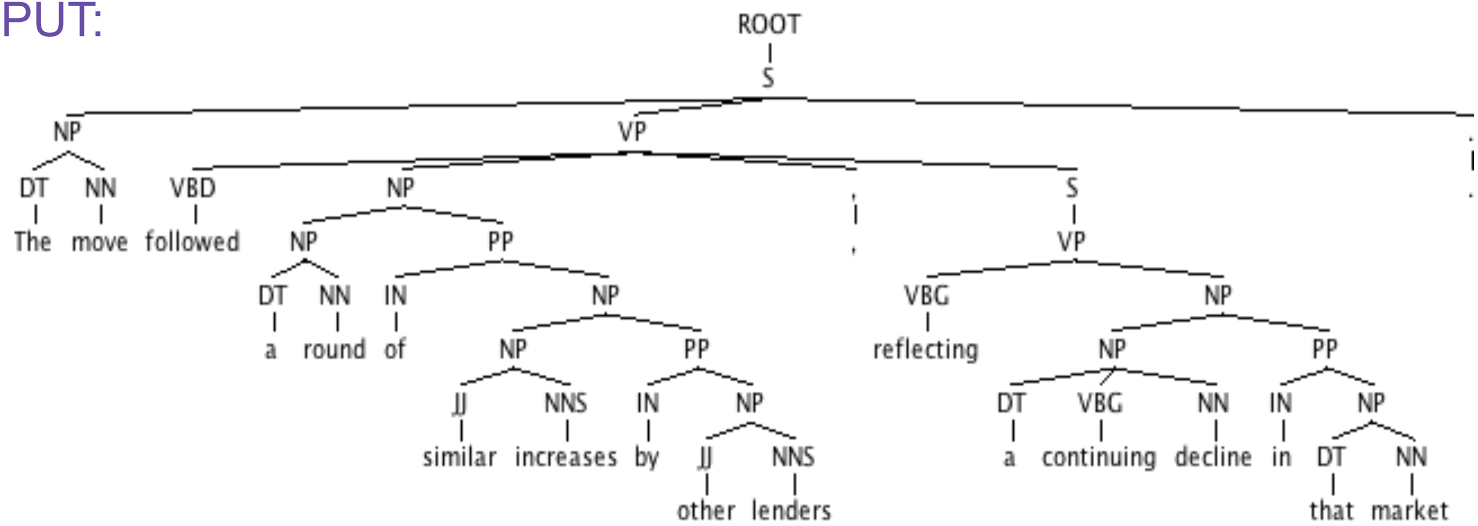
Ambiguity

- I saw a girl with a telescope



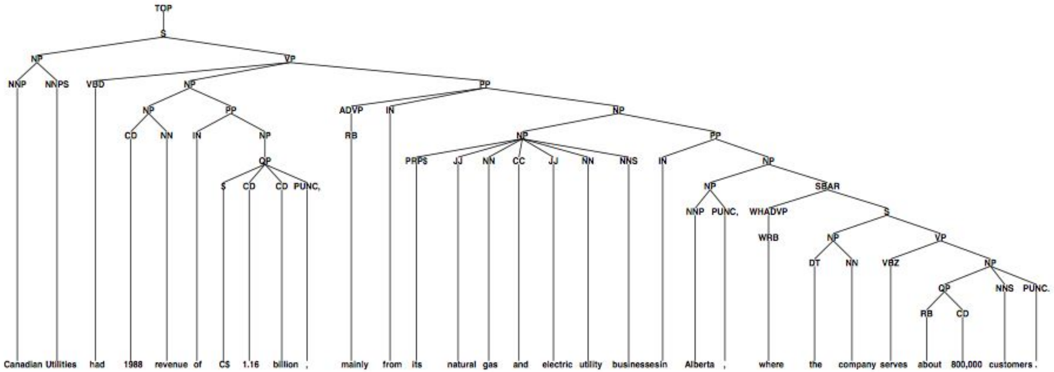
Syntactic Parsing

- INPUT:
  - The move followed a round of similar increases by other lenders, reflecting a continuing decline in that market
- OUTPUT:



A Supervised ML Problem

- Data for parsing experiments:
  - Penn WSJ Treebank = 50,000 sentences with associated trees
  - Usual set-up: 40,000 training, 2,400 test



Canadian Utilities had 1988 revenue of $ 1.16 billion , mainly from its natural gas and electric utility businesses in Alberta , where the company serves about 800,000 customers .

[from Michael Collins slides]

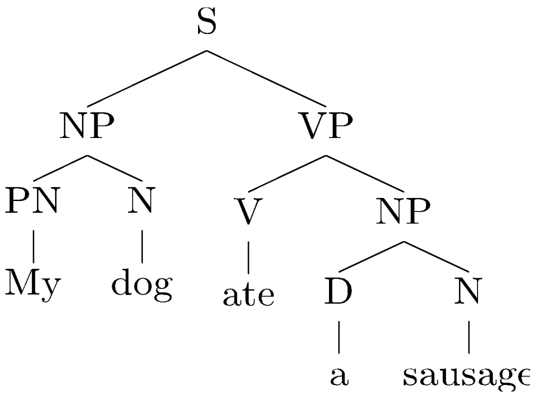
Syntax

- The study of the patterns of formation of sentences and phrases from words

○ my dog	Pron N
○ the dog	Det N
○ the cat	Det N
○ and	Conj
○ the large cat	Det Adj N
○ the black cat	Det Adj N
○ ate a sausage	V Det N

Parsing

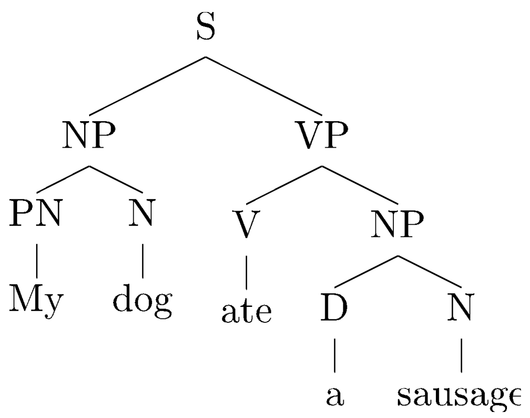
- The process of predicting syntactic representations
- Different types of syntactic representations are possible, for example:



Constituent (a.k.a. phrase-structure) tree

Constituent trees

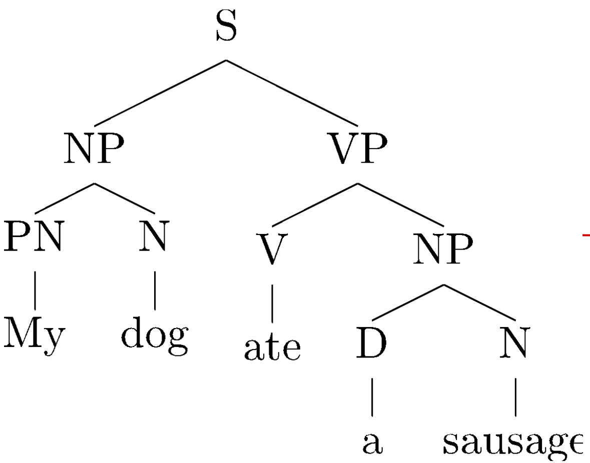
- Internal nodes correspond to phrases
  - S – a sentence
  - NP – Noun Phrase: My dog, a sandwich, lakes,..
  - VP – Verb Phrase: ate a sausage, barked, …
  - PP – Prepositional phrases: with a friend, in a car, …



- Nodes immediately above words are PoS tags (aka preterminals)
  - PN – pronoun
  - D – determiner
  - V – verb
  - N – noun
  - P – preposition

Bracketing notation

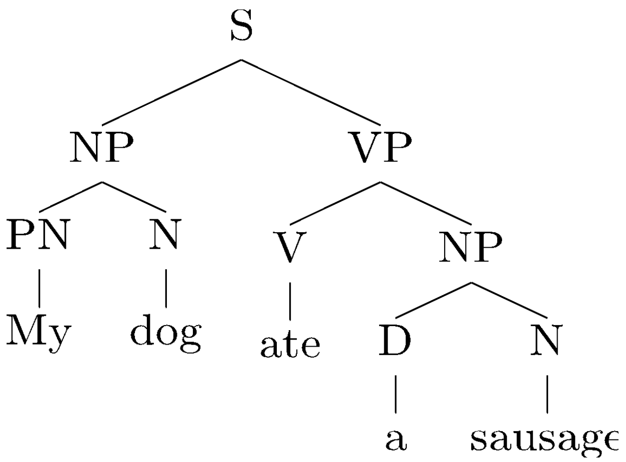
- It is often convenient to represent a tree as a bracketed sequence



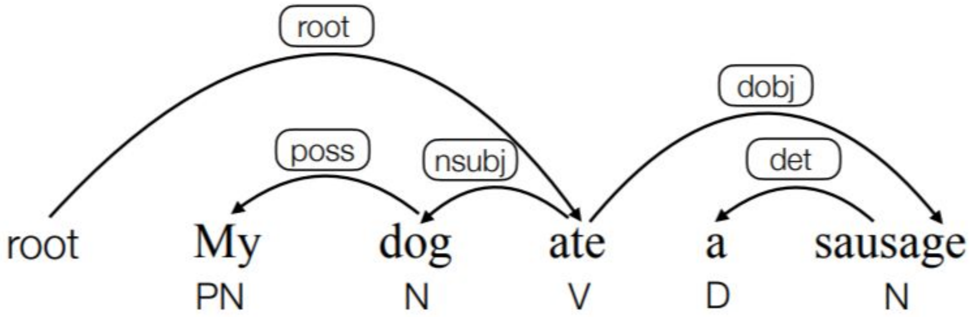

```
(S
   (NP (PN My) (N dog) )
   (VP (V ate)
      (NP (D a) (N sausage) )
   )
)
```

Parsing

- The process of predicting syntactic representations
- Different types of syntactic representations are possible, for example:

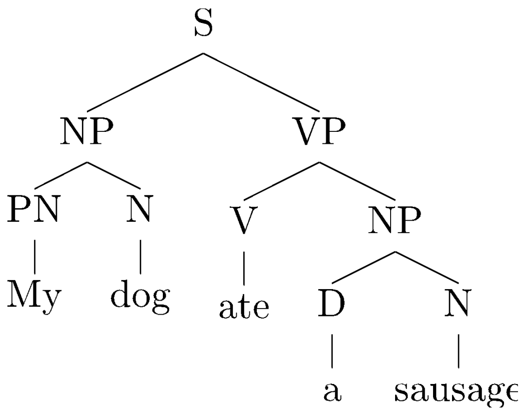


Constituent (a.k.a. phrase-structure) tree



Dependency tree

Constituent trees



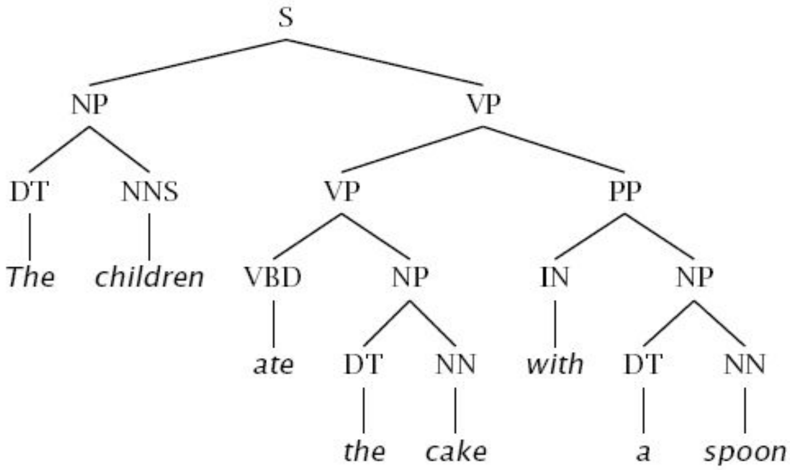
- Internal nodes correspond to phrases
  - S – a sentence
  - NP (Noun Phrase): My dog, a sandwich, lakes,..
  - VP (Verb Phrase): ate a sausage, barked, …
  - PP (Prepositional phrases): with a friend, in a car, …

- Nodes immediately above words are PoS tags (aka preterminals)
  - PN – pronoun
  - D – determiner
  - V – verb
  - N – noun
  - P – preposition

Constituency Tests

- How do we know what nodes go in the tree?

- Classic constituency tests:
  - Replacement
  - Movement
    - Passive
    - Clefting
    - Preposing
  - Substitution by *proform*
  - Modification
  - Coordination/Conjunction
  - Ellipsis/Deletion



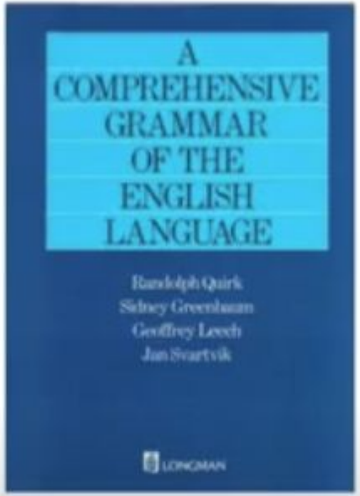
Morphology/Syntax/Semantics

- Syntax: The study of the patterns of formation of sentences and phrases from word
  - Borders with semantics and morphology sometimes blurred

Afyonkarahisarlılaştırabildiklerimizdenmişsinizcesinee

in Turkish means "as if you are one of the people that we thought to be originating from Afyonkarahisar" [wikipedia]

English grammar



Product Details (from Amazon)
Hardcover: 1779 pages
Publisher: Longman; 2nd Revised edition
Language: English
ISBN-10: 0582517346
ISBN-13: 978-0582517349
Product Dimensions: 8.4 x 2.4 x 10 inches
Shipping Weight: 4.6 pounds

Context Free Grammar (CFG)

Context Free Grammar (CFG)

Grammar (CFG)

ROOT $\rightarrow$ S
S $\rightarrow$ NP VP
NP $\rightarrow$ DT NN
NP $\rightarrow$ NN NNS
NP $\rightarrow$ NP PP
VP $\rightarrow$ VBP NP
VP $\rightarrow$ VBP NP PP
PP $\rightarrow$ IN NP

Lexicon

NN $\rightarrow$ interest
NNS $\rightarrow$ raises
VBP $\rightarrow$ interest
VBZ $\rightarrow$ raises
…

Other grammar formalisms: LFG (Lexical functional grammar), HPSG (Head-driven phrase structure grammar), TAG (Tree adjoining grammar), CCG (Combinatory categorial grammar)…

CFGs

S

$$\boxed{S \rightarrow NP \ VP}$$

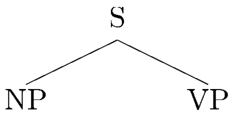
$$VP \rightarrow V$$
$$VP \rightarrow V \ NP$$
$$VP \rightarrow VP \ PP$$

$$NP \rightarrow NP \ PP$$
$$NP \rightarrow D \ N$$
$$NP \rightarrow PN$$

$$PP \rightarrow P \ NP$$

$$N \rightarrow girl$$
$$N \rightarrow telescope$$
$$N \rightarrow sandwich$$
$$PN \rightarrow I$$
$$V \rightarrow saw$$
$$V \rightarrow ate$$
$$P \rightarrow with$$
$$P \rightarrow in$$
$$D \rightarrow a$$
$$D \rightarrow the$$

CFGs



$$S \rightarrow NP \;\; VP$$

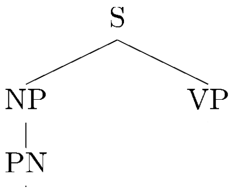
$$VP \rightarrow V$$
$$VP \rightarrow V \;\; NP$$
$$VP \rightarrow VP \;\; PP$$

$$NP \rightarrow NP \;\; PP$$
$$NP \rightarrow D \;\; N$$
$$\boxed{NP \rightarrow PN}$$

$$PP \rightarrow P \;\; NP$$

$$N \rightarrow girl$$
$$N \rightarrow telescope$$
$$N \rightarrow sandwich$$
$$PN \rightarrow I$$
$$V \rightarrow saw$$
$$V \rightarrow ate$$
$$P \rightarrow with$$
$$P \rightarrow in$$
$$D \rightarrow a$$
$$D \rightarrow the$$

CFGs



$$S \rightarrow NP \ \ VP$$

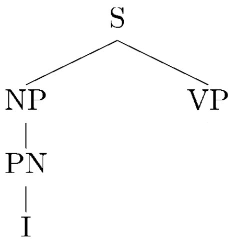
$$VP \rightarrow V$$
$$VP \rightarrow V \ \ NP$$
$$VP \rightarrow VP \ \ PP$$

$$NP \rightarrow NP \ \ PP$$
$$NP \rightarrow D \ \ N$$
$$NP \rightarrow PN$$

$$PP \rightarrow P \ \ NP$$

$$N \rightarrow girl$$
$$N \rightarrow telescope$$
$$N \rightarrow sandwich$$
$$\boxed{PN \rightarrow I}$$
$$V \rightarrow saw$$
$$V \rightarrow ate$$
$$P \rightarrow with$$
$$P \rightarrow in$$
$$D \rightarrow a$$
$$D \rightarrow the$$

CFGs



$$S \rightarrow NP \;\; VP$$

$$VP \rightarrow V$$

$$\boxed{VP \rightarrow V \;\; NP}$$

$$VP \rightarrow VP \;\; PP$$

$$NP \rightarrow NP \;\; PP$$

$$NP \rightarrow D \;\; N$$

$$NP \rightarrow PN$$

$$PP \rightarrow P \;\; NP$$

$$N \rightarrow girl$$

$$N \rightarrow telescope$$

$$N \rightarrow sandwich$$

$$PN \rightarrow I$$

$$V \rightarrow saw$$

$$V \rightarrow ate$$

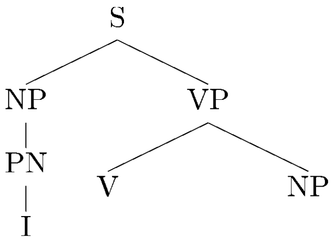
$$P \rightarrow with$$

$$P \rightarrow in$$

$$D \rightarrow a$$

$$D \rightarrow the$$

CFGs



$$S \rightarrow NP\ \ VP$$

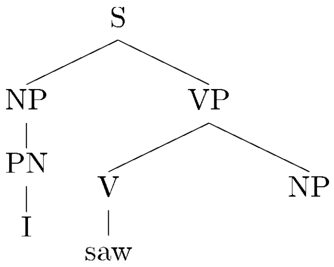
$$VP \rightarrow V$$
$$VP \rightarrow V\ \ NP$$
$$VP \rightarrow VP\ \ PP$$

$$NP \rightarrow NP\ \ PP$$
$$NP \rightarrow D\ \ N$$
$$NP \rightarrow PN$$

$$PP \rightarrow P\ \ NP$$

$$N \rightarrow girl$$
$$N \rightarrow telescope$$
$$N \rightarrow sandwich$$
$$PN \rightarrow I$$
$$V \rightarrow saw$$
$$V \rightarrow ate$$
$$P \rightarrow with$$
$$P \rightarrow in$$
$$D \rightarrow a$$
$$D \rightarrow the$$

CFGs



$$S \rightarrow NP\ \ VP$$

$$VP \rightarrow V$$

$$VP \rightarrow V\ \ NP$$

$$VP \rightarrow VP\ \ PP$$

$$\boxed{NP \rightarrow NP\ \ PP}$$

$$NP \rightarrow D\ \ N$$

$$NP \rightarrow PN$$

$$PP \rightarrow P\ \ NP$$

$$N \rightarrow girl$$

$$N \rightarrow telescope$$

$$N \rightarrow sandwich$$

$$PN \rightarrow I$$

$$V \rightarrow saw$$

$$V \rightarrow ate$$

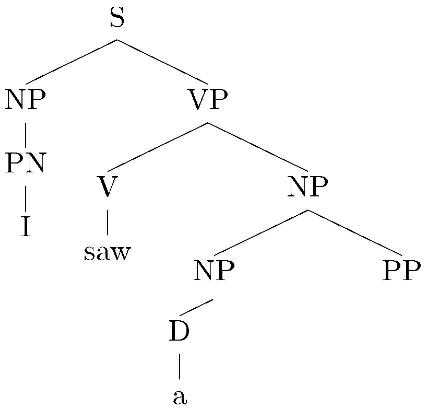
$$P \rightarrow with$$

$$P \rightarrow in$$

$$D \rightarrow a$$

$$D \rightarrow the$$

CFGs



$$S \rightarrow NP \;\; VP$$

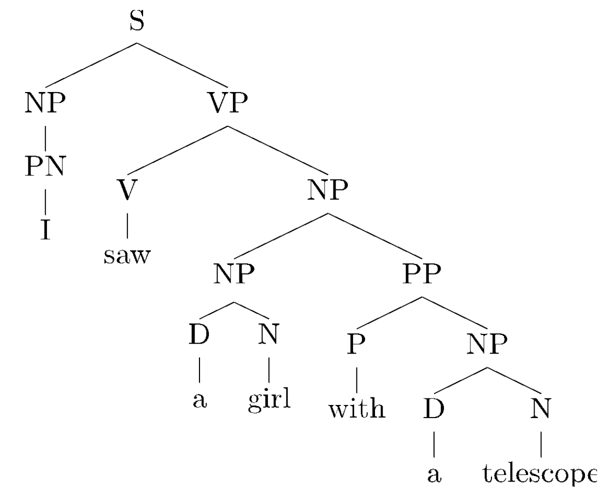
$$VP \rightarrow V$$
$$VP \rightarrow V \;\; NP$$
$$VP \rightarrow VP \;\; PP$$

$$NP \rightarrow NP \;\; PP$$
$$\boxed{NP \rightarrow D \;\; N}$$
$$NP \rightarrow PN$$

$$PP \rightarrow P \;\; NP$$

$$N \rightarrow girl$$
$$N \rightarrow telescope$$
$$N \rightarrow sandwich$$
$$PN \rightarrow I$$
$$V \rightarrow saw$$
$$V \rightarrow ate$$
$$P \rightarrow with$$
$$P \rightarrow in$$
$$D \rightarrow a$$
$$D \rightarrow the$$

CFGs



$$S \rightarrow NP\ \ VP$$

$$VP \rightarrow V$$
$$VP \rightarrow V\ \ NP$$
$$VP \rightarrow VP\ \ PP$$

$$NP \rightarrow NP\ \ PP$$
$$NP \rightarrow D\ \ N$$
$$NP \rightarrow PN$$

$$PP \rightarrow P\ \ NP$$

$$N \rightarrow girl$$
$$N \rightarrow telescope$$
$$N \rightarrow sandwich$$
$$PN \rightarrow I$$
$$V \rightarrow saw$$
$$V \rightarrow ate$$
$$P \rightarrow with$$
$$P \rightarrow in$$
$$D \rightarrow a$$
$$D \rightarrow the$$

Treebank Sentences

```
( (S (NP-SBJ The move)
     (VP followed
         (NP (NP a round)
             (PP of
                 (NP (NP similar increases)
                     (PP by
                         (NP other lenders))
                     (PP against
                         (NP Arizona real estate loans)))))
         ,
         (S-ADV (NP-SBJ *)
                (VP reflecting
                    (NP (NP a continuing decline)
                        (PP-LOC in
                                (NP that market))))))
     .))
```

Context-Free Grammars

- A context-free grammar is a 4-tuple <N, T, S, R>
  - N : the set of non-terminals
    - **Phrasal categories**: S, NP, VP, ADJP, etc.
    - **Parts-of-speech** (pre-terminals): NN, JJ, DT, VB
  - T : the set of terminals (the words)
  - S : the start symbol
    - Often written as ROOT or TOP
    - Not usually the sentence non-terminal S
  - R : the set of rules
    - Of the form $X \rightarrow Y_1\ Y_2 \ldots Y_k$, with $X, Y_i \in N$
    - Examples: $S \rightarrow NP\ VP$, $VP \rightarrow VP\ CC\ VP$
    - Also called rewrites, productions, or local trees

An example grammar

$N = \{S, VP, NP, PP, N, V, PN, P\}$

$T = \{girl, telescope, sandwich, I, saw, ate, with, in, a, the\}$

$S = \{S\}$

$R:$

Called **Inner rules**

$S \rightarrow NP \;\; VP$ (NP A girl) (VP ate a sandwich)

$VP \rightarrow V$

$VP \rightarrow V \;\; NP$ (V ate) (NP a sandwich)

$VP \rightarrow VP \;\; PP$ (VP saw a girl) (PP with a telescope)

$NP \rightarrow NP \;\; PP$ (NP a girl) (PP with a sandwich)

$NP \rightarrow D \;\; N$ (D a) (N sandwich)

$NP \rightarrow PN$

$PP \rightarrow P \;\; NP$ (P with) (NP with a sandwich)

Preterminal rules

$N \rightarrow girl$

$N \rightarrow telescope$

$N \rightarrow sandwich$

$PN \rightarrow I$

$V \rightarrow saw$

$V \rightarrow ate$

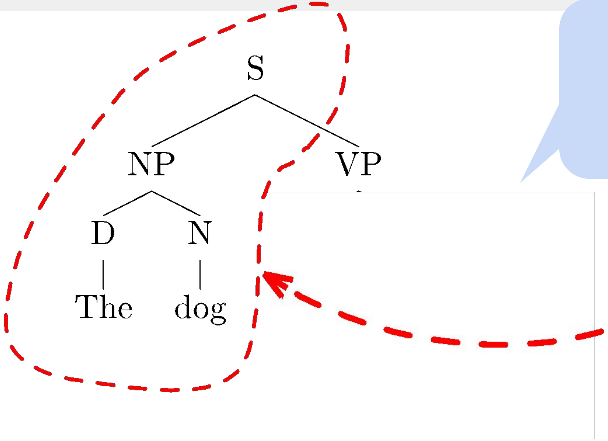
$P \rightarrow with$

$P \rightarrow in$

$D \rightarrow a$

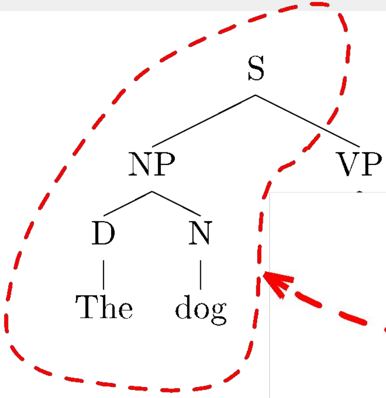
$D \rightarrow the$

Why context-free?

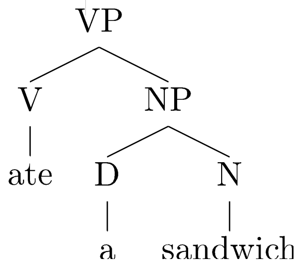


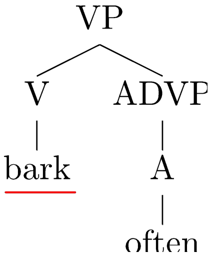
What can be a sub-tree is only affected by what the phrase type is (VP) but not the context

Why context-free?



What can be a sub-tree is only affected by what the phrase type is (VP) but not the context



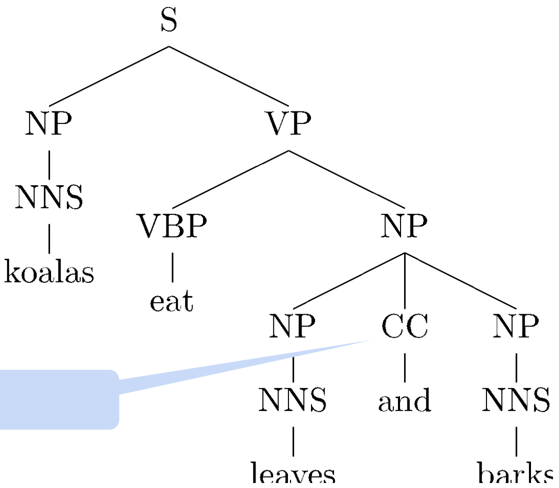


Not grammatical

Ambiguities

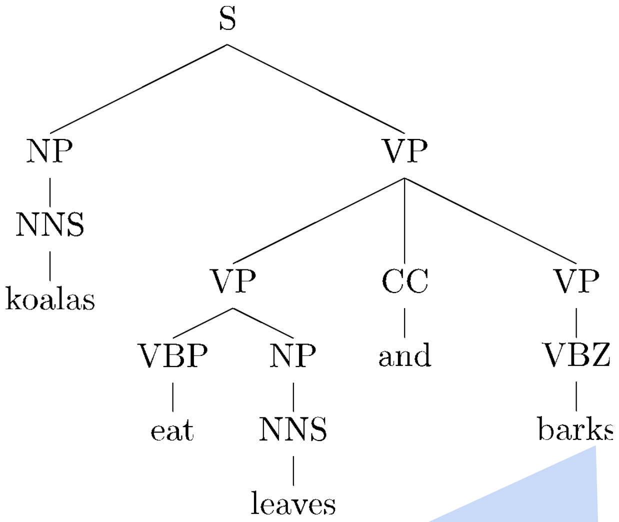
Coordination ambiguity

- Here, the coarse VP and NP categories cannot enforce subject-verb agreement in number resulting in the coordination ambiguity



Coordination

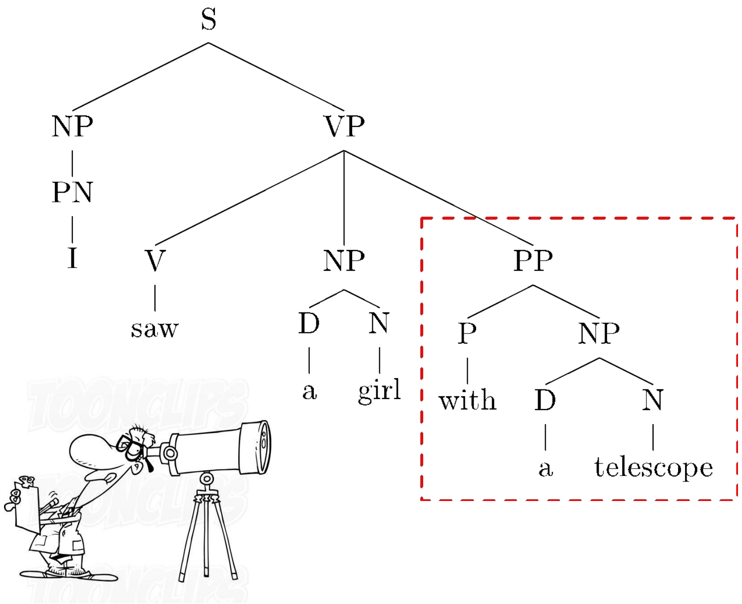
"Bark" can refer both to a noun or a verb

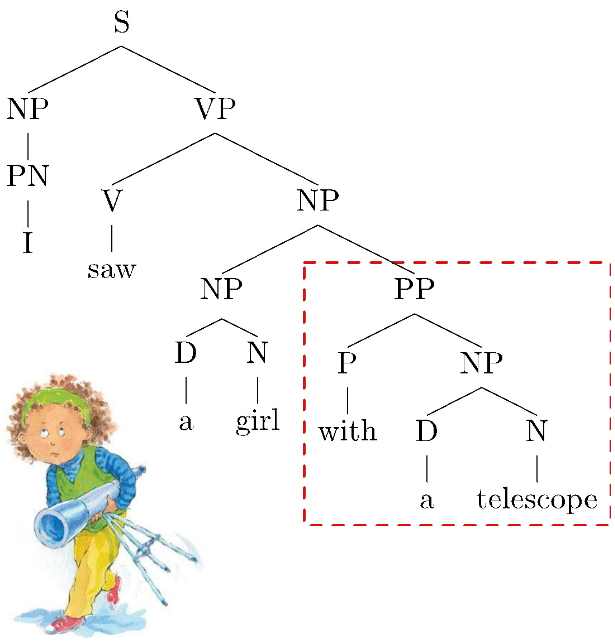


This tree would be ruled out if the context would be somehow captured (subject-verb agreement)

Why is parsing hard?   Ambiguity

- Prepositional phrase attachment ambiguity





PP Ambiguity

Put the block in the box on the table in the kitchen

3 prepositional phrases, 5 interpretations:

- Put the block ((in the box on the table) in the kitchen)
- Put the block (in the box (on the table in the kitchen))
- Put ((the block in the box) on the table) in the kitchen.
- Put (the block (in the box on the table)) in the kitchen.
- Put (the block in the box) (on the table in the kitchen)

PP Ambiguity

Put the block in the box on the table in the kitchen

3 prepositional phrases, 5 interpretations:

- Put the block ((in the box on the table) in the kitchen)
- Put the block (in the box (on the table in the kitchen))
- ...

A general case:

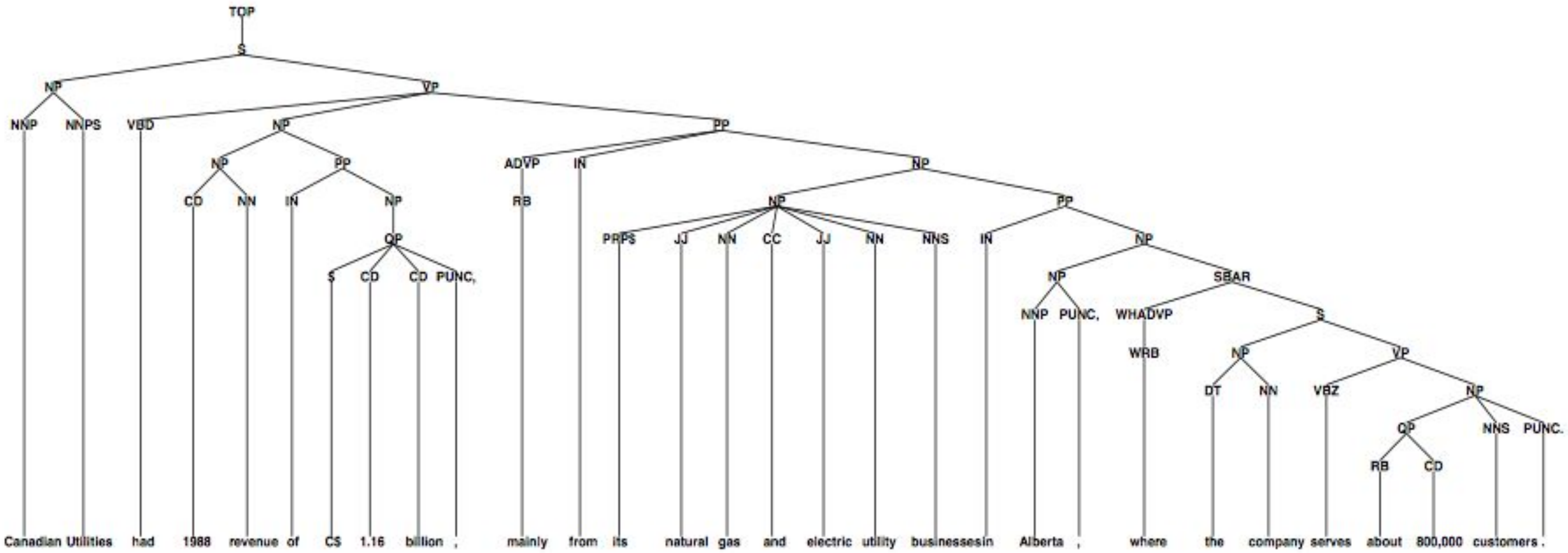
- ((())) ()(()) ()()() (())() (()())

$$Cat_n = \binom{2n}{n} - \binom{2n}{n-1} \sim \frac{4^n}{n^{3/2}\sqrt{\pi}}$$

Catalan numbers

$$1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, \ldots$$

A typical tree from a standard dataset (Penn treebank WSJ)



Canadian Utilities had 1988 revenue of $ 1.16 billion , mainly from its natural gas and electric utility businesses in Alberta , where the company serves about 800,000 customers .

[from Michael Collins slides]

Syntactic Ambiguities I

- Prepositional phrases:
  - They cooked the beans in the pot on the stove with handles.

- Particle vs. preposition:
  - The puppy tore up the staircase.

- Complement structures
  - The tourists objected to the guide that they couldn't hear.
    She knows you like the back of her hand.

- Gerund vs. participial adjective
  - Visiting relatives can be boring.
    Changing schedules frequently confused passengers.

Syntactic Ambiguities II

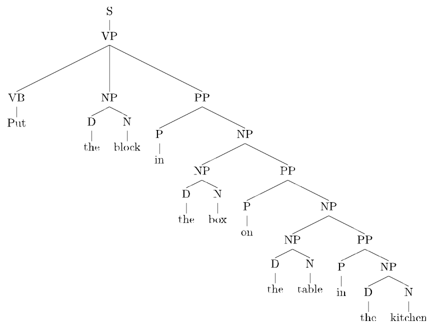
- Modifier scope within NPs
  - impractical design requirements
    plastic cup holder

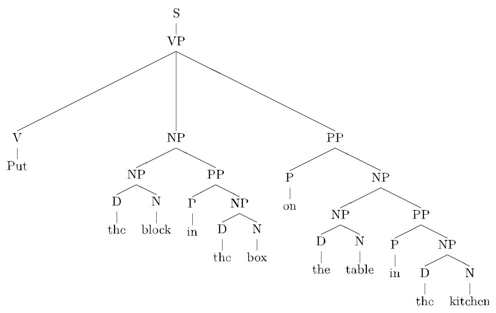
- Multiple gap constructions
  - The chicken is ready to eat.
    The contractors are rich enough to sue.

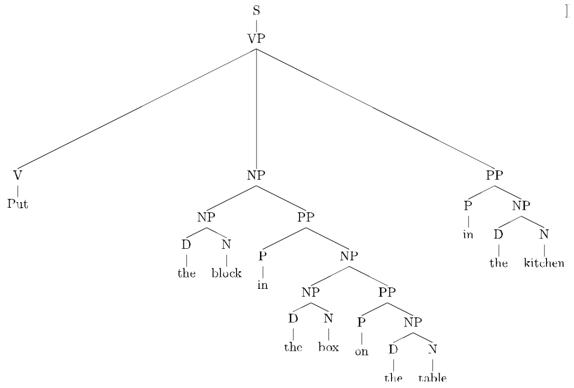
- Coordination scope:
  - Small rats and mice can squeeze into holes or cracks in the wall.

How to Deal with Ambiguity?

- We want to score all the derivations to encode how plausible they are







Put the block in the box on the table in the kitchen

Probabilistic Context Free Grammar (PCFG)

Probabilistic Context-Free Grammars

- A context-free grammar is a 4-tuple <N, T, S, R>
    - N : the set of non-terminals
        - **Phrasal categories**: S, NP, VP, ADJP, etc.
        - **Parts-of-speech** (pre-terminals): NN, JJ, DT, VB
    - T : the set of terminals (the words)
    - S : the start symbol
        - Often written as ROOT or TOP
        - Not usually the sentence non-terminal S
    - R : the set of rules
        - Of the form $X \rightarrow Y_1\ Y_2 \ldots Y_k$, with $X, Y_i \in N$
        - Examples: $S \rightarrow NP\ VP$, $VP \rightarrow VP\ CC\ VP$
        - Also called rewrites, productions, or local trees
- A PCFG adds:
    - A top-down production probability per rule $P(Y_1\ Y_2 \ldots Y_k \mid X)$

PCFGs

Associate probabilities with the rules : $p(X \rightarrow \alpha)$

$$\forall\ X \rightarrow \alpha \in R: \quad 0 \leq p(X \rightarrow \alpha) \leq 1$$
$$\forall X \in N: \quad \sum_{\alpha: X \rightarrow \alpha \in R} p(X \rightarrow \alpha) = 1$$

Now we can score a tree as a product of probabilities corresponding to the used rules

$S \rightarrow NP\ \ VP$	1.0	(NP A girl) (VP ate a sandwich)
$VP \rightarrow V$	0.2	
$VP \rightarrow V\ \ NP$	0.4	(VP ate) (NP a sandwich)
$VP \rightarrow VP\ \ PP$	0.4	(VP saw a girl) (PP with …)
$NP \rightarrow NP\ \ PP$	0.3	(NP a girl) (PP with ….)
$NP \rightarrow D\ \ N$	0.5	(D a) (N sandwich)
$NP \rightarrow PN$	0.2	
$PP \rightarrow P\ \ NP$	1.0	(P with) (NP with a sandwich)

$N \rightarrow girl$	0.2
$N \rightarrow telescope$	0.7
$N \rightarrow sandwich$	0.1
$PN \rightarrow I$	1.0
$V \rightarrow saw$	0.5
$V \rightarrow ate$	0.5
$P \rightarrow with$	0.6
$P \rightarrow in$	0.4
$D \rightarrow a$	0.3
$D \rightarrow the$	0.7

PCFGs

S

$$S \rightarrow NP \;\; VP \;\; \text{1.0}$$

$$VP \rightarrow V \;\; \text{0.2}$$

$$VP \rightarrow V \;\; NP \;\; \text{0.4}$$

$$VP \rightarrow VP \;\; PP \;\; \text{0.4}$$

$$NP \rightarrow NP \;\; PP \;\; \text{0.3}$$

$$NP \rightarrow D \;\; N \;\; \text{0.5}$$

$$NP \rightarrow PN \;\; \text{0.2}$$

$$PP \rightarrow P \;\; NP \;\; \text{1.0}$$

$$N \rightarrow girl \;\; \text{0.2}$$

$$N \rightarrow telescope \;\; \text{0.7}$$

$$N \rightarrow sandwich \;\; \text{0.1}$$

$$PN \rightarrow I \;\; \text{1.0}$$

$$V \rightarrow saw \;\; \text{0.5}$$

$$V \rightarrow ate \;\; \text{0.5}$$

$$P \rightarrow with \;\; \text{0.6}$$

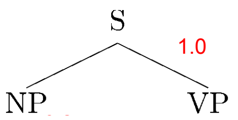
$$P \rightarrow in \;\; \text{0.4}$$

$$D \rightarrow a \;\; \text{0.3}$$

$$D \rightarrow the \;\; \text{0.7}$$

$$p(T) =$$

PCFGs



$$\boxed{S \rightarrow NP \;\; VP \;\; 1.0}$$

$$VP \rightarrow V \;\; 0.2$$
$$VP \rightarrow V \;\; NP \;\; 0.4$$
$$VP \rightarrow VP \;\; PP \;\; 0.4$$

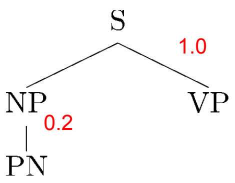
$$NP \rightarrow NP \;\; PP \;\; 0.3$$
$$NP \rightarrow D \;\; N \;\; 0.5$$
$$NP \rightarrow PN \;\; 0.2$$

$$PP \rightarrow P \;\; NP \;\; 1.0$$

$$N \rightarrow girl \;\; 0.2$$
$$N \rightarrow telescope \;\; 0.7$$
$$N \rightarrow sandwich \;\; 0.1$$
$$PN \rightarrow I \;\; 1.0$$
$$V \rightarrow saw \;\; 0.5$$
$$V \rightarrow ate \;\; 0.5$$
$$P \rightarrow with \;\; 0.6$$
$$P \rightarrow in \;\; 0.4$$
$$D \rightarrow a \;\; 0.3$$
$$D \rightarrow the \;\; 0.7$$

$$p(T) = 1.0 \times$$

PCFGs



$S \rightarrow NP \;\; VP$ 1.0

$VP \rightarrow V$ 0.2

$VP \rightarrow V \;\; NP$ 0.4

$VP \rightarrow VP \;\; PP$ 0.4

$NP \rightarrow NP \;\; PP$ 0.3

$NP \rightarrow D \;\; N$ 0.5

$\boxed{NP \rightarrow PN \;\; 0.2}$

$PP \rightarrow P \;\; NP$ 1.0

$N \rightarrow girl$ 0.2

$N \rightarrow telescope$ 0.7

$N \rightarrow sandwich$ 0.1

$PN \rightarrow I$ 1.0

$V \rightarrow saw$ 0.5

$V \rightarrow ate$ 0.5

$P \rightarrow with$ 0.6

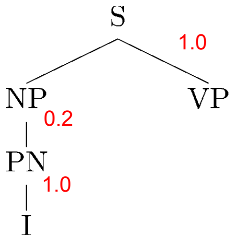
$P \rightarrow in$ 0.4

$D \rightarrow a$ 0.3

$D \rightarrow the$ 0.7

$$p(T) = 1.0 \times 0.2 \times$$

PCFGs



$$S \rightarrow NP \;\; VP \; \textcolor{red}{1.0}$$

$$VP \rightarrow V \; \textcolor{red}{0.2}$$
$$VP \rightarrow V \;\; NP \; \textcolor{red}{0.4}$$
$$VP \rightarrow VP \;\; PP \; \textcolor{red}{0.4}$$

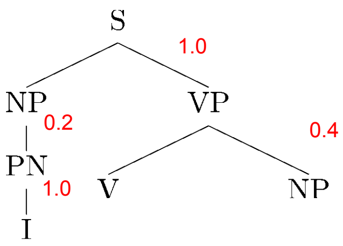
$$NP \rightarrow NP \;\; PP \; \textcolor{red}{0.3}$$
$$NP \rightarrow D \;\; N \; \textcolor{red}{0.5}$$
$$NP \rightarrow PN \; \textcolor{red}{0.2}$$

$$PP \rightarrow P \;\; NP \; \textcolor{red}{1.0}$$

$$N \rightarrow girl \; \textcolor{red}{0.2}$$
$$N \rightarrow telescope \; \textcolor{red}{0.7}$$
$$N \rightarrow sandwich \; \textcolor{red}{0.1}$$
$$\boxed{PN \rightarrow I \; \textcolor{red}{1.0}}$$
$$V \rightarrow saw \; \textcolor{red}{0.5}$$
$$V \rightarrow ate \; \textcolor{red}{0.5}$$
$$P \rightarrow with \; \textcolor{red}{0.6}$$
$$P \rightarrow in \; \textcolor{red}{0.4}$$
$$D \rightarrow a \; \textcolor{red}{0.3}$$
$$D \rightarrow the \; \textcolor{red}{0.7}$$

$$p(T) = 1.0 \times 0.2 \times 1.0 \times$$

PCFGs



$S \rightarrow NP \;\; VP$ 1.0

$VP \rightarrow V$ 0.2

$\boxed{VP \rightarrow V \;\; NP \; 0.4}$

$VP \rightarrow VP \;\; PP$ 0.4

$NP \rightarrow NP \;\; PP$ 0.3

$NP \rightarrow D \;\; N$ 0.5

$NP \rightarrow PN$ 0.2

$PP \rightarrow P \;\; NP$ 1.0

$N \rightarrow girl$ 0.2

$N \rightarrow telescope$ 0.7

$N \rightarrow sandwich$ 0.1

$PN \rightarrow I$ 1.0

$V \rightarrow saw$ 0.5

$V \rightarrow ate$ 0.5

$P \rightarrow with$ 0.6

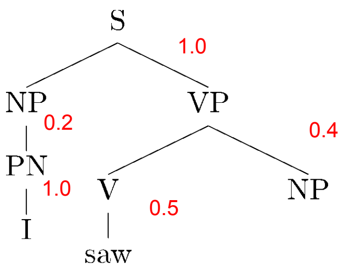
$P \rightarrow in$ 0.4

$D \rightarrow a$ 0.3

$D \rightarrow the$ 0.7

$$p(T) = 1.0 \times 0.2 \times 1.0 \times 0.4 \times$$

PCFGs



$$S \rightarrow NP\ \ VP\ \ 1.0$$

$$VP \rightarrow V\ \ 0.2$$
$$VP \rightarrow V\ \ NP\ \ 0.4$$
$$VP \rightarrow VP\ \ PP\ \ 0.4$$

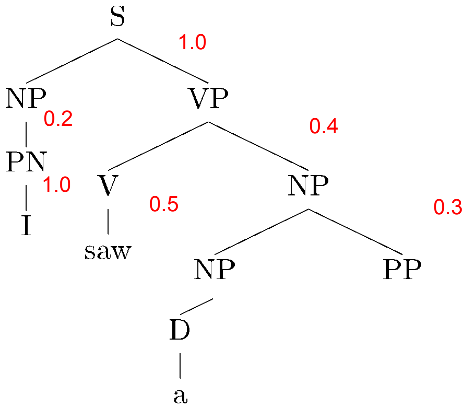
$$NP \rightarrow NP\ \ PP\ \ 0.3$$
$$NP \rightarrow D\ \ N\ \ 0.5$$
$$NP \rightarrow PN\ \ 0.2$$

$$PP \rightarrow P\ \ NP\ \ 1.0$$

$$N \rightarrow girl\ \ 0.2$$
$$N \rightarrow telescope\ \ 0.7$$
$$N \rightarrow sandwich\ \ 0.1$$
$$PN \rightarrow I\ \ 1.0$$
$$\boxed{V \rightarrow saw\ \ 0.5}$$
$$V \rightarrow ate\ \ 0.5$$
$$P \rightarrow with\ \ 0.6$$
$$P \rightarrow in\ \ 0.4$$
$$D \rightarrow a\ \ 0.3$$
$$D \rightarrow the\ \ 0.7$$

$$p(T) = 1.0 \times 0.2 \times 1.0 \times 0.4 \times 0.5 \times$$

PCFGs



$S \rightarrow NP \ \ VP$ 1.0

$VP \rightarrow V$ 0.2

$VP \rightarrow V \ \ NP$ 0.4

$VP \rightarrow VP \ \ PP$ 0.4

$\boxed{NP \rightarrow NP \ \ PP \ \ 0.3}$

$NP \rightarrow D \ \ N$ 0.5

$NP \rightarrow PN$ 0.2

$PP \rightarrow P \ \ NP$ 1.0

$N \rightarrow girl$ 0.2

$N \rightarrow telescope$ 0.7

$N \rightarrow sandwich$ 0.1

$PN \rightarrow I$ 1.0

$V \rightarrow saw$ 0.5

$V \rightarrow ate$ 0.5

$P \rightarrow with$ 0.6

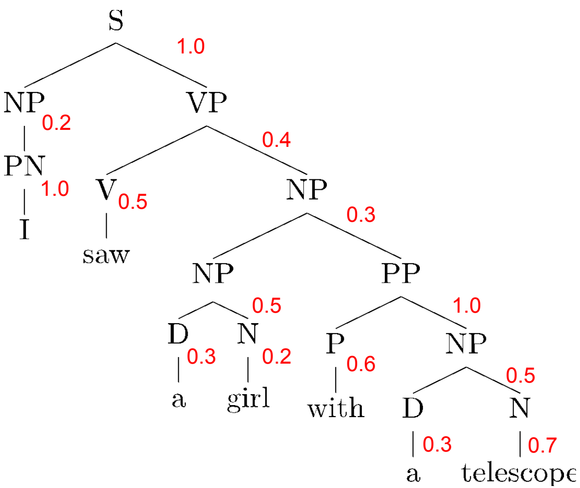
$P \rightarrow in$ 0.4

$D \rightarrow a$ 0.3

$D \rightarrow the$ 0.7

$$p(T) = 1.0 \times 0.2 \times 1.0 \times 0.4 \times 0.5 \times 0.3 \times$$

PCFGs



$$S \to NP \;\; VP \;\; 1.0$$

$$VP \to V \;\; 0.2$$
$$VP \to V \;\; NP \;\; 0.4$$
$$VP \to VP \;\; PP \;\; 0.4$$

$$NP \to NP \;\; PP \;\; 0.3$$
$$NP \to D \;\; N \;\; 0.5$$
$$NP \to PN \;\; 0.2$$

$$PP \to P \;\; NP \;\; 1.0$$

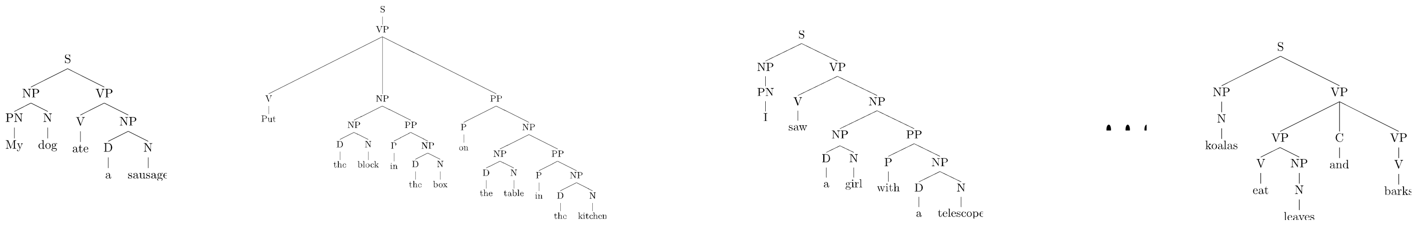
$$N \to girl \;\; 0.2$$
$$N \to telescope \;\; 0.7$$
$$N \to sandwich \;\; 0.1$$
$$PN \to I \;\; 1.0$$
$$V \to saw \;\; 0.5$$
$$V \to ate \;\; 0.5$$
$$P \to with \;\; 0.6$$
$$P \to in \;\; 0.4$$
$$D \to a \;\; 0.3$$
$$D \to the \;\; 0.7$$

$$p(T) = 1.0 \times 0.2 \times 1.0 \times 0.4 \times 0.5 \times 0.3 \times 0.5 \times 0.3 \times 0.2 \times 1.0 \times 0.6 \times 0.5 \times 0.3 \times 0.7 = 2.26 \times 10^{-5}$$

PCFG Estimation

Maximum likelihood estimation

- A treebank: a collection sentences annotated with constituent trees



- An estimated probability of a rule (maximum likelihood estimates)

$$p(X \to \alpha) = \frac{C(X \to \alpha)}{C(X)}$$

The number of times the rule used in the corpus

The number of times the nonterminal X appears in the treebank

- Smoothing is helpful
  - Especially important for preterminal rules

Parsing evaluation

Parsing evaluation

- Intrinsic evaluation:
  - Automatic: evaluate against annotation provided by human experts (gold standard) according to some predefined measure
  - Manual: ... according to human judgment

- Extrinsic evaluation: score syntactic representation by comparing how well a system using this representation performs on some task
  - E.g., use syntactic representation as input for a semantic analyzer and compare results of the analyzer using syntax predicted by different parsers.

Standard evaluation setting in parsing

- Automatic intrinsic evaluation is used: parsers are evaluated against gold standard by provided by linguists
  - There is a standard split into the parts:
    - training set: used for estimation of model parameters
    - development set: used for tuning the model (initial experiments)
    - test set: final experiments to compare against previous work

Automatic evaluation of constituent parsers

- Exact match: percentage of trees predicted correctly
- Bracket score: scores how well individual phrases (and their boundaries) are identified

The most standard measure; we will focus on it

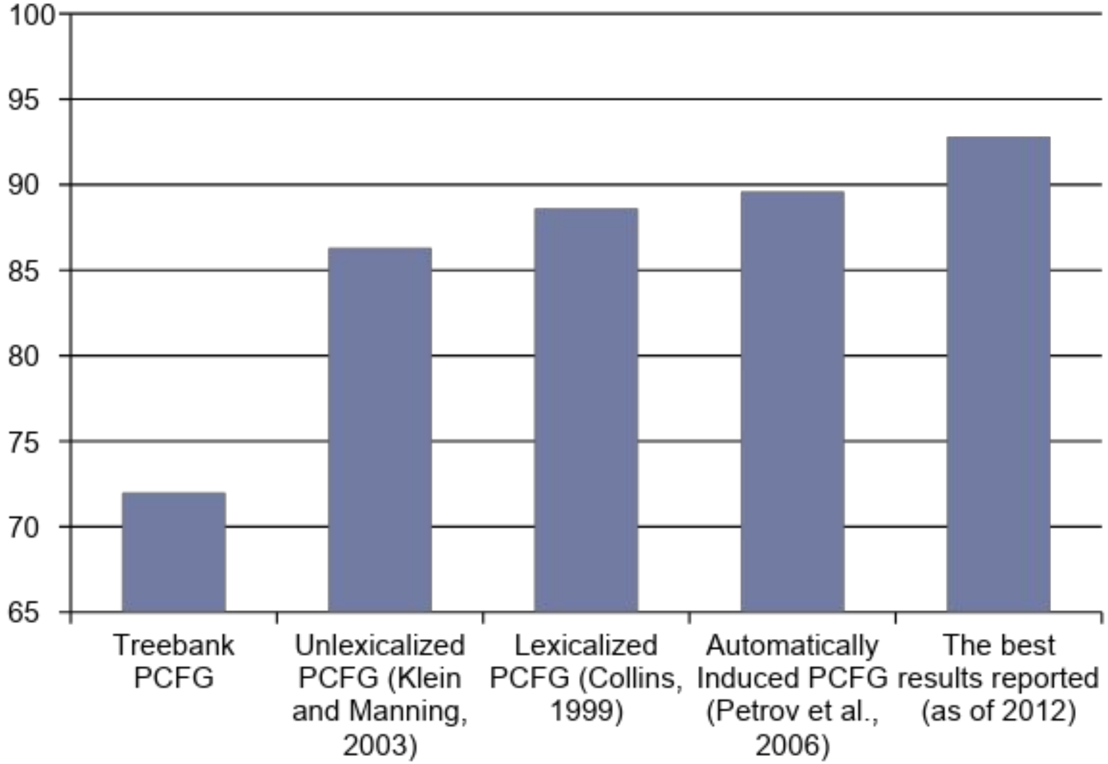
Brackets scores

Subtree signatures for CKY

$$[min, max, C]$$

- The most standard score is bracket score
- It regards a tree as a collection of brackets:
- The set of brackets predicted by a parser is compared against the set of brackets in the tree annotated by a linguist
- Precision, recall and F1 are used as scores

Preview: F1 bracket score



CKY Parsing

Parsing

- Parsing is search through the space of all possible parses
  - e.g., we may want either any parse, all parses or the highest scoring parse (if PCFG):

$$\arg\max_{T \in G(x)} P(T)$$

- Bottom-up:
  - One starts from words and attempt to construct the full tree

- Top-down
  - Start from the start symbol and attempt to expand to get the sentence

CKY algorithm (aka CYK)

- Cocke-Kasami-Younger algorithm
  - Independently discovered in late 60s / early 70s

- An efficient bottom up parsing algorithm for (P)CFGs
  - can be used both for the recognition and parsing problems
  - Very important in NLP (and beyond)

- We will start with the non-probabilistic version

Constraints on the grammar

- The basic CKY algorithm supports only rules in the Chomsky Normal Form (CNF):

$$C \rightarrow x$$

$$C \rightarrow C_1 C_2$$

Unary preterminal rules (generation of words given PoS tags) $N \rightarrow telescope$ $D \rightarrow the$

Binary inner rules $S \rightarrow NP\ VP$ $NP \rightarrow D\ N$

Constraints on the grammar

- The basic CKY algorithm supports only rules in the Chomsky Normal Form (CNF):

$$C \to x$$

$$C \to C_1 C_2$$

- Any CFG can be converted to an equivalent CNF
  - Equivalent means that they define the same language
  - However (syntactic) trees will look differently
  - It is possible to address it by defining such transformations that allows for easy reverse transformation

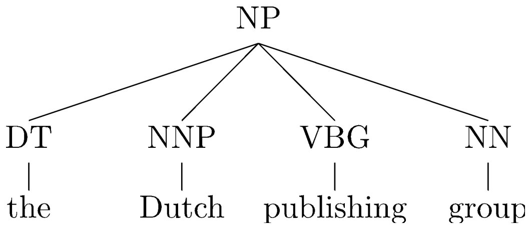
Transformation to CNF form

- What one need to do to convert to CNF form
    - Get rid of rules that mix terminals and non-terminals
    - Get rid of unary rules: $C \rightarrow C_1$
    - Get rid of N-ary rules: $C \rightarrow C_1\ C_2 \ldots C_n \quad (n > 2)$

Crucial to process them, as required for efficient parsing

Transformation to CNF form: binarization

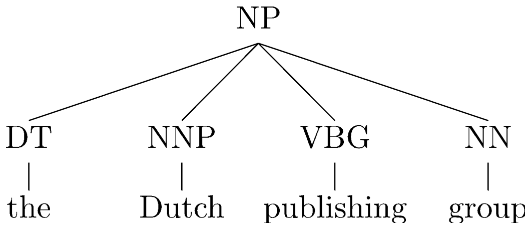
- Consider $$NP \rightarrow DT \;\; NNP \;\; VBG \;\; NN$$



- How do we get a set of binary rules which are equivalent?

Transformation to CNF form: binarization

- Consider $NP \rightarrow DT \;\; NNP \;\; VBG \;\; NN$



- How do we get a set of binary rules which are equivalent?

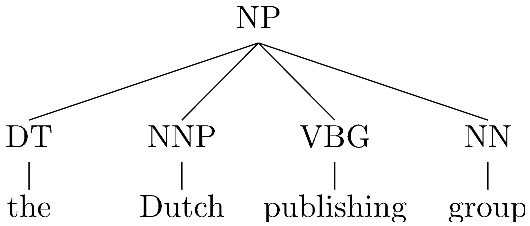
$$NP \rightarrow DT \;\; X$$

$$X \rightarrow NNP \;\; Y$$

$$Y \rightarrow VBG \;\; NN$$

Transformation to CNF form: binarization

- Consider $NP \rightarrow DT \;\; NNP \;\; VBG \;\; NN$



- How do we get a set of binary rules which are equivalent?

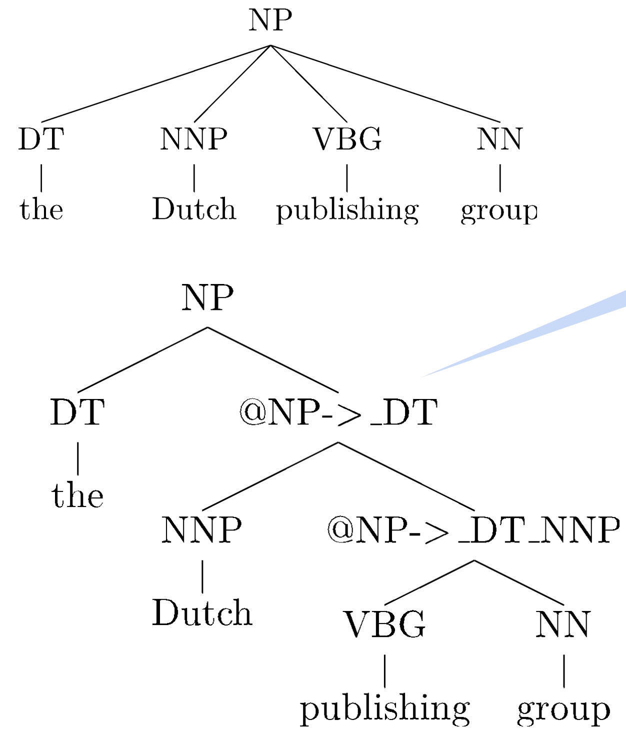
$$NP \rightarrow DT \;\; X$$
$$X \rightarrow NNP \;\; Y$$
$$Y \rightarrow VBG \;\; NN$$

- A more systematic way to refer to new non-terminals

$$NP \rightarrow DT \;\; @NP|DT$$
$$@NP|DT \rightarrow NNP \;\; @NP|DT\_NNP$$
$$@NP|DT\_NNP \rightarrow VBG \;\; NN$$

Transformation to CNF form: binarization

- Instead of binarizing tuples we can binarize trees on preprocessing:



Also known as **lossless Markovization** in the context of PCFGs

Can be easily reversed on postprocessing

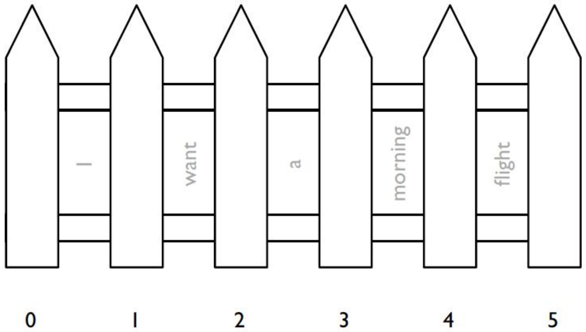
CKY: Parsing task

- We are given
  - a grammar <N, T, S, R>
  - a sequence of words $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$

- Our goal is to produce a parse tree for w

CKY: Parsing task

- We a given
  - a grammar <N, T, S, R>
  - a sequence of words $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$
- Our goal is to produce a parse tree for w
- We need an easy way to refer to substrings of w



indices refer to fenceposts

span (i, j) refers to words between fenceposts i and j

Parsing one word

$$C \rightarrow w_i$$

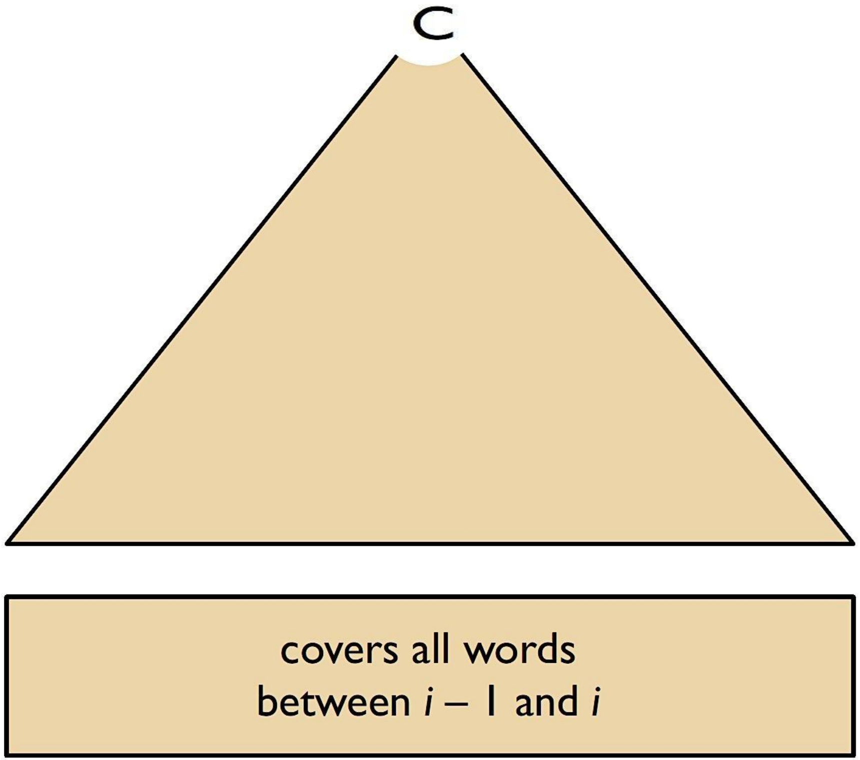
w_i

Parsing one word

$$C \to w_i$$

Parsing one word

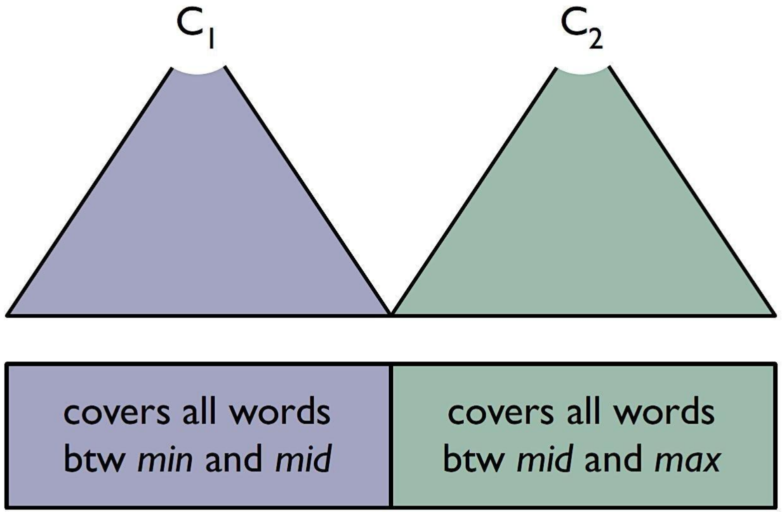
$$C \to w_i$$



Parsing longer spans

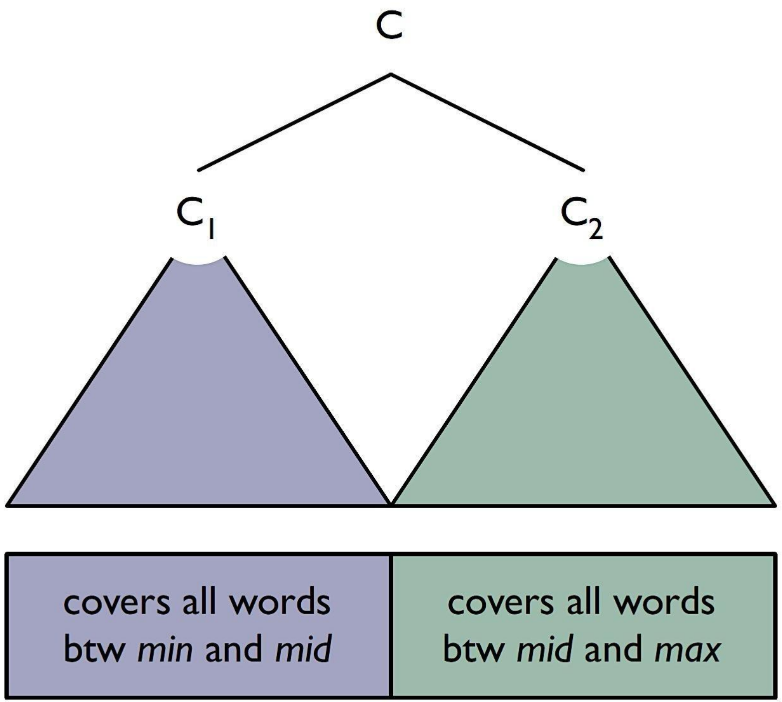
$C \rightarrow C_1 \ \ C_2$

Check through all C1, C2, mid



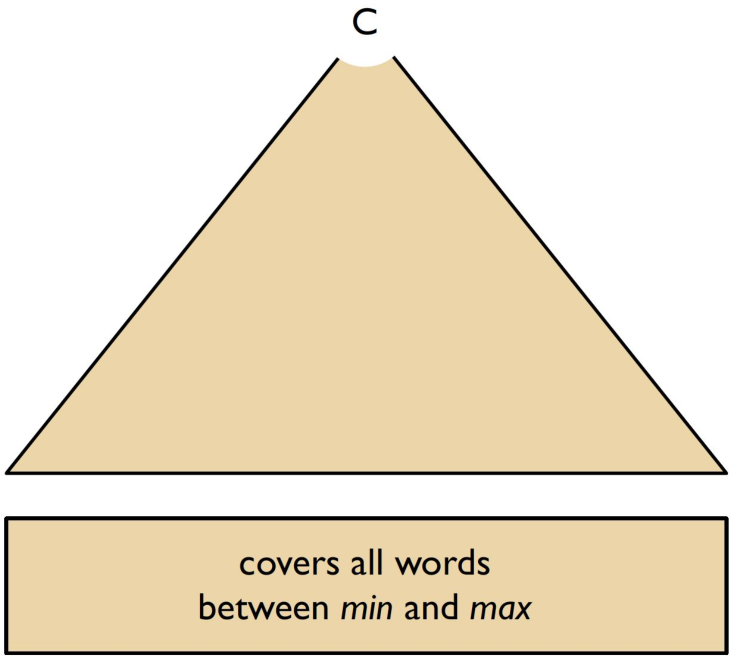
Parsing longer spans

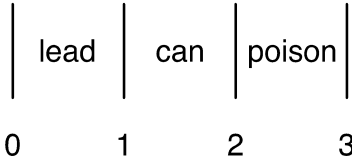
$$C \rightarrow C_1 \;\; C_2$$



Check through all C1, C2, mid

Parsing longer spans





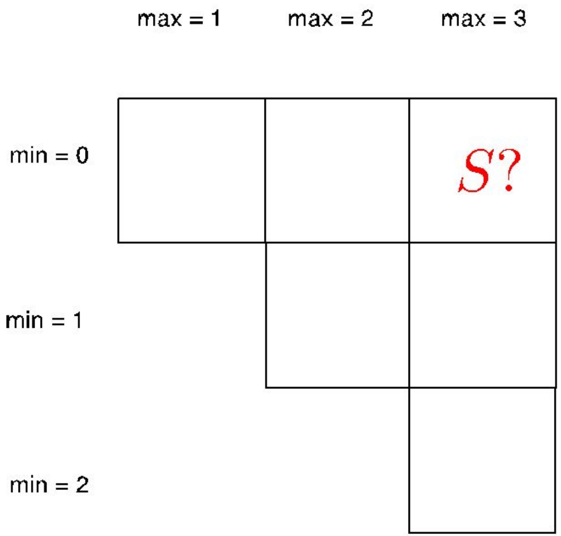


Chart (aka parsing triangle)

$S \to NP \ VP$

$VP \to M \ V$

$VP \to V$

$NP \to N$

$NP \to N \ NP$

Inner rules

$N \to can$

$N \to lead$

$N \to poison$

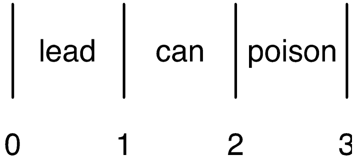
$M \to can$

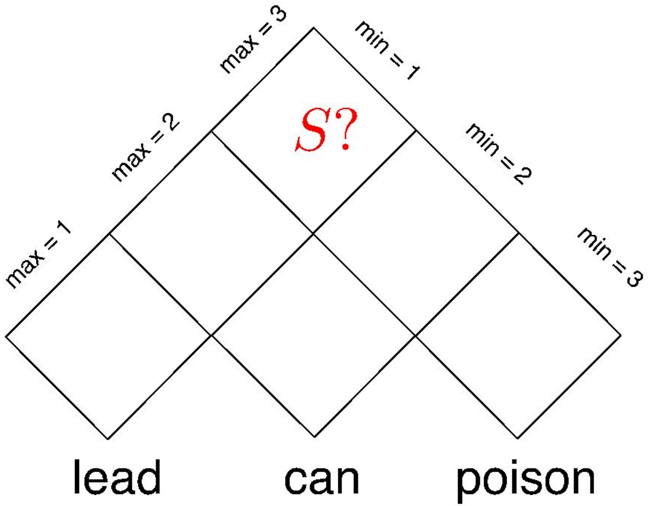
$M \to must$

$V \to poison$

$V \to lead$

Preterminal rules





$$S \rightarrow NP\ VP$$

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

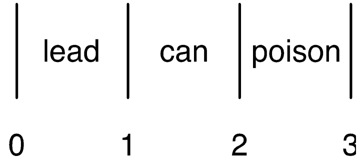
Inner rules

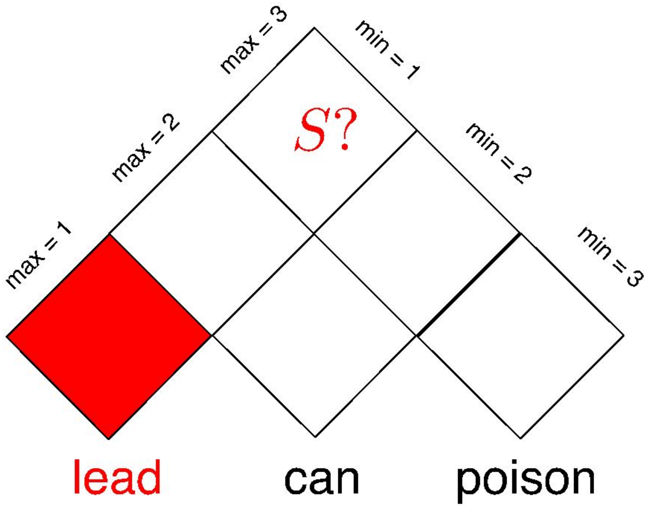
$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules





$$S \to NP \ VP$$

$$VP \to M \ V$$
$$VP \to V$$

$$NP \to N$$
$$NP \to N \ NP$$

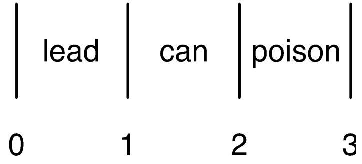
Inner rules

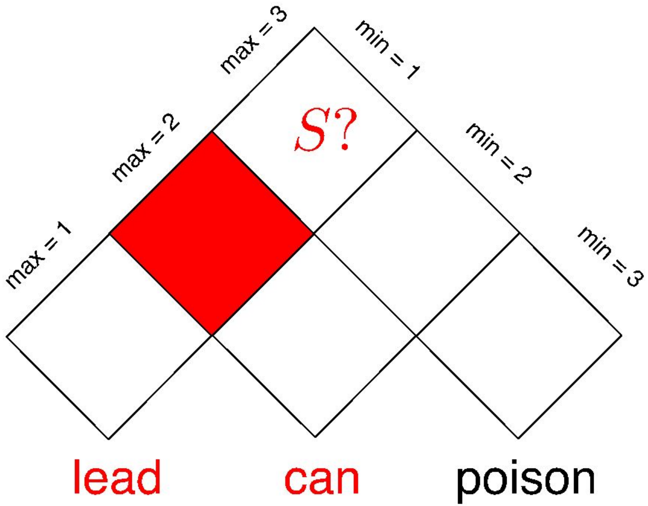
$$N \to can$$
$$N \to lead$$
$$N \to poison$$

$$M \to can$$
$$M \to must$$

$$V \to poison$$
$$V \to lead$$

Preterminal rules





$$S \rightarrow NP \ VP$$

$$VP \rightarrow M \ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N \ NP$$

Inner rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules

lead	can	poison

0 1 2 3

	max = 1	max = 2	max = 3
min = 0			$S?$
min = 1			
min = 2			

$$S \to NP \ VP$$

$$VP \to M \ V$$
$$VP \to V$$

$$NP \to N$$
$$NP \to N \ NP$$

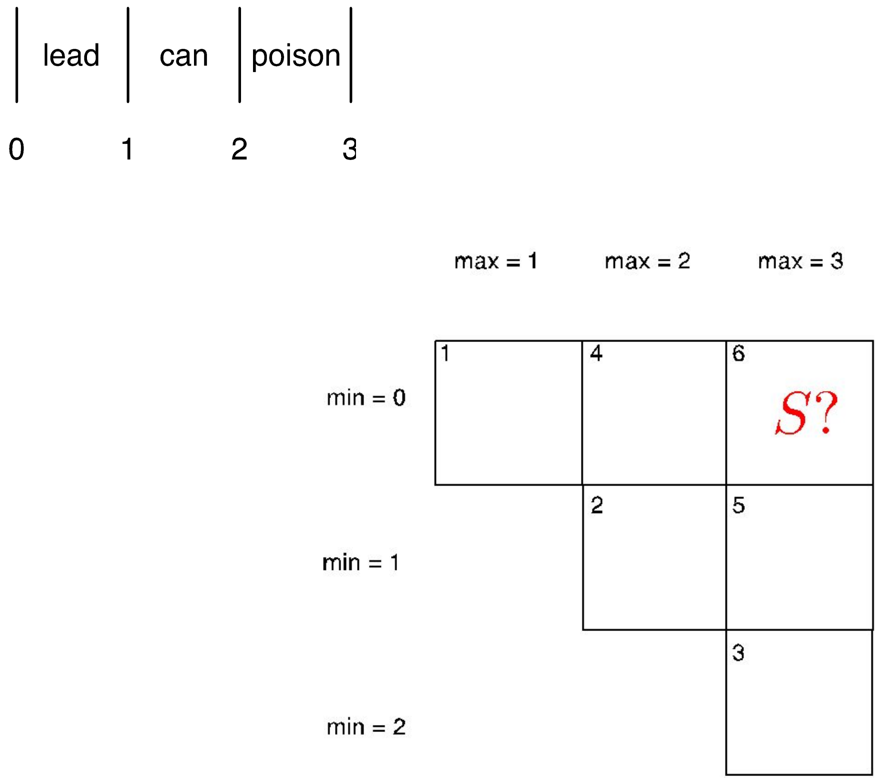
Inner rules

$$N \to can$$
$$N \to lead$$
$$N \to poison$$

$$M \to can$$
$$M \to must$$

$$V \to poison$$
$$V \to lead$$

Preterminal rules



$$S \rightarrow NP\ VP$$

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

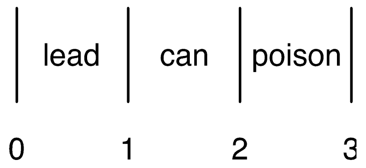
Inner rules

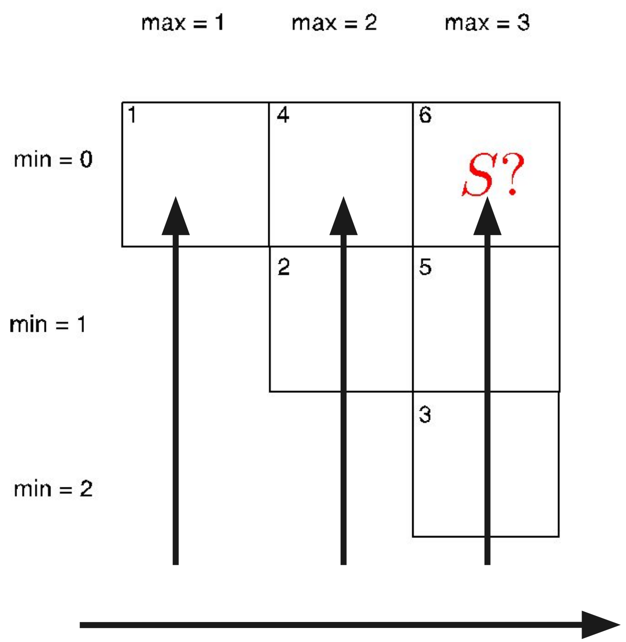
$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules





$$S \rightarrow NP \ VP$$

$$VP \rightarrow M \ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N \ NP$$

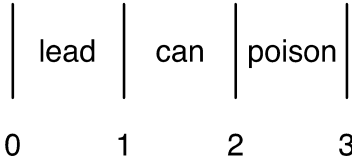
Inner rules

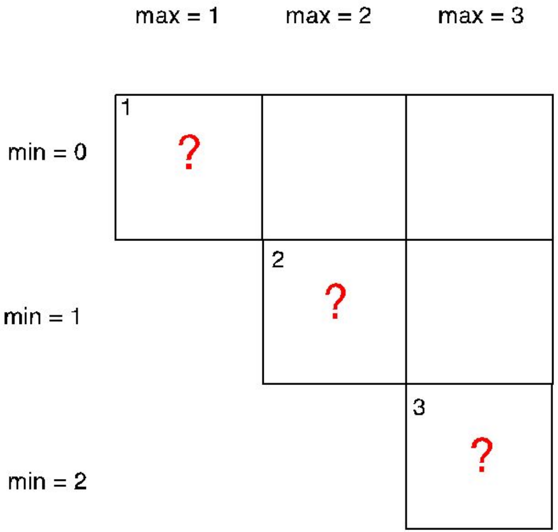
$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules





$$S \rightarrow NP\ VP$$

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

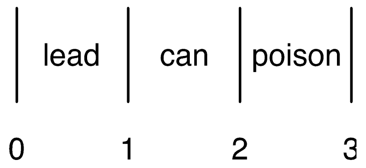
Inner rules

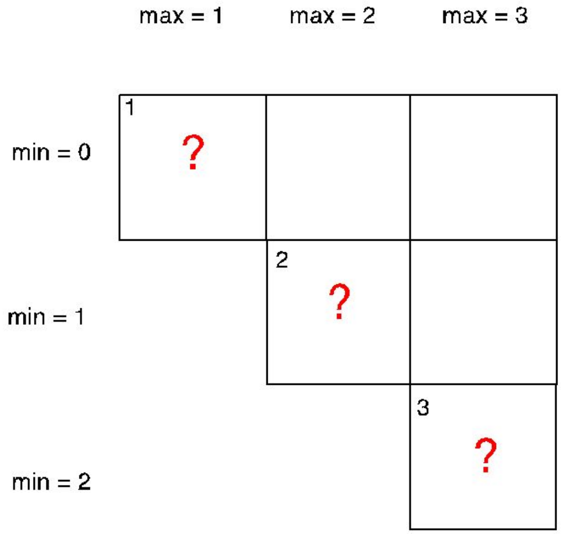
$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules





$S \rightarrow NP\ VP$

$VP \rightarrow M\ V$

$VP \rightarrow V$

$NP \rightarrow N$

$NP \rightarrow N\ NP$

Inner rules

$N \rightarrow can$

$N \rightarrow lead$

$N \rightarrow poison$

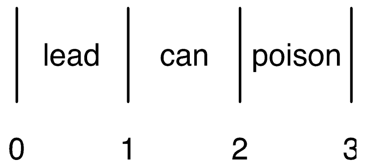
$M \rightarrow can$

$M \rightarrow must$

$V \rightarrow poison$

$V \rightarrow lead$

Preterminal rules



	max = 1	max = 2	max = 3
min = 0	1 N, V		
min = 1		2 N, M	
min = 2			3 N, V

$$S \rightarrow NP\ VP$$

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

Inner rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules

| lead | can | poison |

0 1 2 3

	max = 1	max = 2	max = 3
min = 0	1 N, V NP, VP	4 ?	
min = 1		2 N, M NP	
min = 2			3 N, V NP, VP

$$S \rightarrow NP\ VP$$

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

Inner rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules

| lead | can | poison |

0 1 2 3

	max = 1	max = 2	max = 3
min = 0	1 N, V NP, VP	4 ?	
min = 1		2 N, M NP	
min = 2			3 N, V NP, VP

$$S \rightarrow NP \ VP$$

$$VP \rightarrow M \ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$\boxed{NP \rightarrow N \ NP}$$

Inner rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules

$$S \rightarrow NP\ \ VP$$

| lead | can | poison |

0 1 2 3

	max = 1	max = 2	max = 3
min = 0	1 N, V NP, VP	4 NP	
min = 1		2 N, M NP	
min = 2			3 N, V NP, VP

$$VP \rightarrow M\ \ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$\boxed{NP \rightarrow N\ NP}$$

Inner rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules

lead	can	poison

0 1 2 3

	max = 1	max = 2	max = 3
min = 0	1 N, V NP, VP	4 NP	
min = 1		2 N, M NP	5 ?
min = 2			3 N, V NP, VP

$$S \rightarrow NP\ VP$$

Inner rules

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$
$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

Preterminal rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$
$$M \rightarrow can$$
$$M \rightarrow must$$
$$V \rightarrow poison$$
$$V \rightarrow lead$$

| lead | can | poison |

0 1 2 3

	max = 1	max = 2	max = 3
min = 0	1 N, V NP, VP	4 NP	
min = 1		2 N, M NP	5 $S, VP,$ NP
min = 2			3 N, V NP, VP

$\boxed{S \rightarrow NP\ VP}$

$\boxed{VP \rightarrow M\ V}$

$VP \rightarrow V$

$NP \rightarrow N$

$\boxed{NP \rightarrow N\ NP}$

Inner rules

$N \rightarrow can$

$N \rightarrow lead$

$N \rightarrow poison$

$M \rightarrow can$

$M \rightarrow must$

$V \rightarrow poison$

$V \rightarrow lead$

Preterminal rules

	lead		can		poison	
0		1		2		3

	max = 1	max = 2	max = 3
min = 0	1 N, V NP, VP	4 NP	6 ?
min = 1		2 N, M NP	5 $S, VP,$ NP
min = 2			3 N, V NP, VP

$$S \rightarrow NP\ VP$$

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

Inner rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules

| lead | can | poison |

0 1 2 3

	max = 1	max = 2	max = 3
min = 0	1 N, V NP, VP	4 NP	6 ?
min = 1		2 N, M NP	5 $S, VP,$ NP
min = 2			3 N, V NP, VP

$$S \rightarrow NP\ VP$$

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

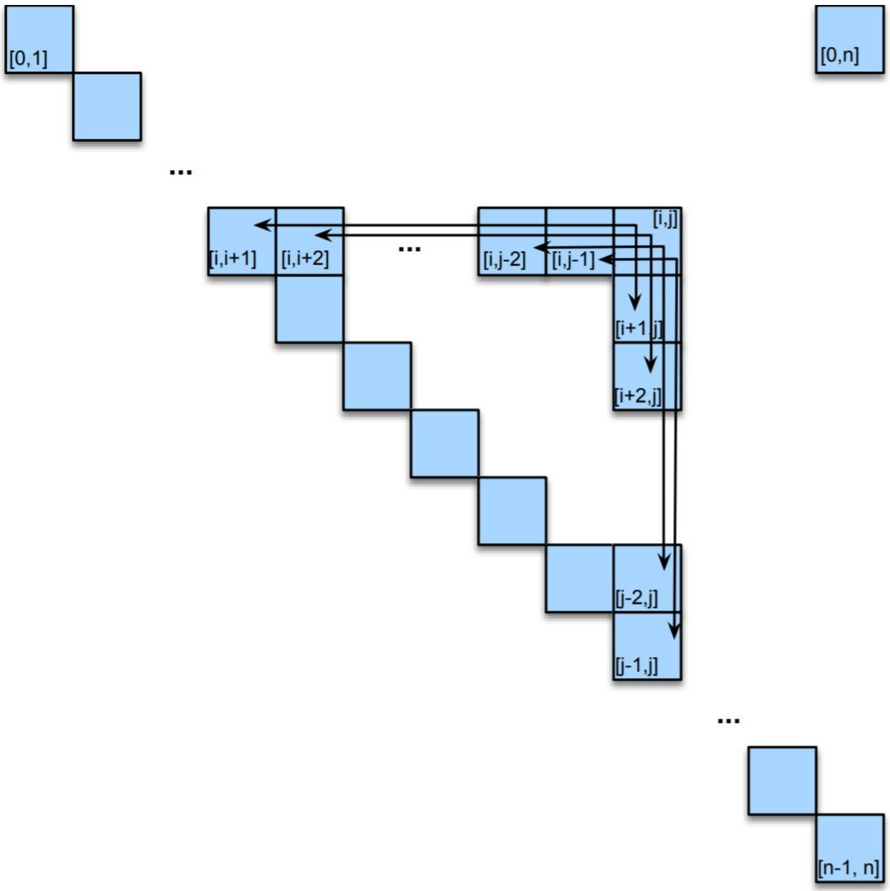
Inner rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules



[0,1]
[0,n]
...
[i,i+1]
[i,i+2]
...
[i,j-2]
[i,j-1]
[i,j]
[i+1,j]
[i+2,j]
[j-2,j]
[j-1,j]
...
[n-1, n]

lead	can	poison

0 1 2 3

mid=1

	max = 1	max = 2	max = 3
min = 0	1 N, V NP, VP	4 NP	6 S, NP
min = 1		2 N, M NP	5 $S, VP,$ NP
min = 2			3 N, V NP, VP

$$\boxed{S \rightarrow NP\ VP}$$

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$\boxed{NP \rightarrow N\ NP}$$

Inner rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

$$M \rightarrow can$$
$$M \rightarrow must$$

$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules

| lead | can | poison |
0 1 2 3

mid=2

	max = 1	max = 2	max = 3
min = 0	1 N, V NP, VP	4 NP	6 S, NP $S(?!)$
min = 1		2 N, M NP	5 $S, VP,$ NP
min = 2			3 N, V NP, VP

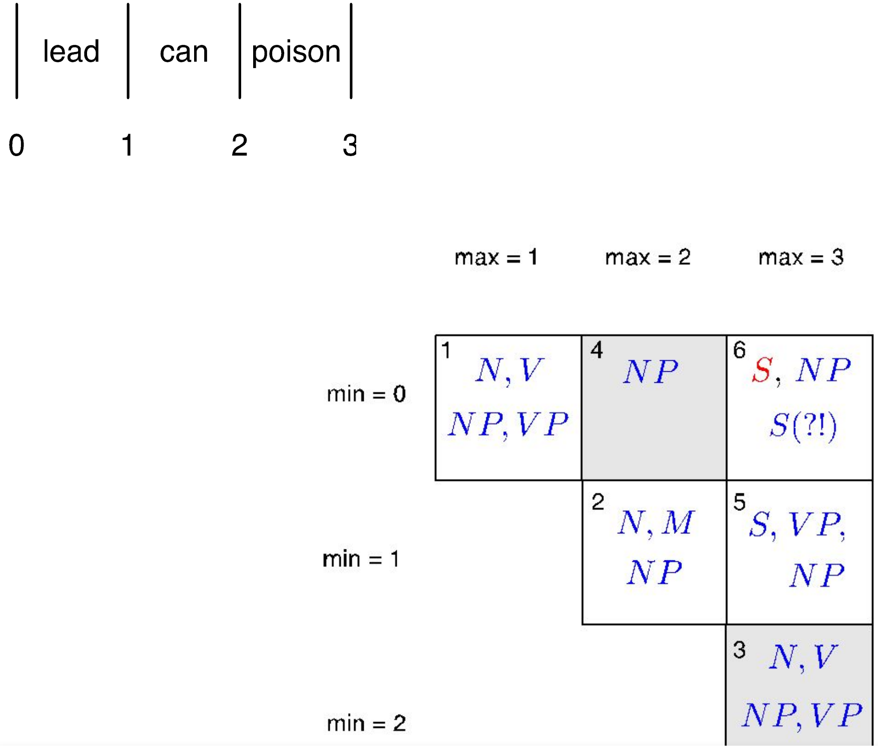
$$\boxed{S \rightarrow NP\ VP}$$

Inner rules

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$
$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

Preterminal rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$
$$M \rightarrow can$$
$$M \rightarrow must$$
$$V \rightarrow poison$$
$$V \rightarrow lead$$



$$S \rightarrow NP\ VP$$

$$VP \rightarrow M\ V$$
$$VP \rightarrow V$$

$$NP \rightarrow N$$
$$NP \rightarrow N\ NP$$

Inner rules

$$N \rightarrow can$$
$$N \rightarrow lead$$
$$N \rightarrow poison$$

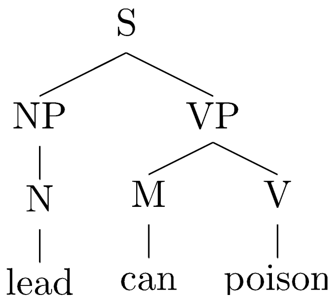
$$M \rightarrow can$$
$$M \rightarrow must$$

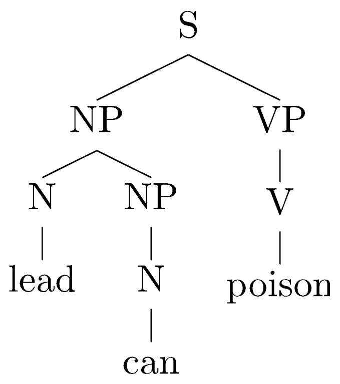
$$V \rightarrow poison$$
$$V \rightarrow lead$$

Preterminal rules

Apparently the sentence is ambiguous for the grammar: (as the grammar overgenerates)

Ambiguity





No subject-verb agreement, and *poison* used as an intransitive verb