Linguistic Structure Prediction

CSE 447 / 517

Feb 24th, 2022 (Week 8)

Logistics

- A7 is due **tomorrow** (**Friday, 2/25**).

Agenda

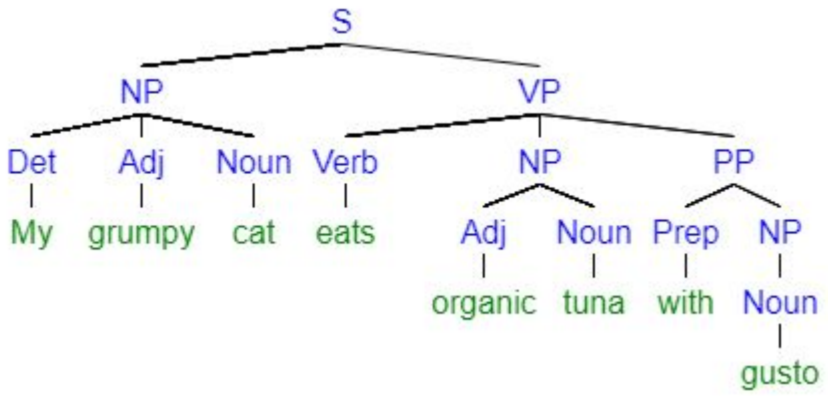
- Syntax
  - What is syntax?
  - Syntactic Constituents
  - CKY
  - Quiz 7, part I: CKY walkthrough
- Semantics
  - What is semantics?
  - Compositional semantics and λ-calculus
- Q & A

Syntax

What Is Syntax?

Linguistic definition: the study of how phrases and sentences are structured

(If this piques your interest, the Linguistics department has several classes that discuss syntax at length; see also Emily Bender's book on linguistic fundamentals for NLP.)



Syntactic Constituents

Running example: *My grumpy cat eats organic tuna with gusto.*

Constituents are groups of words (e.g., *my grumpy cat*, *organic tuna*) that:

- Can move (*It is organic tuna that my grumpy cat eats.*)
- Can be coordinated (*My grumpy cat eats organic tuna and Pacific salmon.*)
- Can answer a question (*What does my grumpy cat eat? Organic tuna.*)

Syntactic Constituents

Running example: *My grumpy cat eats organic tuna with gusto.*

Types of constituents

- Noun Phrase (NP) - *my grumpy cat, organic tuna*
- Verb Phrase (VP) - *eats organic tuna with gusto*
- Prepositional Phrase (PP) - *with gusto*
- …

Constituents and Recursion

this is the house

this is the house that Jack built

this is the cat that lives in the house that Jack built

this is the dog that chased the cat that lives in the house that Jack built

…

Context-Free Grammars (CFG)

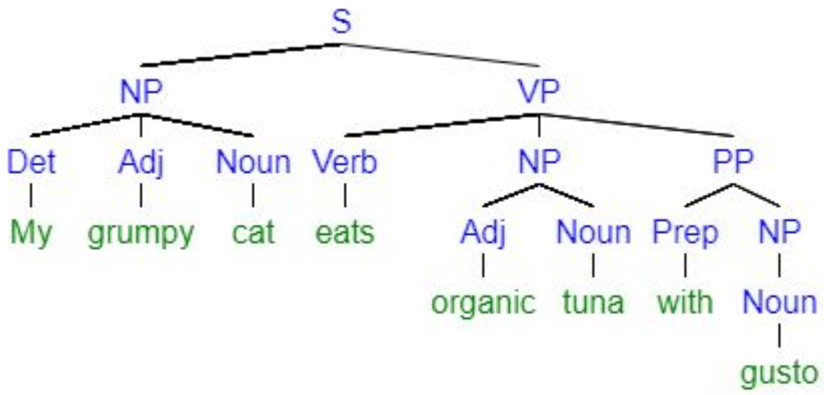
A **context-free grammar** consists of:

- Finite set of nonterminals $\mathcal{N}$
- Start symbol S
- Finite set of terminals (words) $\boldsymbol{\Sigma}$
- Production rule set $\mathcal{R}$, containing rules $\mathrm{N} \rightarrow \boldsymbol{\alpha}$
    - N - nonterminal from $\mathcal{N}$
    - $\boldsymbol{\alpha}$ - sequence of 0 or more terminals/nonterminals
        - Chomsky normal form - $\boldsymbol{\alpha}$ must be either a single terminal or two nonterminals

Phrase Structure Trees

We can use the rules from our CFG to build a phrase structure tree for a given sentence. This represents both the **syntactic structure** of the sentence and its **derivation** from our CFG.

S -> NP VP
NP -> Det Adj Noun
NP -> Adj Noun
NP -> Noun
VP -> Verb NP PP
PP -> Prep Noun
Det -> my
Adj -> grumpy | organic
Noun -> cat | tuna | gusto
Verb -> eats
Prep -> with



Syntax in NLP

Given a CFG and a sentence ***x***:

Recognition - is ***x*** in the CFG?

Parsing - how can we generate ***x*** from the rules of the CFG?

PCFGs and CKY

Slides from Yejin Choi, Chris Manning, Dan Klein, Michael Collins, Luke Zettlemoyer, Ray Mooney, and Graham Neubig

PCFG: Probabilistic Context Free Grammar

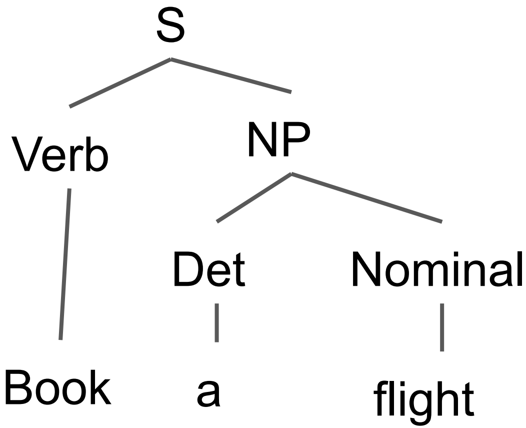
- A context-free grammar is a tuple <*N*, Σ ,*S*, *R*>
  - *N* : the set of non-terminals
    - Phrasal categories: S, NP, VP, ADJP, etc.
    - Parts-of-speech (pre-terminals): NN, JJ, DT, VB, etc.
  - Σ : the set of terminals (the words)
  - *S* : the start symbol
    - Often written as ROOT or TOP
    - *Not* usually the sentence non-terminal S
  - *R* : the set of rules
    - Of the form $X \rightarrow Y_1 Y_2 \ldots Y_n$, with $X \in N$, $n \geq 0$, $Y_i \in (N \cup \Sigma)$
    - Examples: S → NP VP, VP → VP CC VP
- A PCFG adds a distribution q:
  - Probability q(r) for each $r \in R$, such that for all $X \in N$:

$$\sum_{\alpha \rightarrow \beta \in R : \alpha = X} q(\alpha \rightarrow \beta) = 1$$

(Part of) A PCFG

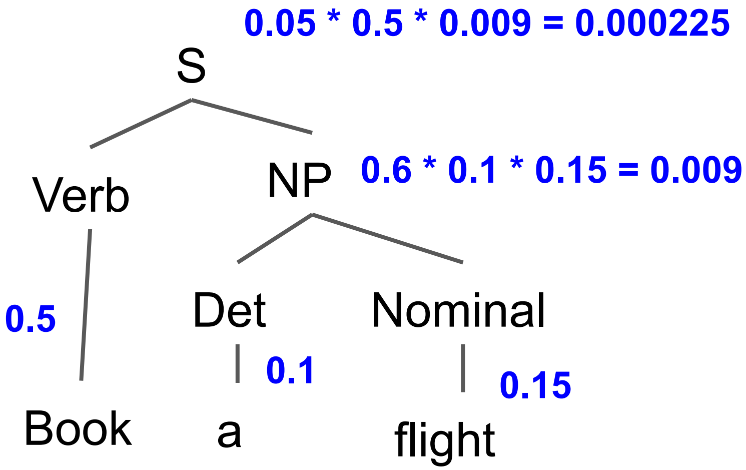
S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer 0.01 0.004 0.006	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me 0.1 0.02 0.02 0.06	
NP → Houston | NWA 0.16 .04	
Det → the | a | an 0.6 0.1 0.05	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money 0.03 0.15 0.06 0.06	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer 0.5 0.04 0.06	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from 0.2 0.3 0.3	
PP → Prep NP	1.0

(Part of) A PCFG



S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer 0.01 0.004 0.006	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me 0.1 0.02 0.02 0.06	
NP → Houston | NWA 0.16 .04	
Det→ the | a | an 0.6 0.1 0.05	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money 0.03 0.15 0.06 0.06	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb→ book | include | prefer 0.5 0.04 0.06	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from 0.2 0.3 0.3	
PP → Prep NP	1.0

(Part of) A PCFG



S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer 0.01 0.004 0.006	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me 0.1 0.02 0.02 0.06	
NP → Houston | NWA 0.16 .04	
Det→ the | a | an 0.6 0.1 0.05	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money 0.03 0.15 0.06 0.06	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb→ book | include | prefer 0.5 0.04 0.06	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from 0.2 0.3 0.3	
PP → Prep NP	1.0

Why do we need CKY?

- Find the best tree t* such that t* = argmax P(t) over all possible trees.
- It’s a decoding algorithm (analogy: Viterbi)
- Polynomial time (Dynamic programming)

Dynamic Programming

- We will store: score of the max parse of x_i to x_j with root non-terminal X

$$\pi(i, j, X)$$

- So we can compute the most likely parse:

$$\pi(1, n, S) = \max_{t \in \mathcal{T}_G(s)} p(t)$$

- Via the recursion:

$$\pi(i, j, X) = \max_{\substack{X \to YZ \in R, \\ s \in \{i \ldots (j-1)\}}} (q(X \to YZ) \times \pi(i, s, Y) \times \pi(s+1, j, Z))$$

- With base case:

$$\pi(i, i, X) = \begin{cases} q(X \to x_i) & \text{if } X \to x_i \in R \\ 0 & \text{otherwise} \end{cases}$$

CKY Algorithm

- Input: a sentence s = x_1 .. x_n and a PCFG = <N, Σ ,S, R, q>
- Initialization: For i = 1 … n and all X in N

Bottom-up:
Starting from (i, i, X)

$$\pi(i,i,X) = \begin{cases} q(X \rightarrow x_i) & \text{if } X \rightarrow x_i \in R \\ 0 & \text{otherwise} \end{cases}$$

- For l = 1 … (n-1) [iterate all phrase lengths]
  - For i = 1 … (n-l) and j = i+l [iterate all phrases of length l + 1]
    - For all X in N [iterate all non-terminals]

By this point, every span with length < l is already computed

$$\pi(i,j,X) = \max_{\substack{X \rightarrow YZ \in R, \\ s \in \{i \ldots (j-1)\}}} \left(q(X \rightarrow YZ) \times \pi(i,s,Y) \times \pi(s+1,j,Z) \right)$$

Previous states:
Best score for span (i,s) and non-terminal Y, and best score for span(s+1, j) and non-terminal Z

- also, store back pointers

$$bp(i,j,X) = \arg \max_{\substack{X \rightarrow YZ \in R, \\ s \in \{i \ldots (j-1)\}}} \left(q(X \rightarrow YZ) \times \pi(i,s,Y) \times \pi(s+1,j,Z) \right)$$

Quiz 7, Part I CKY

1. For the tree you build, what is the nonterminal for the constituent “prefer the flight to Houston”?

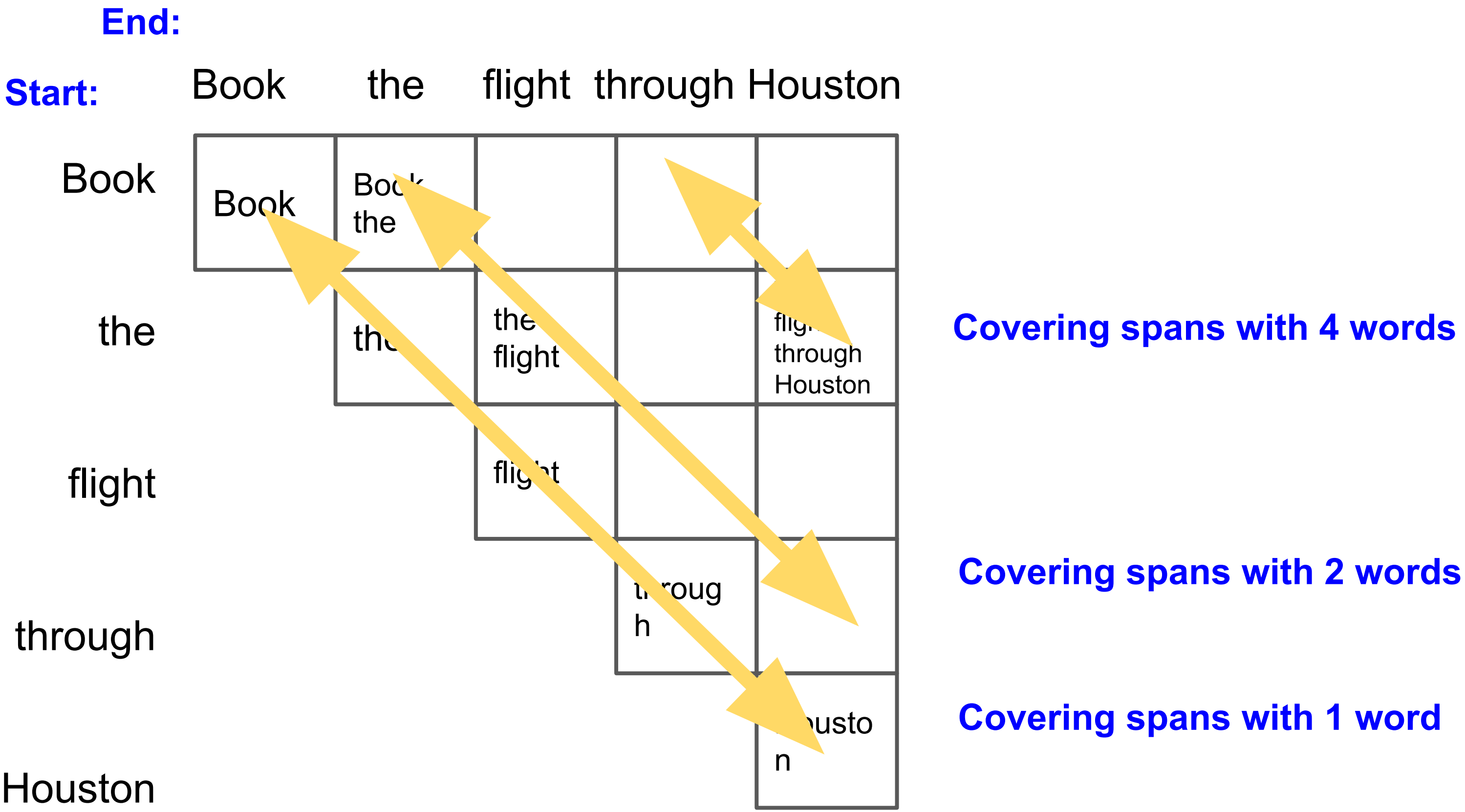
2. If “prefer” has a “Verb” tag, what could be a possible nonterminal for the constituent “the flight to Houston”?

3. If “prefer the flight” has a “VP” tag, what could be a possible nonterminal for the constituent “to Houston”?

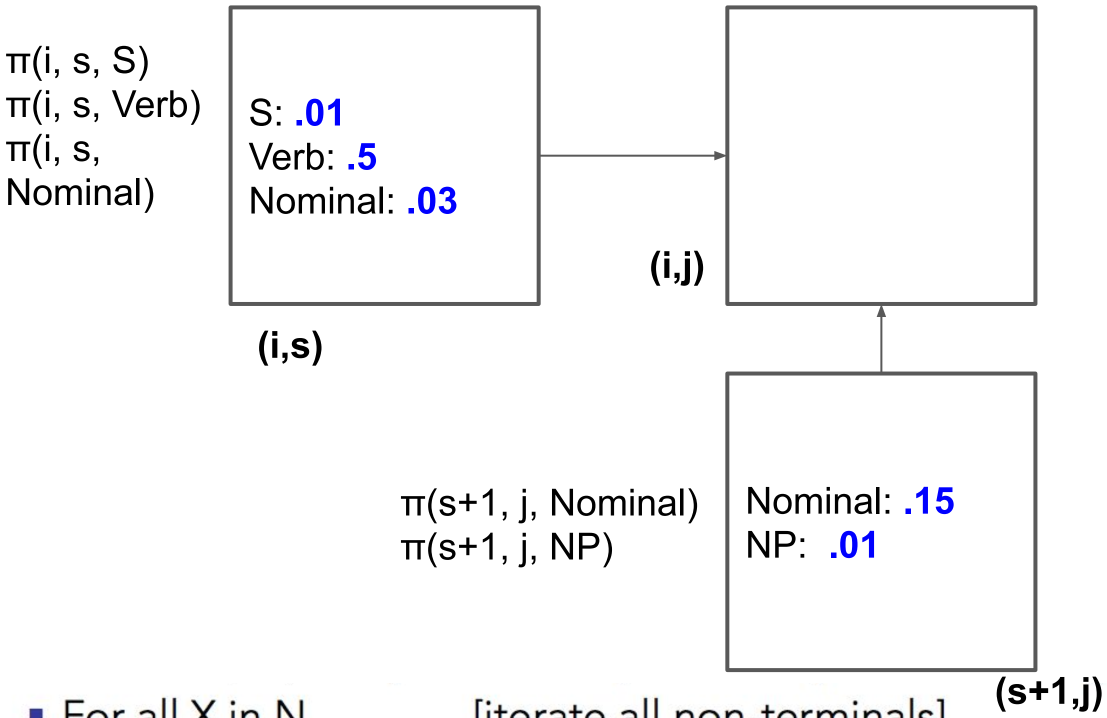
CKY Walkthrough: the Chart

Start: \ **End:**	Book	the	flight	through	Houston
Book					
the					
flight					
through					
Houston					

CKY Walkthrough



Compute the (Max) Scores

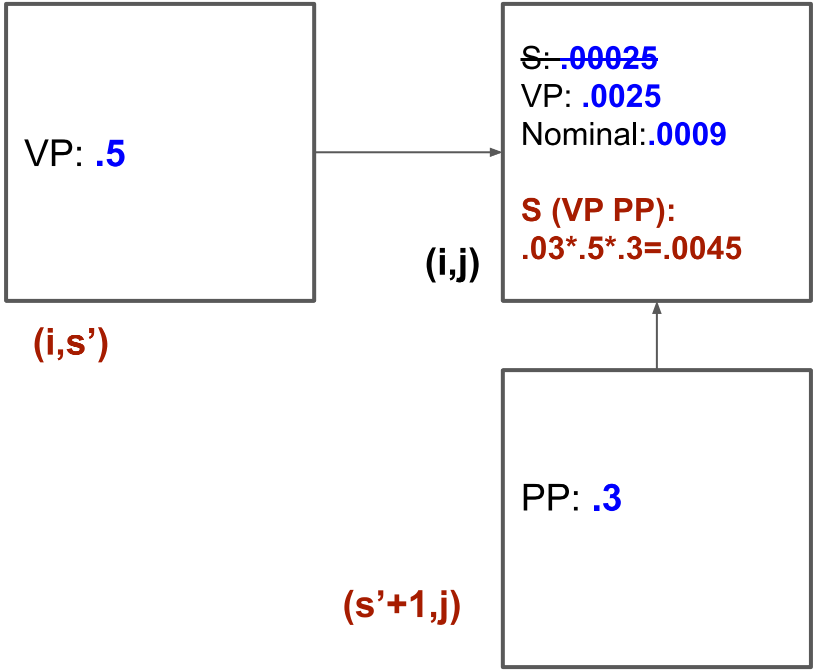


- For all X in N [iterate all non-terminals]

$$\pi(i,j,X) = \max_{\substack{X \to Y Z \in R,\\ s \in \{i \ldots (j-1)\}}} \left(q(X \to Y Z) \times \pi(i,s,Y) \times \pi(s+1,j,Z)\right)$$

S → NP VP 0.8
S → X1 VP 0.1
X1 → Aux NP 1.0
S → book | include | prefer
0.01 0.004 0.006
S → Verb NP 0.05
S → VP PP 0.03
NP → I | he | she | me
0.1 0.02 0.02 0.06
NP → Houston | NWA
0.16 .04
Det→ the | a | an
0.6 0.1 0.05
NP → Det Nominal 0.6
Nominal → book | flight | meal | money
0.03 0.15 0.06 0.06
Nominal → Nominal Nominal 0.2
Nominal → Nominal PP 0.5
Verb→ book | include | prefer
0.5 0.04 0.06
VP → Verb NP 0.5
VP → VP PP 0.3
Prep → through | to | from
0.2 0.3 0.3
PP → Prep NP 1.0

Compute the (Max) Scores (2)

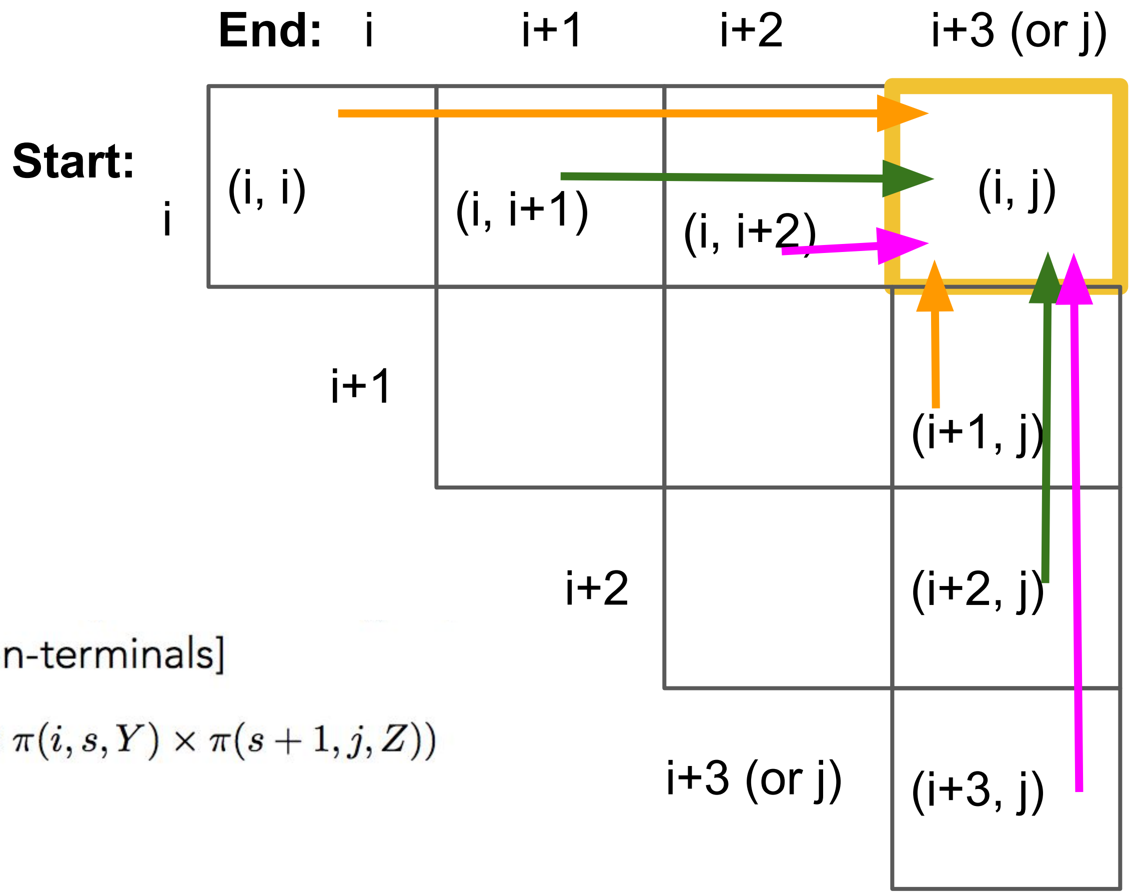


- For all X in N [iterate all non-terminals]

$$\pi(i,j,X) = \max_{\substack{X \to Y Z \in R, \\ s \in \{i \ldots (j-1)\}}} \left(q(X \to Y Z) \times \pi(i,s,Y) \times \pi(s+1,j,Z) \right)$$

S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer 0.01 0.004 0.006	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me 0.1 0.02 0.02 0.06	
NP → Houston | NWA 0.16 .04	
Det→ the | a | an 0.6 0.1 0.05	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money 0.03 0.15 0.06 0.06	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb→ book | include | prefer 0.5 0.04 0.06	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from 0.2 0.3 0.3	
PP → Prep NP	1.0

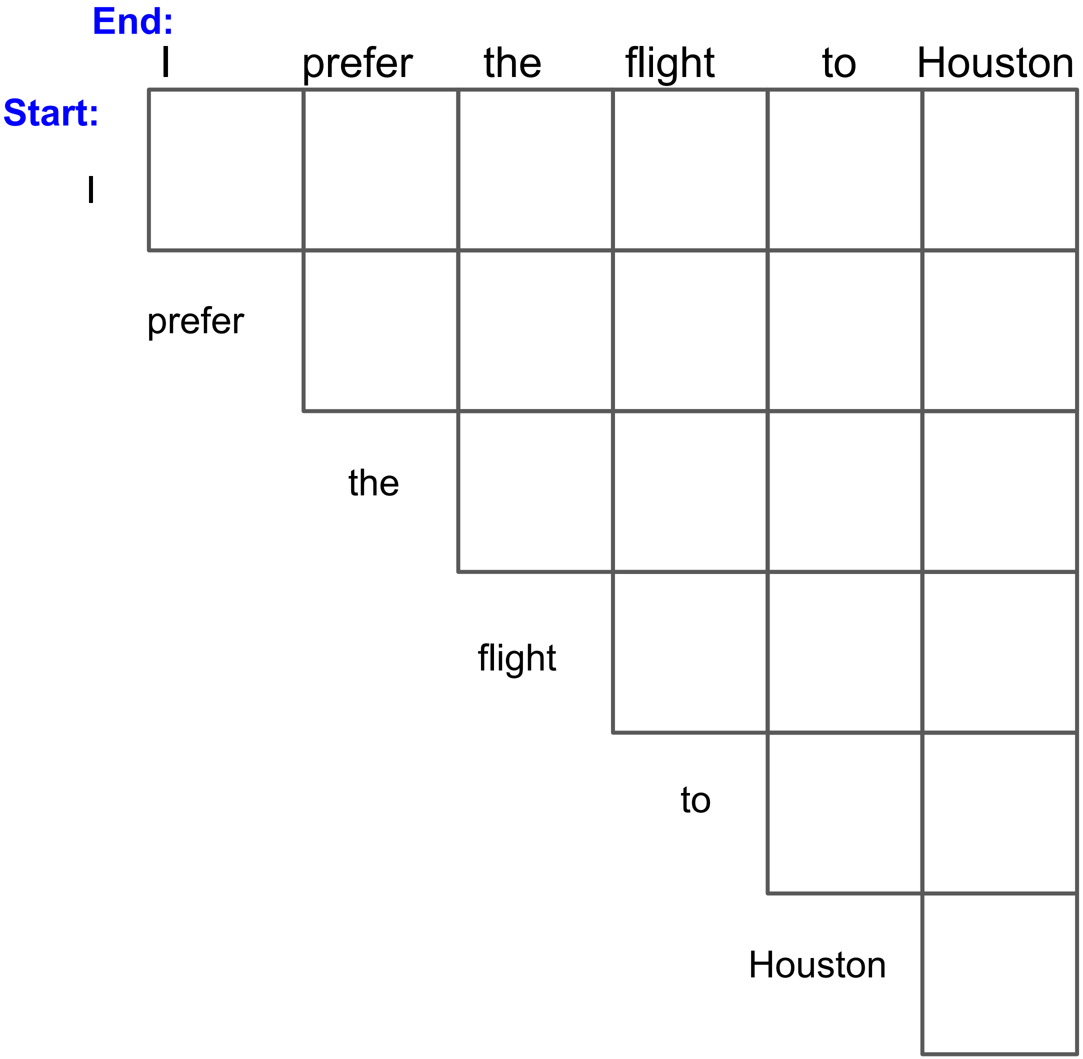
Enumerate split point *s* for state (i, j)



- For all X in N [iterate all non-terminals]

$$\pi(i,j,X) = \max_{\substack{X \to YZ \in R,\\ s \in \{i \ldots (j-1)\}}} \left(q(X \to YZ) \times \pi(i,s,Y) \times \pi(s+1,j,Z) \right)$$

Structure	Sequence	Tree
Algorithm	Viterbi	CKY (or CYK)
State space	|Tokens| x |Tags|	|Spans| x |Symbols| = O(N^2*|Symbols|)
Time complexity	|Tokens| x |Tags|^2	|Spans| x |Tokens| x |Rules| = O(N^3*|Rules|)
Filling the chart	Left-to-right	Bottom-up
Recursive definition	$\pi(i, y_i) = \max_{y_{i-1}} e(x_i \mid y_i) q(y_i \mid y_{i-1}) \pi(i-1, y_{i-1})$	$\pi(i, j, X) = \max_{\substack{X \rightarrow YZ \in R, \\ s \in \{i \ldots (j-1)\}}} (q(X \rightarrow YZ) \times \pi(i, s, Y) \times \pi(s+1, j, Z))$



S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer 0.01 0.004 0.006	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me 0.1 0.02 0.02 0.06	
NP → Houston | NWA 0.16 .04	
Det → the | a | an 0.6 0.1 0.05	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money 0.03 0.15 0.06 0.06	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer 0.5 0.04 0.06	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from 0.2 0.3 0.3	
PP → Prep NP	1.0

Start: \ **End:**	I	prefer	the	flight	to	Houston
I	NP: .1					
prefer		S: .006 Verb: .06				
the			Det: .6			
flight				Nominal: .15		
to					Prep: .3	
Houston						NP: .16

S → NP VP 0.8
S → X1 VP 0.1
X1 → Aux NP 1.0
S → book | include | prefer
0.01 0.004 0.006
S → Verb NP 0.05
S → VP PP 0.03
NP → I | he | she | me
0.1 0.02 0.02 0.06
NP → Houston | NWA
0.16 .04
Det → the | a | an
0.6 0.1 0.05
NP → Det Nominal 0.6
Nominal → book | flight | meal | money
0.03 0.15 0.06 0.06
Nominal → Nominal Nominal 0.2
Nominal → Nominal PP 0.5
Verb → book | include | prefer
0.5 0.04 0.06
VP → Verb NP 0.5
VP → VP PP 0.3
Prep → through | to | from
0.2 0.3 0.3
PP → Prep NP 1.0

Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1					
prefer		S: .006 Verb: .06				
the			Det: .6			
flight				Nominal: .15		
to					Prep: .3	
Houston						NP: .16

Rule	Probability
S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer 0.01 0.004 0.006	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me 0.1 0.02 0.02 0.06	
NP → Houston | NWA 0.16 .04	
Det → the | a | an 0.6 0.1 0.05	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money 0.03 0.15 0.06 0.06	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer 0.5 0.04 0.06	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from 0.2 0.3 0.3	
PP → Prep NP	1.0

Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1					
prefer		S: .006 Verb: .06				
the			Det: .6			
flight				Nominal: .15		
to					Prep: .3	
Houston						NP: .16

Rule	Probability
S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer	0.01 | 0.004 | 0.006
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me	0.1 | 0.02 | 0.02 | 0.06
NP → Houston | NWA	0.16 | .04
Det → the | a | an	0.6 | 0.1 | 0.05
NP → Det Nominal	0.6
Nominal → book | flight | meal | money	0.03 | 0.15 | 0.06 | 0.06
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer	0.5 | 0.04 | 0.06
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from	0.2 | 0.3 | 0.3
PP → Prep NP	1.0

Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1					
prefer		S: .006 Verb: .06				
the			Det: .6	NP: .6*.6 *.15 = **.054**		
flight				Nominal: .15		
to					Prep: .3	
Houston						NP: .16

Rule	Probability
S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer **0.01 0.004 0.006**	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me **0.1 0.02 0.02 0.06**	
NP → Houston | NWA **0.16 .04**	
Det → the | a | an **0.6 0.1 0.05**	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money **0.03 0.15 0.06 0.06**	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer **0.5 0.04 0.06**	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from **0.2 0.3 0.3**	
PP → Prep NP	1.0

Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1					
prefer		S: .006 Verb: .06				
the			Det: .6	NP: .6*.6 *.15 = **.054**		
flight				Nominal: .15		
to					Prep: .3	
Houston						NP: .16

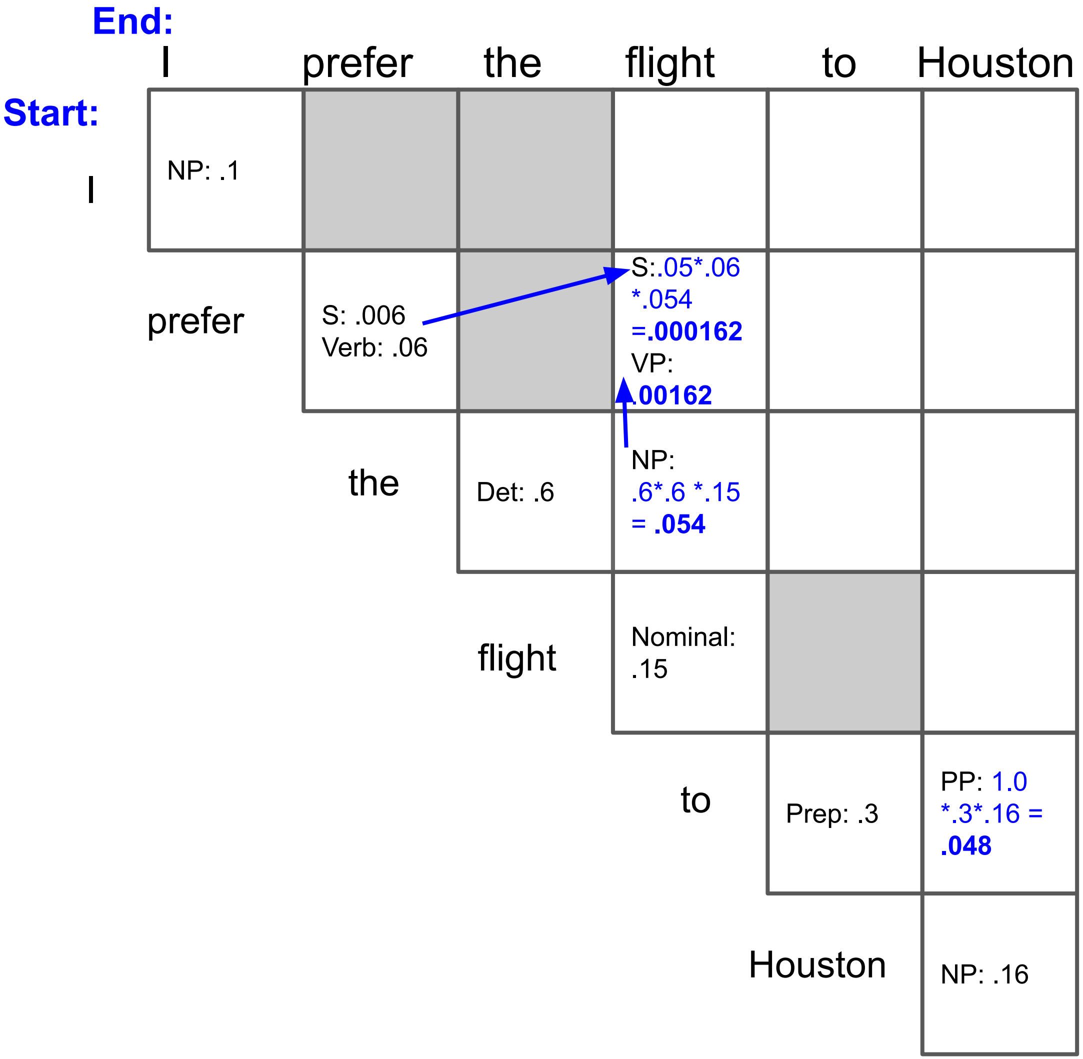
S → NP VP 0.8
S → X1 VP 0.1
X1 → Aux NP 1.0
S → book | include | prefer
0.01 0.004 0.006
S → Verb NP 0.05
S → VP PP 0.03
NP → I | he | she | me
0.1 0.02 0.02 0.06
NP → Houston | NWA
0.16 .04
Det→ the | a | an
0.6 0.1 0.05
NP → Det Nominal 0.6
Nominal → book | flight | meal | money
0.03 0.15 0.06 0.06
Nominal → Nominal Nominal 0.2
Nominal → Nominal PP 0.5
Verb→ book | include | prefer
0.5 0.04 0.06
VP → Verb NP 0.5
VP → VP PP 0.3
Prep → through | to | from
0.2 0.3 0.3
PP → Prep NP 1.0

Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1					
prefer		S: .006 Verb: .06				
the			Det: .6	NP: .6*.6 *.15 = **.054**		
flight				Nominal: .15		
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16

S → NP VP 0.8
S → X1 VP 0.1
X1 → Aux NP 1.0
S → book | include | prefer
0.01 0.004 0.006
S → Verb NP 0.05
S → VP PP 0.03
NP → I | he | she | me
0.1 0.02 0.02 0.06
NP → Houston | NWA
0.16 .04
Det→ the | a | an
0.6 0.1 0.05
NP → Det Nominal 0.6
Nominal → book | flight | meal | money
0.03 0.15 0.06 0.06
Nominal → Nominal Nominal 0.2
Nominal → Nominal PP 0.5
Verb→ book | include | prefer
0.5 0.04 0.06
VP → Verb NP 0.5
VP → VP PP 0.3
Prep → through | to | from
0.2 0.3 0.3
PP → Prep NP 1.0

Start \ End:	I	prefer	the	flight	to	Houston
I	NP: .1					
prefer		S: .006 Verb: .06				
the			Det: .6	NP: .6*.6 *.15 = **.054**		
flight				Nominal: .15		
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16

S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer **0.01 0.004 0.006**	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me **0.1 0.02 0.02 0.06**	
NP → Houston | NWA **0.16 .04**	
Det→ the | a | an **0.6 0.1 0.05**	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money **0.03 0.15 0.06 0.06**	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb→ book | include | prefer **0.5 0.04 0.06**	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from **0.2 0.3 0.3**	
PP → Prep NP	1.0



S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer 0.01 0.004 0.006	
S → Verb NP	**0.05**
S → VP PP	0.03
NP → I | he | she | me 0.1 0.02 0.02 0.06	
NP → Houston | NWA 0.16 .04	
Det → the | a | an 0.6 0.1 0.05	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money 0.03 0.15 0.06 0.06	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer 0.5 0.04 0.06	
VP → Verb NP	**0.5**
VP → VP PP	0.3
Prep → through | to | from 0.2 0.3 0.3	
PP → Prep NP	1.0

Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1					
prefer		S: .006 Verb: .06		S:.05*.06 *.054 =.000162 VP: .00162		
the			Det: .6	NP: .6*.6 *.15 = .054		
flight				Nominal: .15		
to					Prep: .3	PP: 1.0 *.3*.16 = .048
Houston						NP: .16

S → NP VP 0.8
S → X1 VP 0.1
X1 → Aux NP 1.0
S → book | include | prefer
0.01 0.004 0.006
S → Verb NP 0.05
S → VP PP 0.03
NP → I | he | she | me
0.1 0.02 0.02 0.06
NP → Houston | NWA
0.16 .04
Det → the | a | an
0.6 0.1 0.05
NP → Det Nominal 0.6
Nominal → book | flight | meal | money
0.03 0.15 0.06 0.06
Nominal → Nominal Nominal 0.2
Nominal → Nominal PP 0.5
Verb → book | include | prefer
0.5 0.04 0.06
VP → Verb NP 0.5
VP → VP PP 0.3
Prep → through | to | from
0.2 0.3 0.3
PP → Prep NP 1.0

Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1					
prefer		S: .006 Verb: .06		S:.05*.06*.054 =.000162 VP: .00162		
the			Det: .6	NP: .6*.6 *.15 = .054		
flight				Nominal: .15		Nominal: .5*.15*.048=.0036
to					Prep: .3	PP: 1.0 *.3*.16 = .048
Houston						NP: .16

Rule	Probability
S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer	0.01 0.004 0.006
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me	0.1 0.02 0.02 0.06
NP → Houston | NWA	0.16 .04
Det → the | a | an	0.6 0.1 0.05
NP → Det Nominal	0.6
Nominal → book | flight | meal | money	0.03 0.15 0.06 0.06
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	**0.5**
Verb → book | include | prefer	0.5 0.04 0.06
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from	0.2 0.3 0.3
PP → Prep NP	1.0

End:

Start:

	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		
the			Det: .6	NP: .6*.6 *.15 = **.054**		
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16

S → NP VP 0.8
S → X1 VP 0.1
X1 → Aux NP 1.0
S → book | include | prefer
0.01 0.004 0.006
S → Verb NP 0.05
S → VP PP 0.03
NP → I | he | she | me
0.1 0.02 0.02 0.06
NP → Houston | NWA
0.16 .04
Det → the | a | an
0.6 0.1 0.05
NP → Det Nominal 0.6
Nominal → book | flight | meal | money
0.03 0.15 0.06 0.06
Nominal → Nominal Nominal 0.2
Nominal → Nominal PP 0.5
Verb → book | include | prefer
0.5 0.04 0.06
VP → Verb NP 0.5
VP → VP PP 0.3
Prep → through | to | from
0.2 0.3 0.3
PP → Prep NP 1.0

Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		
the			Det: .6	NP: .6*.6 *.15 = **.054**		
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16

Rule	Probability
S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer 0.01 0.004 0.006	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me 0.1 0.02 0.02 0.06	
NP → Houston | NWA 0.16 .04	
Det → the | a | an 0.6 0.1 0.05	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money 0.03 0.15 0.06 0.06	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer 0.5 0.04 0.06	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from 0.2 0.3 0.3	
PP → Prep NP	1.0

End: (columns) / **Start:** (rows)

Start \ End	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16

S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer	0.01 0.004 0.006
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me	0.1 0.02 0.02 0.06
NP → Houston | NWA	0.16 .04
Det → the | a | an	0.6 0.1 0.05
NP → Det Nominal	**0.6**
Nominal → book | flight | meal | money	0.03 0.15 0.06 0.06
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer	0.5 0.04 0.06
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from	0.2 0.3 0.3
PP → Prep NP	1.0

End: / **Start:**

Start \ End	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16

Rule	Probability
S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer	0.01 0.004 0.006
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me	0.1 0.02 0.02 0.06
NP → Houston | NWA	0.16 .04
Det → the | a | an	0.6 0.1 0.05
NP → Det Nominal	0.6
Nominal → book | flight | meal | money	0.03 0.15 0.06 0.06
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer	0.5 0.04 0.06
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from	0.2 0.3 0.3
PP → Prep NP	1.0

Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		S:4e-5 VP:.5*.06*.001296 = **.000039**
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16

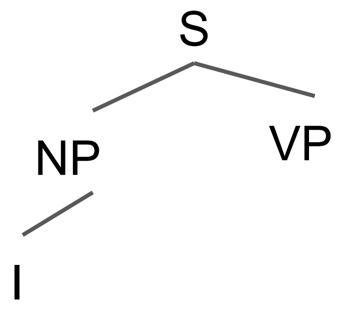
S->Verb NP: **.05*.06*.01296=4e-5**
S->VP PP: .03*.00162*.48=2e-5
VP->Verb NP: **.5*.06*.01296=.00039**
VP->VP PP: .3*.00162*.48=.00023

S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer (0.01, 0.004, 0.006)	
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me (0.1, 0.02, 0.02, 0.06)	
NP → Houston | NWA (0.16, .04)	
Det → the | a | an (0.6, 0.1, 0.05)	
NP → Det Nominal	0.6
Nominal → book | flight | meal | money (0.03, 0.15, 0.06, 0.06)	
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer (0.5, 0.04, 0.06)	
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from (0.2, 0.3, 0.3)	
PP → Prep NP	1.0

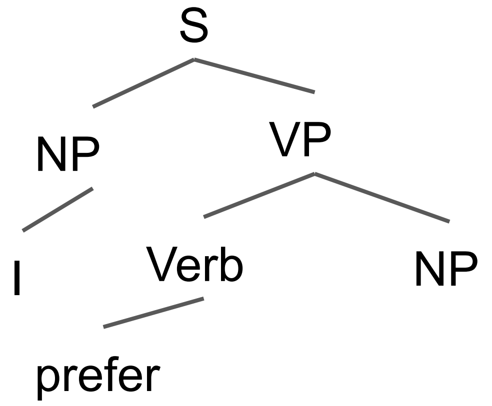
Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		S: .8*.1*.000039=**3e-6**
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		S:4e-5 VP:.5*.06*.001296 = **.000039**
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16

Rule	Probability
S → NP VP	0.8
S → X1 VP	0.1
X1 → Aux NP	1.0
S → book | include | prefer	0.01 0.004 0.006
S → Verb NP	0.05
S → VP PP	0.03
NP → I | he | she | me	0.1 0.02 0.02 0.06
NP → Houston | NWA	0.16 .04
Det → the | a | an	0.6 0.1 0.05
NP → Det Nominal	0.6
Nominal → book | flight | meal | money	0.03 0.15 0.06 0.06
Nominal → Nominal Nominal	0.2
Nominal → Nominal PP	0.5
Verb → book | include | prefer	0.5 0.04 0.06
VP → Verb NP	0.5
VP → VP PP	0.3
Prep → through | to | from	0.2 0.3 0.3
PP → Prep NP	1.0

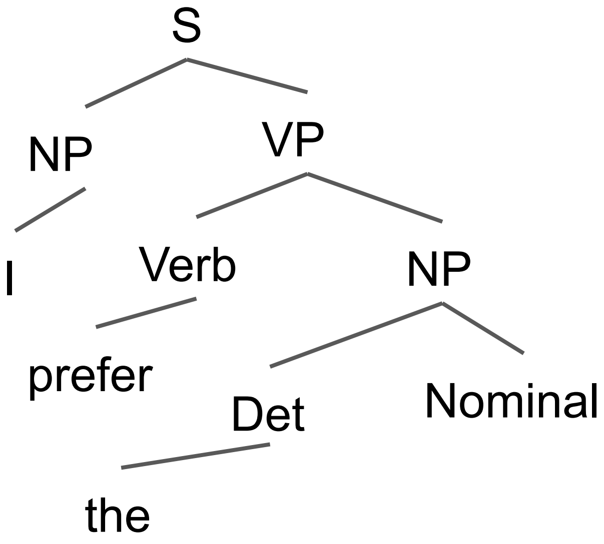
Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		S: .8*.1*.000039=**3e-6**
prefer		S: .006 Verb: .06		S:.05*.06 *.054 =**.000162** VP: **.00162**		S:4e-5 VP:.5*.06 *.001296 = **.000039**
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16



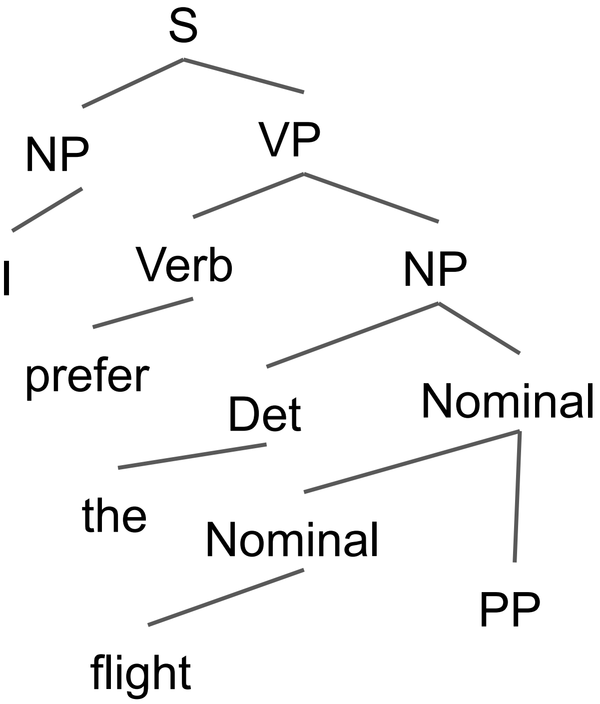
Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		S: .8*.1*.000039=**3e-6**
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		S:4e-5 VP:.5*.06*.001296 = **.000039**
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16



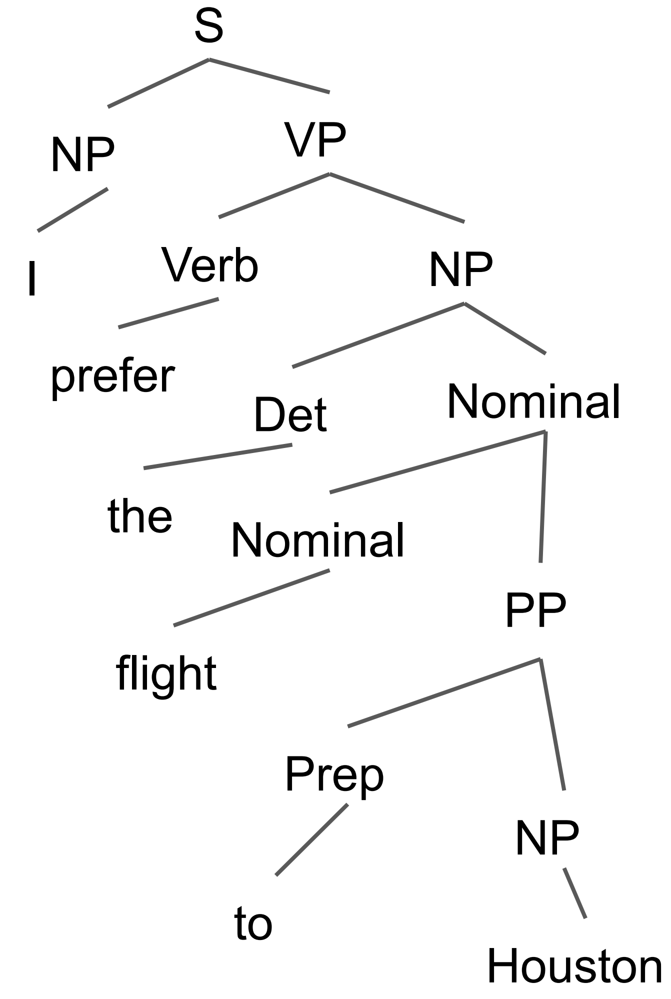
Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		S: .8*.1*.000039=**3e-6**
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		S:4e-5 VP:.5*.06*.001296 = **.000039**
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16



Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		S: .8*.1*.000039=**3e-6**
prefer		S: .006 Verb: .06		S:.05*.06 *.054 =**.000162** VP: **.00162**		S:4e-5 VP:.5*.06 *.001296 = **.000039**
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16



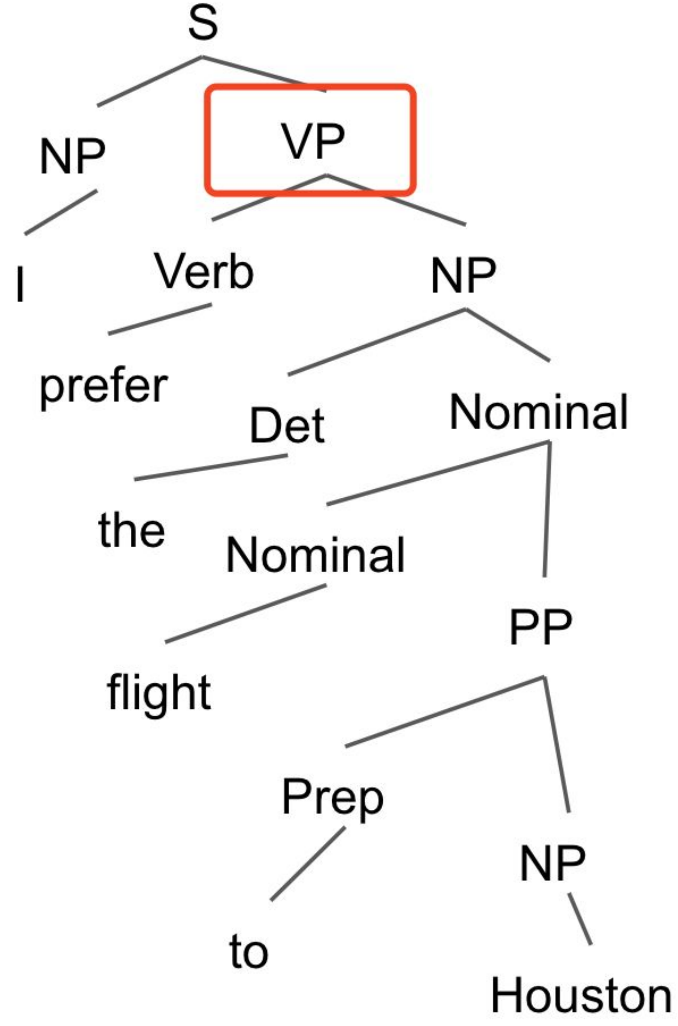
Start: \ End:	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		S: .8*.1*.000039=**3e-6**
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		S:4e-5 VP:.5*.06*.001296 = **.000039**
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0*.3*.16 = **.048**
Houston						NP: .16



Quiz 7, Part I CKY

1. For the tree you build, what is the nonterminal for the constituent “prefer the flight to Houston”?

   VP



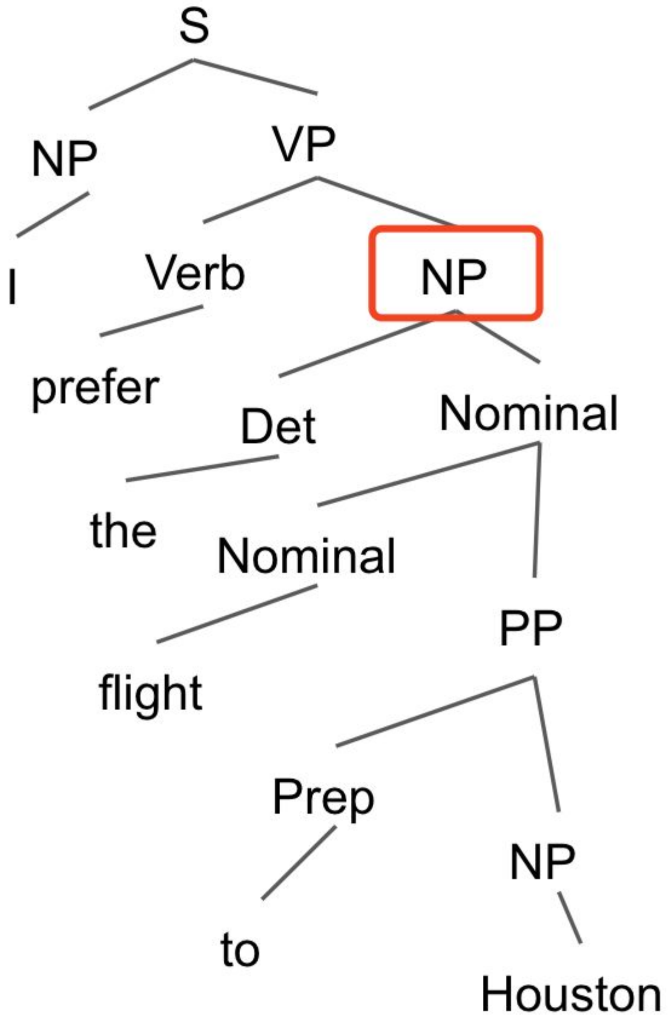
Quiz 7, Part I CKY

1. For the tree you build, what is the nonterminal for the constituent “prefer the flight to Houston”?

   VP

2. If “prefer” has a “Verb” tag, what could be a possible nonterminal for the constituent “the flight to Houston”?

   NP



Quiz 7, Part I CKY

1. For the tree you build, what is the nonterminal for the constituent "prefer the flight to Houston"?

   VP

2. If "prefer" has a "Verb" tag, what could be a possible nonterminal for the constituent "the flight to Houston"?

   NP

3. If "prefer the flight" has a "VP" tag, what could be a possible nonterminal for the constituent "to Houston"?

   PP

End: / **Start:**

Start \ End	I	prefer	the	flight	to	Houston
I	NP: .1			S: .8*.1*.00162= **.00013**		
prefer		S: .006 Verb: .06		S:.05*.06*.054 =**.000162** VP: **.00162**		
the			Det: .6	NP: .6*.6 *.15 = **.054**		NP: .6*.6*.0036 = **.001296**
flight				Nominal: .15		Nominal: .5*.15*.048=**.0036**
to					Prep: .3	PP: 1.0 *.3*.16 = **.048**
Houston						NP: .16

Semantics

What Is Semantics?

- Syntax concerns the structure of sentences, i.e., does a sentence conform to the rules of a language?
    - PL parallel: “Does a program compile?”
- Semantics concerns the meaning of sentences, usually using syntax as a scaffold
    - PL parallel: “What is the output of a program?”
- There are syntactically well-formed yet semantically infelicitous sentences...
    - Famous example due to Chomsky: “Colorless green ideas sleep furiously.”
    - “The present King of France is bald.”
- ... and arguably semantically felicitous yet syntactically ill-formed ones
    - “Dog loyal animal.”
        - Cf. “Dogs are loyal animals.”

First-Order logic

- **Term:** a constant (ss) or a variable
- **Formula:** defined inductively ...
  - If R is an n-ary relation and $t_1, \ldots, t_n$ are terms, then $R(t_1, \ldots, t_n)$ is a formula.
  - If ϕ is a formula, then its negation, $\neg\phi$, is a formula.
  - If ϕ and ψ are formulas, then binary logical connectives can be used to create formulas:
    - $\phi \wedge \psi$
    - $\phi \vee \psi$
    - $\phi \Rightarrow \psi$
  - If ϕ is a formula and v is a variable, then quantifiers can be used to create formulas:
    - Universal quantifier: $\forall v, \phi$
    - Existential quantifier: $\exists v, \phi$

From lecture slide 80

Bonus Question → English Translation

Q3: $\forall x, Quiz(x) \Rightarrow (Hard(x) \lor \neg Does(Adrian, x))$

A3: Every quiz is hard or is not done by Adrian.

Q4: $\exists y, \forall x, Quiz(y) \Rightarrow \neg Does(x, y)$

A4: There exists some quizzes that no one does

OR not everything in the world is a quiz!

Q4’: $\exists y, \forall x, Quiz(y) \land \neg Does(x, y)$

Bonus English → FOL Translation

- “Every farmer who owns a donkey beats it.”
- This type of sentence, called “donkey sentences”, is well-known in semantic theory
- Answer in https://en.wikipedia.org/wiki/Donkey_sentence (but try it first -- it’s fun!)

Compositional Semantics

- How do we get to FOL (or some other meaning representation) from natural language?
- Theory: compositionality
  - Human can't remember the meaning of all sentences, as there are infinitely many of them (e.g., recursion)
  - So sentence meanings are composed from smaller parts, with rules stitching them together
    - Often on top of the syntax tree
  - The most basic pieces of meaning come from the lexicon
- Tool: λ-calculus

λ-Calculus

- Anonymized functions; like the lambdas in Python/Java/etc.
- Syntax of λ-calculus (which is in a sense arbitrary): λv . φ
  - In Python: `lambda v: φ`
- Usually, instead of a value as the output (e.g., λx . x + 1), we have a statement (e.g., λx . x runs)
  - You can think of the statement as a "truth-condition" that evaluates to either true or false given the world state
  - So the output is just a special type of object, if you will
- In compositional semantics, as you walk up the tree, the argument(s) of the λs get filled in and (necessarily) ending up with a arity/valency of 0 at the root

John: John
Mary: Mary
Likes: λx . λy . y likes x

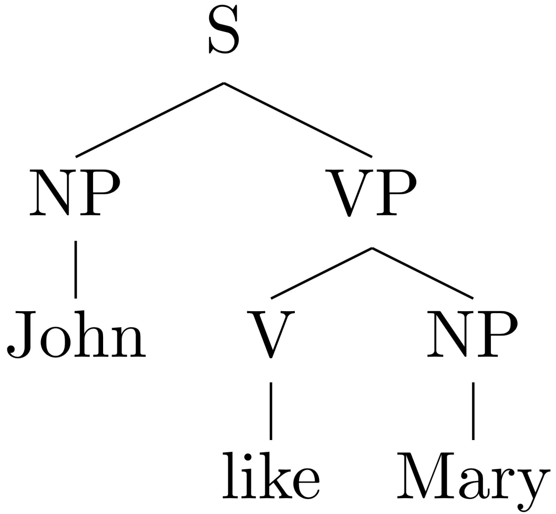
Example: Compositional Semantics

- Input: “John likes Mary”
- Goal: The input is true iff John likes Mary
    - Not so interesting in this case, but we still want to see how we derive this compositionally

John: John
Mary: Mary
Likes: λx . λy . y likes x

Example: Compositional Semantics

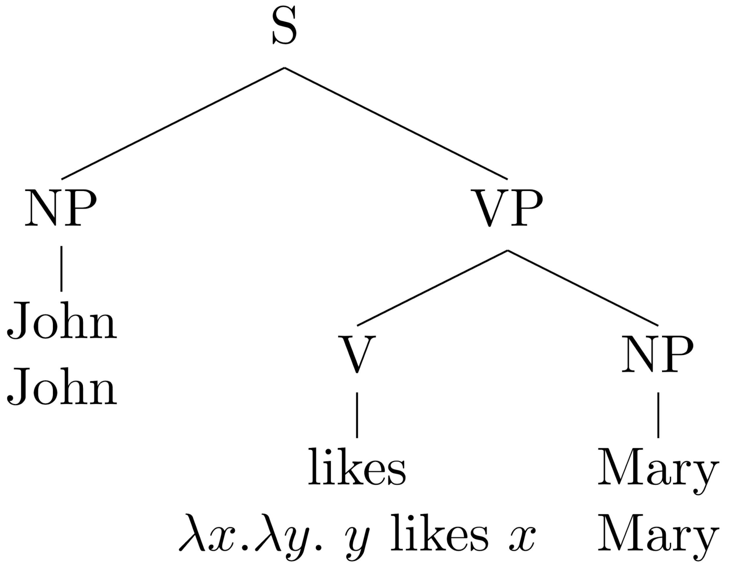
- Input: “John likes Mary”
- Goal: The input is true iff John likes Mary
    - Not so interesting in this case, but we still want to see how we derive this compositionally
- Step 1: parse



John: John
Mary: Mary
Likes: λx . λy . y likes x

Example: Compositional Semantics

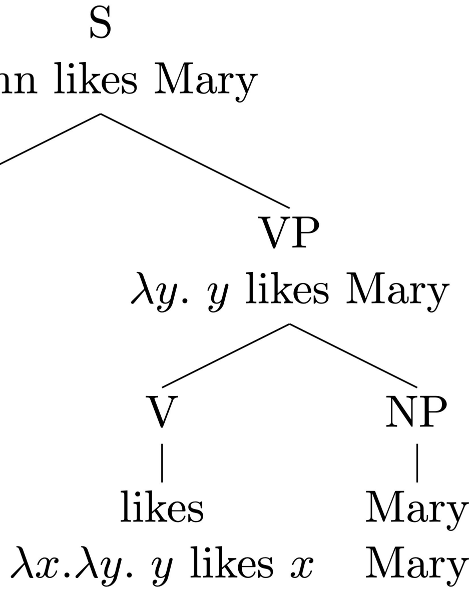
- Input: “John likes Mary”
- Goal: The input is true iff John likes Mary
    - Not so interesting in this case, but we still want to see how we derive this compositionally
- Step 2: apply the lexicon



John: John
Mary: Mary
Likes: λx . λy . y likes x

Example: Compositional Semantics

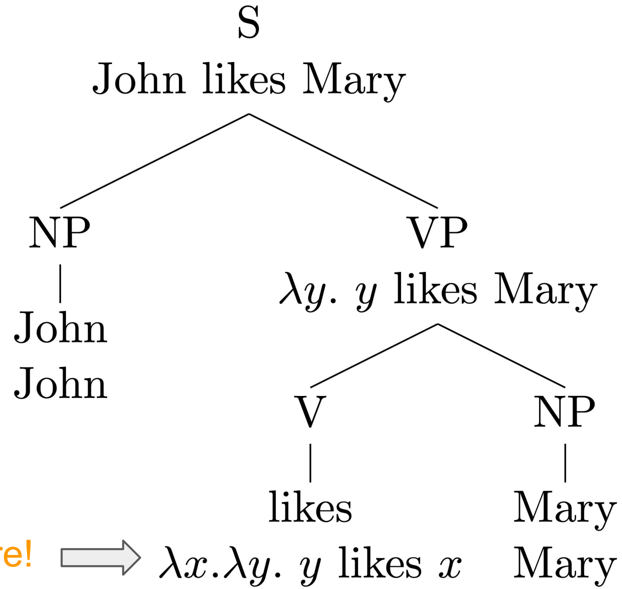
- Input: "John likes Mary"
- Goal: The input is true iff John likes Mary
    - Not so interesting in this case, but we still want to see how we derive this compositionally
- Step 3: compose



John: John
Mary: Mary
Likes: λx . λy . y likes x

Example: Compositional Semantics

- Input: “John likes Mary”
- Goal: The input is true iff John likes Mary
  - Not so interesting in this case, but we still want to see how we derive this compositionally
- Step 3: compose



Note the importance of the order here!

Q & A