Mining Unstructured Data
6. Constituent parsing

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat d'Informàtica de Barcelona

Outline

Outline

Goal and motivation

- Syntax studies the combination of words in a sentence.
- Syntactic parsing provides information of the combination of words in a sentence (the syntactic structure).
- Syntactic information is relevant for many NLP applications:
  - Authorship recognition
  - Grammar checking

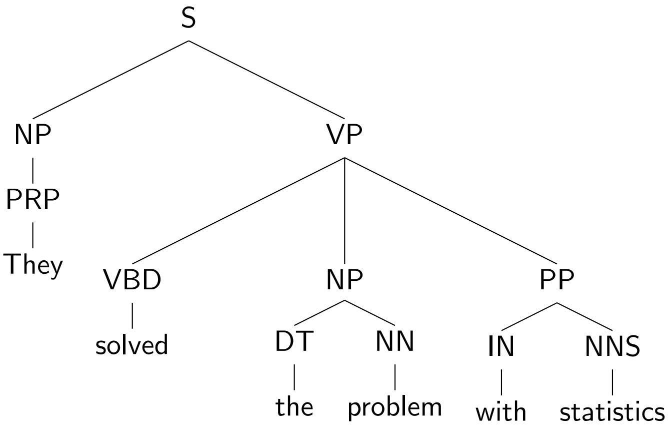
    Ex: 3th-Singular-noun + basic-verb $\Longrightarrow$ error
  - Machine Translation

    Ex: [es] NN+JJ $\Longrightarrow$ [en] JJ+NN
  - Information Extraction

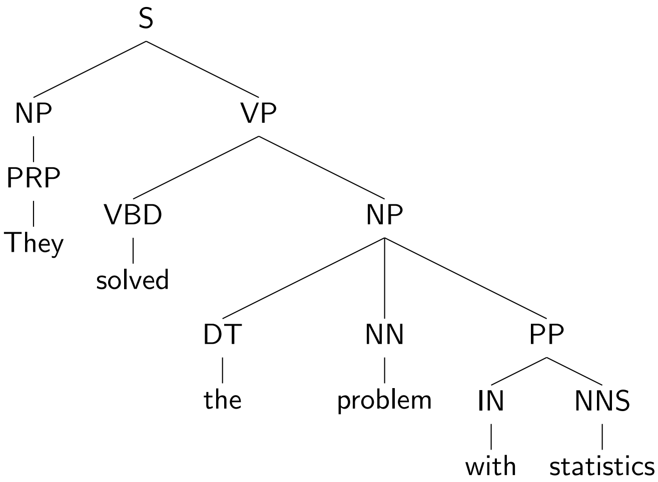
    Ex: $X - [subj] \rightarrow$ visited $\leftarrow [dobj] - Y \Longrightarrow$ visit(X,Y)
  - ...
- **Goal**: find the syntactic structure associated to a sentence.

Outline

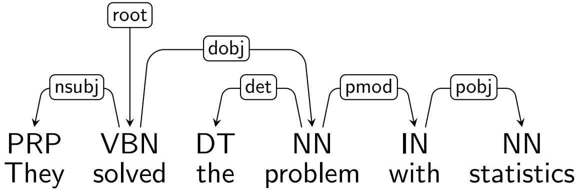
A Syntactic Tree



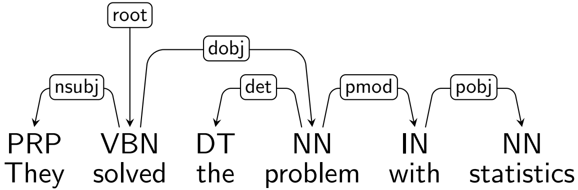
Another Syntactic Tree

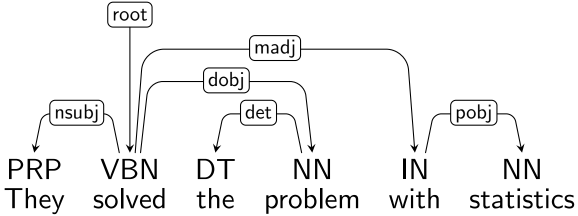


Dependency Trees

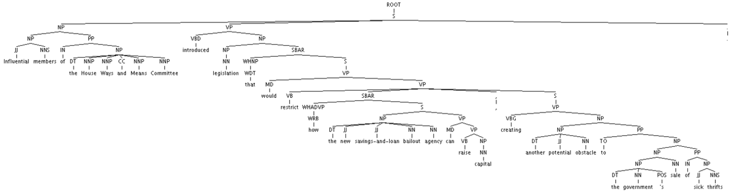


Dependency Trees





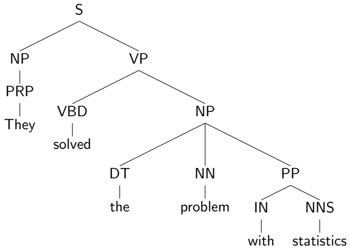
A "real" sentence



Influential members of the House Ways and Means Committee introduced legislation that would restrict how the new savings-and-loan bailout agency can raise capital, creating another potential obstacle to the government's sale of sick thrifts.

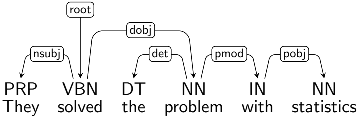
Theories of Syntactic Structure

Constituent Trees



- Main element: constituents (or phrases, or bracketings)
- Constituents = abstract linguistic units
- Results in nested trees

Dependency Trees



- Main element: dependency
- Focus on relations between words
- Handles *free word order* nicely.

Context Free Grammars (CFGs)

A context-free grammar is defined as a tuple $G = \langle N, \Sigma, R, S \rangle$ where:

- N is a set of non-terminal symbols
- $S \in N$ is a distinguished start symbol
- Σ is a set of terminal symbols
- R is a set of rules of the form $X \to Y_1 Y_2 \ldots Y_n$ where $n \geq 0,\ X \in N,\ Y_i \in N \cup \Sigma$

Context Free Grammars, Example

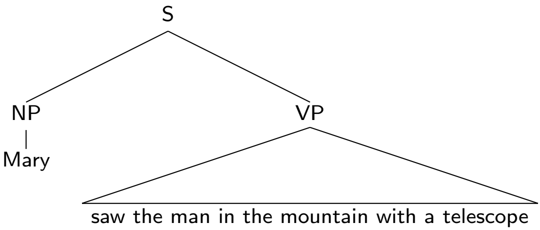
$$
\begin{aligned}
N &= \{\mathrm{S}, \mathrm{VP}, \mathrm{NP}, \mathrm{PP}, \mathrm{DT}, \mathrm{Vi}, \mathrm{Vt}, \mathrm{NN}, \mathrm{IN}\}^1 \\
S &= \{\mathrm{S}\} \\
\Sigma &= \{\mathrm{sleeps}, \mathrm{saw}, \mathrm{man}, \mathrm{woman}, \mathrm{telescope}, \mathrm{the}, \mathrm{with}, \mathrm{in}\} \\
R &= \left\{ \begin{array}{ll}
\mathrm{S} \to \mathrm{NP\ VP} & \mathrm{Vi} \to \mathrm{sleeps} \\
\mathrm{S} \to \mathrm{NP\ Vi} & \mathrm{Vt} \to \mathrm{saw} \\
\mathrm{NP} \to \mathrm{DT\ NN} & \mathrm{NN} \to \mathrm{man} \\
\mathrm{NP} \to \mathrm{NP\ PP} & \mathrm{NN} \to \mathrm{woman} \\
\mathrm{PP} \to \mathrm{IN\ NP} & \mathrm{NN} \to \mathrm{telescope} \\
\mathrm{VP} \to \mathrm{Vt\ NP} & \mathrm{DT} \to \mathrm{the} \\
\mathrm{VP} \to \mathrm{VP\ PP} & \mathrm{IN} \to \mathrm{with} \\
\mathrm{VP} \to \mathrm{Vi\ PP} & \mathrm{IN} \to \mathrm{in}
\end{array} \right\}
\end{aligned}
$$

[1] S=sentence, VP=verb phrase, NP=noun phrase, PP=prepositional phrase, DT=determiner, Vi=intransitive verb, Vt=transitive verb, NN=noun, IN=preposition

Properties of CFGs

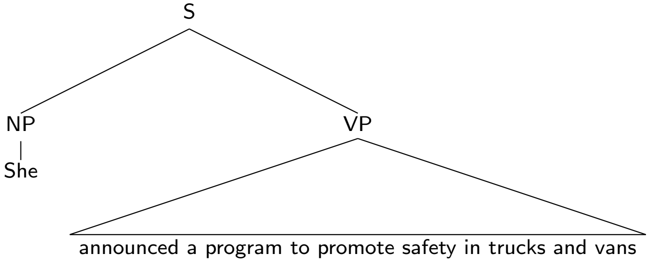
- A CFG defines a set of possible *derivations* (i.e. unique trees)
- A sequence of terminals $s \in \Sigma^*$ is *generated* by the CFG (or *recognized* by it, or *belongs* to the language defined by it) if there is at least a derivation that produces s.
- Some sequences of terminals generated by the CFG may have more than one derivation (*ambiguity*).

Ambiguity



- *Mary used a telescope to see a man who was in the mountain*
- *Mary saw a man who was in the mountain and carried a telescope*
- *Mary was in the mountain and used a telescope to see a man*
- *Mary was in the mountain that has a telescope and saw a man*
- *Mary saw a man who was in the mountain that has a telescope*
- *Mary was in the mountain and saw a man carrying a telescope*

Ambiguity



- *She announced a program aimed to make trucks and vans safer*
- *She used trucks and vans to announce a program aimed to promote safety*
- *She announced a program aimed to make trucks safer. She also announced vans*
- *She used trucks to announce a program aimed to promote safety. She also announced vans*
- *She announced a program. She did so in order to promote satefy in trucks and vans*
- *She used trucks and vans to announce a program. She did so in order to promote satefy*
- *...*

Ambiguity

Some trees are more likely than others...

Ambiguity

Some trees are more likely than others...

Can we model that?

Context Free Grammar (CFGs)

A context-free grammar is defined as a tuple $G = \langle N, \Sigma, R, S \rangle$ where:

- N is a set of non-terminal symbols
- $S \in N$ is a distinguished start symbol
- Σ is a set of terminal symbols
- R is a set of rules of the form $X \rightarrow Y_1 Y_2 \ldots Y_n$ where $n \geq 0,\ X \in N,\ Y_i \in N \cup \Sigma$

Context Free Grammar (CFGs)

A context-free grammar is defined as a tuple $G = \langle N, \Sigma, R, S \rangle$ where:

- N is a set of non-terminal symbols
- $S \in N$ is a distinguished start symbol
- Σ is a set of terminal symbols
- R is a set of rules of the form $X \rightarrow Y_1 Y_2 \ldots Y_n$ where $n \geq 0,\ X \in N,\ Y_i \in N \cup \Sigma$

Probabilistic Context Free Grammar (**P**CFGs)

A context-free grammar is defined as a tuple $G = \langle N, \Sigma, R, S \quad \rangle$ where:

- N is a set of non-terminal symbols
- $S \in N$ is a distinguished start symbol
- Σ is a set of terminal symbols
- R is a set of rules of the form $X \to Y_1 Y_2 \dots Y_n$ where $n \geq 0,\ X \in N,\ Y_i \in N \cup \Sigma$

Probabilistic Context Free Grammar (**P**CFGs)

A probabilistic context-free grammar is defined as a tuple $G = \langle N, \Sigma, R, S \quad \rangle$ where:

- N is a set of non-terminal symbols
- $S \in N$ is a distinguished start symbol
- Σ is a set of terminal symbols
- R is a set of rules of the form $X \to Y_1 Y_2 \ldots Y_n$ where $n \geq 0,\ X \in N,\ Y_i \in N \cup \Sigma$

Probabilistic Context Free Grammar (PCFGs)

A probabilistic context-free grammar is defined as a tuple $G = \langle N, \Sigma, R, S, q \rangle$ where:

- N is a set of non-terminal symbols
- $S \in N$ is a distinguished start symbol
- Σ is a set of terminal symbols
- R is a set of rules of the form $X \to Y_1 Y_2 \ldots Y_n$ where $n \geq 0, \; X \in N, \; Y_i \in N \cup \Sigma$

Probabilistic Context Free Grammar (PCFGs)

A probabilistic context-free grammar is defined as a tuple $G = \langle N, \Sigma, R, S, q \rangle$ where:

- N is a set of non-terminal symbols
- $S \in N$ is a distinguished start symbol
- Σ is a set of terminal symbols
- R is a set of rules of the form $X \to Y_1 Y_2 \ldots Y_n$ where $n \geq 0,\ X \in N,\ Y_i \in N \cup \Sigma$
- q is a set of non-negative parameters, one for each rule $X \to \alpha \in R$ such that, for any $X \in N$,

$$\sum_{(X \to \alpha) \in R} q(X \to \alpha) = 1$$

Context Free Grammars, Example

$$
\begin{aligned}
N &= \{\mathrm{S}, \mathrm{VP}, \mathrm{NP}, \mathrm{PP}, \mathrm{DT}, \mathrm{Vi}, \mathrm{Vt}, \mathrm{NN}, \mathrm{IN}\}^{1} \\
S &= \{\mathrm{S}\} \\
\Sigma &= \{\mathrm{sleeps}, \mathrm{saw}, \mathrm{man}, \mathrm{woman}, \mathrm{telescope}, \mathrm{the}, \mathrm{with}, \mathrm{in}\} \\
R &= \left\{ \begin{array}{ll}
\mathrm{S} \rightarrow \mathrm{NP\ VP} & \mathrm{Vi} \rightarrow \mathrm{sleeps} \\
\mathrm{S} \rightarrow \mathrm{NP\ Vi} & \mathrm{Vt} \rightarrow \mathrm{saw} \\
\mathrm{NP} \rightarrow \mathrm{DT\ NN} & \mathrm{NN} \rightarrow \mathrm{man} \\
\mathrm{NP} \rightarrow \mathrm{NP\ PP} & \mathrm{NN} \rightarrow \mathrm{woman} \\
\mathrm{PP} \rightarrow \mathrm{IN\ NP} & \mathrm{NN} \rightarrow \mathrm{telescope} \\
\mathrm{VP} \rightarrow \mathrm{Vt\ NP} & \mathrm{DT} \rightarrow \mathrm{the} \\
\mathrm{VP} \rightarrow \mathrm{VP\ PP} & \mathrm{IN} \rightarrow \mathrm{with} \\
\mathrm{VP} \rightarrow \mathrm{Vi\ PP} & \mathrm{IN} \rightarrow \mathrm{in}
\end{array} \right\}
\end{aligned}
$$

[1] S=sentence, VP=verb phrase, NP=noun phrase, PP=prepositional phrase, DT=determiner, Vi=intransitive verb, Vt=transitive verb, NN=noun, IN=preposition

Probabilistic Context Free Grammars, Example

$$N = \{\mathrm{S}, \mathrm{VP}, \mathrm{NP}, \mathrm{PP}, \mathrm{DT}, \mathrm{Vi}, \mathrm{Vt}, \mathrm{NN}, \mathrm{IN}\}$$ [1]

$$S = \{\mathrm{S}\}$$

$$\Sigma = \{\mathrm{sleeps}, \mathrm{saw}, \mathrm{man}, \mathrm{woman}, \mathrm{telescope}, \mathrm{the}, \mathrm{with}, \mathrm{in}\}$$

$$R = \left\{ \begin{array}{llll} \mathrm{S} \rightarrow \mathrm{NP\ VP} & 0.5 & \mathrm{Vi} \rightarrow \mathrm{sleeps} & 1.0 \\ \mathrm{S} \rightarrow \mathrm{NP\ Vi} & 0.5 & \mathrm{Vt} \rightarrow \mathrm{saw} & 1.0 \\ \mathrm{NP} \rightarrow \mathrm{DT\ NN} & 0.4 & \mathrm{NN} \rightarrow \mathrm{man} & 0.7 \\ \mathrm{NP} \rightarrow \mathrm{NP\ PP} & 0.6 & \mathrm{NN} \rightarrow \mathrm{woman} & 0.2 \\ \mathrm{PP} \rightarrow \mathrm{IN\ NP} & 1.0 & \mathrm{NN} \rightarrow \mathrm{telescope} & 0.1 \\ \mathrm{VP} \rightarrow \mathrm{Vt\ NP} & 0.4 & \mathrm{DT} \rightarrow \mathrm{the} & 1.0 \\ \mathrm{VP} \rightarrow \mathrm{VP\ PP} & 0.1 & \mathrm{IN} \rightarrow \mathrm{with} & 0.5 \\ \mathrm{VP} \rightarrow \mathrm{Vi\ PP} & 0.5 & \mathrm{IN} \rightarrow \mathrm{in} & 0.5 \end{array} \right\}$$

[1] S=sentence, VP=verb phrase, NP=noun phrase, PP=prepositional phrase, DT=determiner, Vi=intransitive verb, Vt=transitive verb, NN=noun, IN=preposition

Properties of PCFGs

- The probability of a parse tree $t \in \mathcal{T}_G$ is computed as:

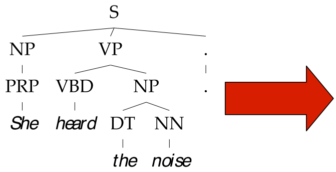
$$p(t) = \prod_{r \in t} q(r)$$

- If there is more than one tree for a sentence, we can rank them by probability.
- The most likely tree for a sentence s is:

$$\arg\max_{t \in \mathcal{T}(s)} p(t)$$

Learning Treebank Grammars

- Read the grammar rules from a treebank



S → NP VP .	1
NP → PRP	0.5
NP → DT NN	0.5
VP → VBD NP	1
PRP → She	1
…	

- Set rule weights by maximum likelihood

$$q(\alpha \rightarrow \beta) = \frac{\text{Count}(\alpha \rightarrow \beta)}{\text{Count}(\alpha)}$$

- Smoothing issues apply
- Having the appropriate CFG is critical to success

Outline

Outline

Parsing Natural Language Sentences

Possible goals of a parser:

Find all possible trees

Parsing Natural Language Sentences

Possible goals of a parser:

Find all possible trees, maybe ranked by probability

Parsing Natural Language Sentences

Possible goals of a parser:

> Find all possible trees, maybe ranked by probability or find the most likely tree.

Parsing Natural Language Sentences

Possible goals of a parser:

> Find all possible trees, maybe ranked by probability or find the most likely tree.

Parsing performance depends on many aspects:

Parsing Natural Language Sentences

Possible goals of a parser:

> Find all possible trees, maybe ranked by probability or find the most likely tree.

Parsing performance depends on many aspects:

- Grammar expressivity (combination of symbols)

Parsing Natural Language Sentences

Possible goals of a parser:

> Find all possible trees, maybe ranked by probability or find the most likely tree.

Parsing performance depends on many aspects:

- Grammar expressivity (combination of symbols)
- Coverage (words)

Parsing Natural Language Sentences

Possible goals of a parser:

> Find all possible trees, maybe ranked by probability or find the most likely tree.

Parsing performance depends on many aspects:

- Grammar expressivity (combination of symbols)
- Coverage (words)
- Parsing strategy (bottom-up, top-down)

Parsing Natural Language Sentences

Possible goals of a parser:

Find all possible trees, maybe ranked by probability or find the most likely tree.

Parsing performance depends on many aspects:

- Grammar expressivity (combination of symbols)
- Coverage (words)
- Parsing strategy (bottom-up, top-down)
- Rule application order (largest rule, most likely rule)

Parsing Natural Language Sentences

Possible goals of a parser:

Find all possible trees, maybe ranked by probability or find the most likely tree.

Parsing performance depends on many aspects:

- Grammar expressivity (combination of symbols)
- Coverage (words)
- Parsing strategy (bottom-up, top-down)
- Rule application order (largest rule, most likely rule)
- Ambiguity management (keep all, select one - probabilities, semantics, pragmatics)
- ...

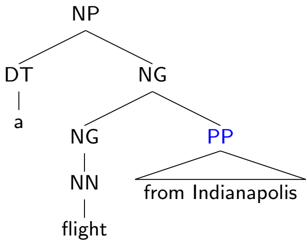
The problem of repeating derivations

- Top-down and bottom-up strategies both lead to repeated derivations when using backtracking

  Ex: "a flight from Indianapolis to Houston [on TWA...]"

  NG → NN

  NG → NG PP



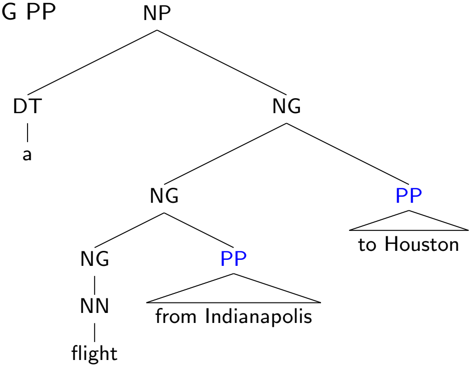
The problem of repeating derivations

- Top-down and bottom-up strategies both lead to repeated derivations when using backtracking

  Ex: "a flight from Indianapolis to Houston"

  NG → NN

  NG → NG PP



Outline

Properties

- They avoid re-doing derivations using dynamic programming.
- They represent derivations as a directed graph named chart.
- They use a dynamic programming table to build the chart.

Chart

- Nodes: positions between words of the input sentence
- Edges: dotted rules subsuming a sequence of words of the input sentence

- Dotted rules represent rules states:
  - Passive rules: $A \rightarrow B_1 \ldots B_k \bullet$
  - Active rules: $A \rightarrow B_1 \ldots B_i \bullet B_{i+1} \ldots B_k$

Ex:

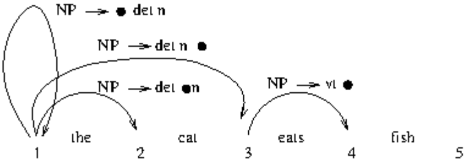
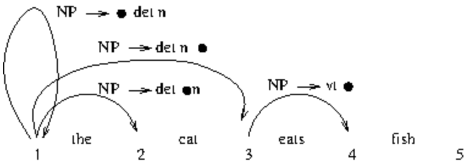


Chart as a dynamic programming table



				[1,5]
			[1,4]	[2,5]
		[1,3] NP → det n ●	[2,4]	[3,5]
	[1,2] NP → det ● n	[2,3]	[3,4] VP → v_i ●	[4,5]
[1,1] NP → ● det n	[2,2]	[3,3]	[4,4]	[5,5]
1	**2**	**3**	**4**	**5**

| the | cat | eats | fish

Popular chart-based algorithms

- CKY algorithm
  - introduced dynamic programming
  - limited to CFGs in Chomsky Normal Form
  - passive bottom-up chart parser (only passive rules)
  - straightforward probabilistic version
- Earley algorithm
  - any CFG
  - active top-down parser (active/passive rules)
  - non-straightforward probabilistic version
- Generalized chart parsing

Outline

CKY Algorithm properties

- Bottom-up
- Requires a grammar in Chomsky Normal Form (CNF).
- Dynammic programming: Store partial results that can be reused in different candidate solutions.
- Analogous to Viterbi in HMMs.
- Intermediate results stored in a *chart* structure.

Chomsky Normal Form (CNF)

A CFG $G = (N, \Sigma, R, S)$ expressed in CNF is as follows:

- N is a set of non-terminal symbols
- Σ is a set of terminal symbols
- R is a set of rules which take one of two forms:
  - $X \rightarrow Y_1 Y_2$ for $X, Y_1, Y_2 \in N$
  - $X \rightarrow \alpha$ for $X \in N$ and $\alpha \in \Sigma$
- $S \in N$ is a start symbol

Any CFG can be converted into CNF

CNF conversion

1. Convert Hybrid rules: replace terminals with new non-terminals

   Ex: $INF\_VP \rightarrow to\ VP \quad (p_1) \quad \Longrightarrow \quad INF\_VP \rightarrow TO\ VP \quad (p_1)$
   $TO \rightarrow to \quad (1.0)$

2. Convert non-binary rules:

   Ex: $S \rightarrow VP\ NP\ PP \quad (p_1) \quad \Longrightarrow \quad S \rightarrow VP\ X \quad (p_1)$
   $X \rightarrow NP\ PP \quad (1.0)$

3. Convert unit productions: $A \rightarrow^* B$ and $B \rightarrow \alpha \Longrightarrow A \rightarrow \alpha$

   Ex: $NP \rightarrow N \quad (p_1)$
   $N \rightarrow dog \quad (p_2) \quad \Longrightarrow \quad NP \rightarrow dog \quad (p_1 * p_2)$

Exercise

Convert the following PCFG to CNF

1. $S \rightarrow NP\ VP \quad (1.0)$
2. $NP \rightarrow det\ n \quad (0.6)$
3. $NP \rightarrow n \quad (0.4)$
4. $VP \rightarrow vt\ NP\ PP \quad (0.7)$
5. $VP \rightarrow vi \quad (0.3)$
6. $PP \rightarrow with\ NP \quad (1.0)$
7. $det \rightarrow the|a \quad (0.6|0.4)$
8. $n \rightarrow cat|fish|knife \quad (0.3|0.5|0.2)$
9. $vt \rightarrow eats \quad (1.0)$
10. $vi \rightarrow eats \quad (1.0)$

CKY Algorithm

Chart content:

- Maximum probability of a subtree with root X spanning words $i \dots j$:
$$\pi(i, j, X)$$
- Backpath to recover which rules produced the maximum probability tree:
$$\psi(i, j, X)$$

The goal is to compute:

- $\max_{t \in \mathcal{T}(s)} p(t) = \pi(1, n, S)$
- $\psi(1, n, S)$
- It is possible to use it without probabilities to get all parse trees (with higher complexity)

CKY Algorithm

Base case: Tree leaves

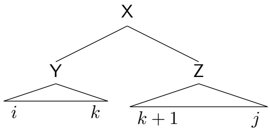
- $\forall i = 1 \dots n$, $\forall X \to w_i \in R$, $\pi(i, i, X) = q(X \to w_i)$

Recursive case: Non-terminal nodes

- $\forall i = 1 \dots n$, $\forall j = (i+1) \dots n$, $\forall X \in N$

$$\pi(i, j, X) = \max_{\substack{X \to Y Z \in R \\ k : i \le k < j}} q(X \to Y Z) \times \pi(i, k, Y) \times \pi(k+1, j, Z)$$

$$\psi(i, j, X) = \arg\max_{\substack{X \to Y Z \in R \\ k : i \le k < j}} q(X \to Y Z) \times \pi(i, k, Y) \times \pi(k+1, j, Z)$$



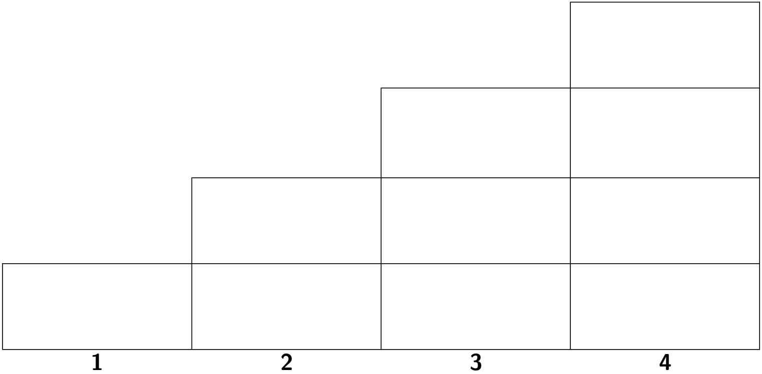
Output:

- Return $\pi(1, n, S)$ and recover backpath trough $\psi(1, n, S)$

CKY Algorithm

Supose $s = w_1 w_2 w_3 w_4$ and $G =< N, \Sigma, S, R, q >$ a PCFG

$R = \{X_k \to Y_s\, Z_t\} \cup \{X_k \to \alpha\}$



CKY Algorithm

Supose $s = w_1 w_2 w_3 w_4$ and $G =< N, \Sigma, S, R, q >$ a PCFG

$R = \{X_k \to Y_s\ Z_t\} \cup \{X_k \to \alpha\}$

Base case:

$\forall_k$ shot $\{X_k \to w_i\}$ to compute $\pi(i, i, X_k)$

11 $\{X_k \to w_1\}_k$

22 $\{X_k \to w_2\}_k$

33 $\{X_k \to w_3\}_k$

44 $\{X_k \to w_4\}_k$

1 **2** **3** **4**

CKY Algorithm

Supose $s = w_1 w_2 w_3 w_4$ and $G =< N, \Sigma, S, R, q >$ a PCFG

$R = \{X_k \rightarrow Y_s\ Z_t\} \cup \{X_k \rightarrow \alpha\}$

Recursive case:

$\forall_k$ shot $\{X_k \rightarrow Y_s\ Z_t\}$ to get $\varphi(i, j, X_k)$ and compute $\pi(i, j, X_k)$

	12	23	34
11 $\{X_k \rightarrow w_1\}_k$	22 $\{X_k \rightarrow w_2\}_k$	33 $\{X_k \rightarrow w_3\}_k$	44 $\{X_k \rightarrow w_4\}_k$
1	**2**	**3**	**4**

CKY Algorithm

Supose $s = w_1 w_2 w_3 w_4$ and $G =< N, \Sigma, S, R, q >$ a PCFG

$R = \{X_k \rightarrow Y_s\, Z_t\} \cup \{X_k \rightarrow \alpha\}$

Recursive case:

$\forall_k$ shot $\{X_k \rightarrow Y_s\, Z_t\}$ to get $\varphi(i, j, X_k)$ and compute $\pi(i, j, X_k)$

		13	24
	12	23	34
11 $\{X_k \rightarrow w_1\}_k$	22 $\{X_k \rightarrow w_2\}_k$	33 $\{X_k \rightarrow w_3\}_k$	44 $\{X_k \rightarrow w_4\}_k$
1	**2**	**3**	**4**

CKY Algorithm

Supose $s = w_1 w_2 w_3 w_4$ and $G =< N, \Sigma, S, R, q >$ a PCFG

$R = \{X_k \to Y_s\, Z_t\} \cup \{X_k \to \alpha\}$

Recursive case:

$\forall_k$ shot $\{X_k \to Y_s\, Z_t\}$ to get $\varphi(i, j, X_k)$ and compute $\pi(i, j, X_k)$

			14
		13	24
	12	23	34
11 $\{X_k \to w_1\}_k$	22 $\{X_k \to w_2\}_k$	33 $\{X_k \to w_3\}_k$	44 $\{X_k \to w_4\}_k$
1	**2**	**3**	**4**

CKY Algorithm

Supose $s = w_1 w_2 w_3 w_4$ and $G =< N, \Sigma, S, R, q >$ a PCFG

$R = \{X_k \rightarrow Y_s \, Z_t\} \cup \{X_k \rightarrow \alpha\}$

Recursive case:

$\forall_k$ shot $\{X_k \rightarrow Y_s \, Z_t\}$ to get $\varphi(i, j, X_k)$ and compute $\pi(i, j, X_k)$

Example for (1,3)

		13 $\{X_k \rightarrow Y_{s,11} \, Z_{t,23}\}_k$	24
	12	23	34
11 $\{X_k \rightarrow w_1\}_k$	22 $\{X_k \rightarrow w_2\}_k$	33 $\{X_k \rightarrow w_3\}_k$	44 $\{X_k \rightarrow w_4\}_k$
1	**2**	**3**	**4**

CKY Algorithm

Supose $s = w_1 w_2 w_3 w_4$ and $G =< N, \Sigma, S, R, q >$ a PCFG

$R = \{X_k \rightarrow Y_s\, Z_t\} \cup \{X_k \rightarrow \alpha\}$

Recursive case:

$\forall_k$ shot $\{X_k \rightarrow Y_s\, Z_t\}$ to get $\varphi(i, j, X_k)$ and compute $\pi(i, j, X_k)$

Example for (1,3)

1	2	3	4
		13 $\{X_k \rightarrow Y_{s,12}\, Z_{t,33}\}_k$	24
	12	23	34
11 $\{X_k \rightarrow w_1\}_k$	22 $\{X_k \rightarrow w_2\}_k$	33 $\{X_k \rightarrow w_3\}_k$	44 $\{X_k \rightarrow w_4\}_k$

Exercise

Compute the best parse tree and its probability for the following input sentence using the PCFG:

"the woman saw the man with the telescope"

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
11	22	33	44	55	66	77	88
DT → the	NN → woman	Vt → saw	DT → the	NN → man	IN → with	DT → the	NN → telescope
1.0	0.2	1.0	1.0	0.7	0.5	1.0	0.1

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08						
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23					
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34				
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5		Vi → sleeps	1.0
S → NP Vi	0.5		Vt → saw	1.0
NP → DT NN	0.4		NN → man	0.7
NP → NP PP	0.6		NN → woman	0.2
PP → IN NP	1.0		NN → telescope	0.1
VP → Vt NP	0.4		DT → the	1.0
VP → VP PP	0.1		IN → with	0.5
VP → Vi PP	0.5		IN → in	0.5

1	2	3	4	5	6	7	8
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28			
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56		
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

		13					
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

		13					
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

		13	24				
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

		13	24				
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112			
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112			
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5		Vi → sleeps	1.0
S → NP Vi	0.5		Vt → saw	1.0
NP → DT NN	0.4		NN → man	0.7
NP → NP PP	0.6		NN → woman	0.2
PP → IN NP	1.0		NN → telescope	0.1
VP → Vt NP	0.4		DT → the	1.0
VP → VP PP	0.1		IN → with	0.5
VP → Vi PP	0.5		IN → in	0.5

		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.11	46		
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.11	46		
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
		13	24	35 $VP \rightarrow Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 $PP \rightarrow IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 $NP \rightarrow DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 $NP \rightarrow DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 $NP \rightarrow DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 PP → $IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

			14				
		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 PP → $IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5		Vi → sleeps	1.0	
S → NP Vi	0.5		Vt → saw	1.0	
NP → DT NN	0.4		NN → man	0.7	
NP → NP PP	0.6		NN → woman	0.2	
PP → IN NP	1.0		NN → telescope	0.1	
VP → Vt NP	0.4		DT → the	1.0	
VP → VP PP	0.1		IN → with	0.5	
VP → Vi PP	0.5		IN → in	0.5	

			14				
		13	24	35 $VP \rightarrow Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 $PP \rightarrow IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 $NP \rightarrow DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 $NP \rightarrow DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 $NP \rightarrow DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
			14				
		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 PP → $IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

			14	25	36	47	58
		13	24	35 $VP \rightarrow Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 $PP \rightarrow IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 $NP \rightarrow DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 $NP \rightarrow DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 $NP \rightarrow DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

				15 $S \rightarrow NP_{12} VP_{35}$ 0.5*0.08*0.112= 4.48e-3			
			14	25	36	47	58
		13	24	35 $VP \rightarrow Vt_{33} NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 $PP \rightarrow IN_{66} NP_{78}$ 1.0*0.5*0.04=0.02
	12 $NP \rightarrow DT_{11} NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 $NP \rightarrow DT_{44} NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 $NP \rightarrow DT_{77} NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
				15 S → $NP_{12}VP_{35}$ 0.5*0.08*0.112= 4.48e-3	46	37	48 NP → $NP_{45}PP_{68}$ 0.6*0.28*0.02=3.36e-3
			14	25	36	47	58
		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 PP → $IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

S → NP VP	0.5		Vi → sleeps	1.0
S → NP Vi	0.5		Vt → saw	1.0
NP → DT NN	0.4		NN → man	0.7
NP → NP PP	0.6		NN → woman	0.2
PP → IN NP	1.0		NN → telescope	0.1
VP → Vt NP	0.4		DT → the	1.0
VP → VP PP	0.1		IN → with	0.5
VP → Vi PP	0.5		IN → in	0.5

					16	27	38 VP → $Vt_{33} NP_{48}$ 0.4*1.0*3.36e-3=1.34e-3
				15 S → $NP_{12} VP_{35}$ 0.5*0.08*0.112= 4.48e-3	46	37	48 NP → $NP_{45} PP_{68}$ 0.6*0.28*0.02=3.36e-3
			14	25	36	47	58
		13	24	35 VP → $Vt_{33} NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 PP → $IN_{66} NP_{78}$ 1.0*0.5*0.04=0.02
	12 NP → $DT_{11} NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44} NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77} NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

					16	27	38 VP → $Vt_{33} NP_{48}$ 0.4*1.0*3.36e-3=1.34e-3 VP → $VP_{35} PP_{68}$ 0.1*0.112*0.02=2.24e-4
				15 S → $NP_{12} VP_{35}$ 0.5*0.08*0.112= 4.48e-3	46	37	48 NP → $NP_{45} PP_{68}$ 0.6*0.28*0.02=3.36e-3
			14	25	36	47	58
		13	24	35 VP → $Vt_{33} NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 PP → $IN_{66} NP_{78}$ 1.0*0.5*0.04=0.02
	12 NP → $DT_{11} NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44} NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77} NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

					16	27	38 VP → $Vt_{33}NP_{48}$ 0.4*1.0*3.36e-3=1.34e-3
				15 S → $NP_{12}VP_{35}$ 0.5*0.08*0.112= 4.48e-3	46	37	48 NP → $NP_{45}PP_{68}$ 0.6*0.28*0.02=3.36e-3
			14	25	36	47	58
		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 PP → $IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**

CKY Algorithm - Example

S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

1	2	3	4	5	6	7	8
							18 S → $NP_{12}VP_{38}$ 0.5*0.08*1.34e-3=5.38e-5
						17	28
					16	27	38 VP → $Vt_{33}NP_{48}$ 0.4*1.0*3.36e-3=1.34e-3
				15 S → $NP_{12}VP_{35}$ 0.5*0.08*0.112= 4.48e-3	46	37	48 NP → $NP_{45}PP_{68}$ 0.6*0.28*0.02=3.36e-3
			14	25	36	47	58
		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 PP → $IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1

CKY Algorithm - Example

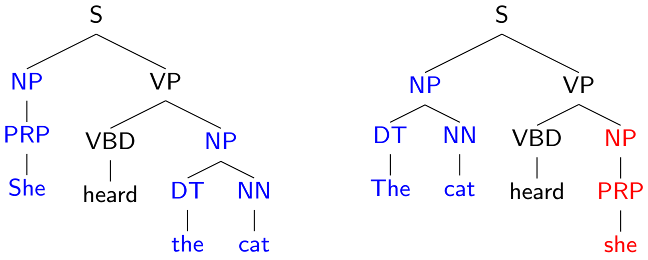
S → NP VP	0.5		Vi → sleeps	1.0
S → NP Vi	0.5		Vt → saw	1.0
NP → DT NN	0.4		NN → man	0.7
NP → NP PP	0.6		NN → woman	0.2
PP → IN NP	1.0		NN → telescope	0.1
VP → Vt NP	0.4		DT → the	1.0
VP → VP PP	0.1		IN → with	0.5
VP → Vi PP	0.5		IN → in	0.5

Final result

							18 S → $NP_{12}VP_{38}$ 0.5*0.08*1.34e-3=5.38e-5
						17	28
					16	27	38 VP → $Vt_{33}NP_{48}$ 0.4*1.0*3.36e-3=1.34e-3
				15 S → $NP_{12}VP_{35}$ 0.5*0.08*0.112= 4.48e-3	46	37	48 NP → $NP_{45}PP_{68}$ 0.6*0.28*0.02=3.36e-3
			14	25	36	47	58
		13	24	35 VP → $Vt_{33}NP_{45}$ 0.4*1.0*0.28=0.112	46	57	68 PP → $IN_{66}NP_{78}$ 1.0*0.5*0.04=0.02
	12 NP → $DT_{11}NN_{22}$ 0.4*1.0*0.2=0.08	23	34	45 NP → $DT_{44}NN_{55}$ 0.4*1.0*0.7=0.28	56	67	78 NP → $DT_{77}NN_{88}$ 0.4*1.0*0.1=0.04
11 DT → the 1.0	22 NN → woman 0.2	33 Vt → saw 1.0	44 DT → the 1.0	55 NN → man 0.7	66 IN → with 0.5	77 DT → the 1.0	88 NN → telescope 0.1
1	2	3	4	5	6	7	8

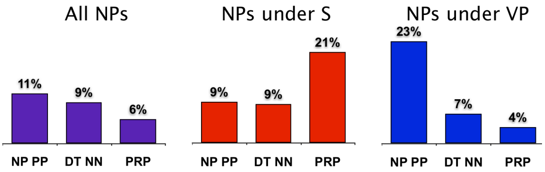
Why *context-free* ?

- Context-free means *context independent*, i.e, assumes that any expansion of a non-terminal is applicable, regardless of the context in which it occurs.



Natural Language is not Context-Free

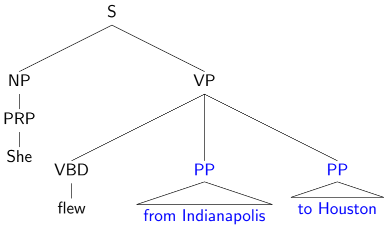
- NP expansion (for instance) is highly dependent on the parent of the NP

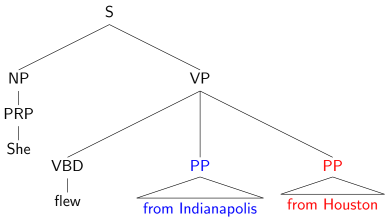


- Complete context independence is a too strong independence assumption for natural language.

Natural Language is not Context-Free

- The application of a rule may affect the applicability of others





Natural Language is not Context-Free

- May contain non-projective structures:

John saw the dog yesterday which was a Yorkshire Terrier