Natural Language Processing

Constituency Parsing

Joakim Nivre

Uppsala University
Department of Linguistics and Philology
joakim.nivre@lingfil.uu.se

Context-Free Grammar

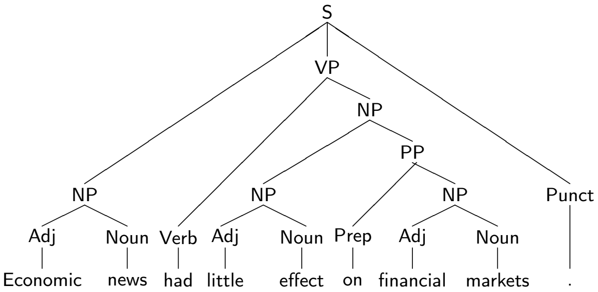
- A context-free grammar (CFG) consists of
  - a finite set of nonterminal symbols
  - a finite set of terminal symbols
  - a distinguished nonterminal symbol S (for Start)
  - a finite set of rules of the form

$$A \to \alpha$$

  where A is a nonterminal and α is a (possibly empty) sequence of nonterminal and terminal symbols

Context-Free Grammar

S	→	NP VP Punct	PP	→	Prep NP	Adj	→	Economic
VP	→	VP PP	Verb	→	had	Adj	→	little
VP	→	Verb NP	Noun	→	news	Adj	→	financial
NP	→	NP PP	Noun	→	effect	Prep	→	on
NP	→	Adj Noun	Noun	→	markets	Punct	→	.



CFG Parsing

- Input:
  - Sentence: $S = w_1, \ldots, w_n$
  - Grammar: G
- Output:
  - Phrase structure tree for S generated by G
- Parsing:
  - Construct a derivation of S in G
  - Read off phrase structure tree from derivation

Quiz

S	→	NP VP Punct	PP	→	Prep NP	Adj	→	Economic
VP	→	VP PP	Verb	→	had	Adj	→	little
VP	→	Verb NP	Noun	→	news	Adj	→	financial
NP	→	NP PP	Noun	→	effect	Prep	→	on
NP	→	Adj Noun	Noun	→	markets	Punct	→	.

- Which sentences can be derived in the grammar above?
  1. Economic news had effect .
  2. Economic news had little effect .
  3. Economic news had little effect on markets.
  4. Economic news had little effect on little effect on little effect .

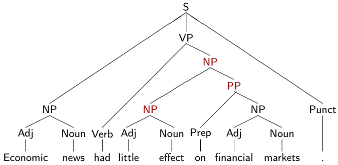
Top-Down Parsing

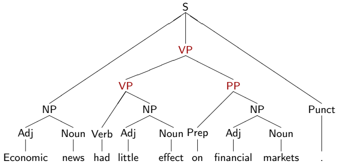
- Basic idea:
  - Start at the root node (start symbol)
  - Expand tree by matching the left-hand side of rules
  - Derive a tree whose terminal nodes match the input sentence
- Potential problems:
  - May use rules that could never match the input
  - May loop on recursive rules: VP $\rightarrow$ VP PP

Bottom-Up Parsing

- Basic idea:
  - Start with the terminal nodes (words)
  - Expand tree by matching the right-hand side of rules
  - Build a tree whose with the start symbol at the root
- Potential problems:
  - May build structures that could never be in a tree
  - May loop on empty productions NP $\rightarrow \epsilon$

Ambiguity





- A sentence may be assigned more than one tree
- This corresponds to syntactic ambiguity
- What should the parser do in this case?

Dealing with Ambiguity

- Combinatorial explosion:
  - The number of possible trees grows exponentially
  - We need a smart algorithm – dynamic programming
- Disambiguation:
  - We need some way of selecting the best tree
  - Probabilistic grammars rank trees by decreasing plausibility

Quiz

- Which of the following sentences are syntactically ambiguous?
  1. I caught a bat
  2. I caught a bat with a hat
  3. I caught a bat happily
  4. I caught small bats and rats

Natural Language Processing

Probabilistic Context-Free Grammar

Joakim Nivre

Uppsala University
Department of Linguistics and Philology
joakim.nivre@lingfil.uu.se

Probabilistic Context-Free Grammar

$$G = (N, \Sigma, R, S, Q)$$

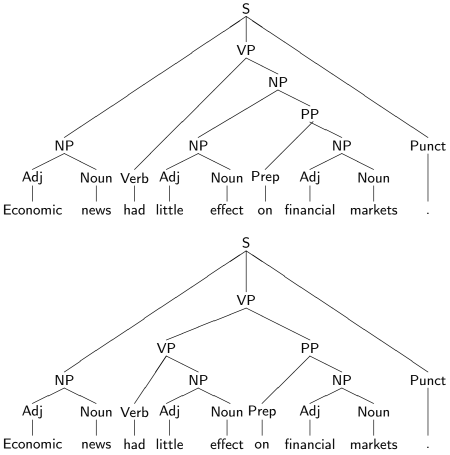
- N is a finite (non-terminal) alphabet
- Σ is a finite (terminal) alphabet
- R is a finite set of rules $A \to \alpha$ ($A \in N$, $\alpha \in (\Sigma \cup N)^*$)
- $S \in N$ is the start symbol

- Q is function from R to real numbers in the interval $[0, 1]$
  - For every nonterminal $A \in N$:

$$\sum_{r \in R\,:\,\mathbf{LHS}(r) = A} Q(r) = 1$$

Example Grammar

S	→	NP VP Punct	1.00
VP	→	VP PP	0.50
VP	→	Verb NP	0.50
NP	→	NP PP	0.25
NP	→	Adj Noun	0.75
PP	→	Prep NP	1.00
Punct	→	.	1.00
Adj	→	Economic	0.33
Adj	→	little	0.33
Adj	→	financial	0.33
Noun	→	news	0.33
Noun	→	effect	0.33
Noun	→	markets	0.33
Verb	→	had	1.00
Prep	→	on	1.00



Probability Model for Trees

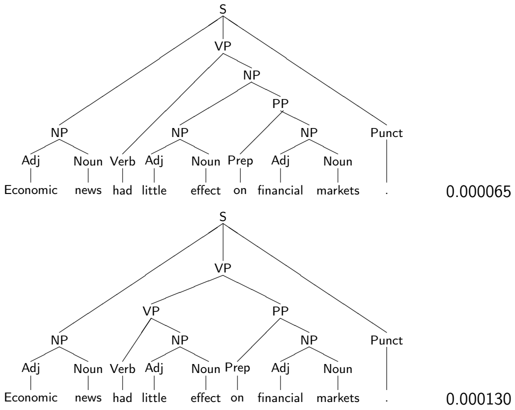
- Probability of a tree T generated by G:

$$P(T) = \prod_{r \in R} Q(r)^{f(r,T)}$$

  - $f(r, T)$ = the frequency of r in T
- Intuition:
  - The tree probability is the joint probability of rule instances
  - Rule instances are dependent only on their left-hand side
  - Thus, the joint probability is the product of rule probabilities

Example Trees

S	→	NP VP Punct	1.00
VP	→	VP PP	0.50
VP	→	Verb NP	0.50
NP	→	NP PP	0.25
NP	→	Adj Noun	0.75
PP	→	Prep NP	1.00
Punct	→	.	1.00
Adj	→	Economic	0.33
Adj	→	little	0.33
Adj	→	financial	0.33
Noun	→	news	0.33
Noun	→	effect	0.33
Noun	→	markets	0.33
Verb	→	had	1.00
Prep	→	on	1.00



Quiz

NP	$\rightarrow$	Noun	0.1
NP	$\rightarrow$	Det Noun	0.4
NP	$\rightarrow$	Adj Noun	0.2
NP	$\rightarrow$	Det Adj Noun	?.?

- If these are all the NP rules, what is the missing probability?
  1. 0.1
  2. 0.3
  3. 0.6
  4. 1.0

Treebank Grammars

- Given a treebank $\mathcal{T} = \{(S_1, T_1), \ldots, (S_n, T_n)\}$
- Extract grammar $G = (N, \Sigma, R, S)$:
  - N = the set of all nonterminals occurring in some tree
  - Σ = the set of all terminals occurring in some tree
  - R = the set of all rules needed to derive some tree
  - S = the nonterminal at the root of every tree
- Estimate Q using relative frequencies (MLE):

$$Q(A \rightarrow \alpha) = \frac{\sum_{i=1}^{n} f(A \rightarrow \alpha, T_i)}{\sum_{i=1}^{n} \sum_{r \in R:\textsc{lhd}(r)=A} f(r, T_i)}$$

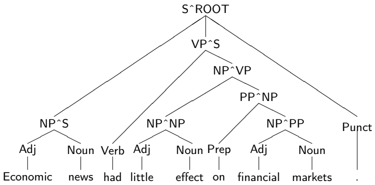
Limitations of Treebank Grammars

- Insensitive to structural context:

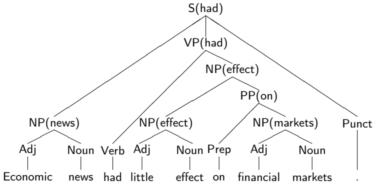
Tree Context	NP PP	DT NN	PRP
Anywhere	11%	9%	6%
NP under S	9%	9%	21%
NP under VP	23%	7%	4%

- Insensitive to lexical information:
  - She ate pizza with mushrooms
  - She ate pizza with a fork

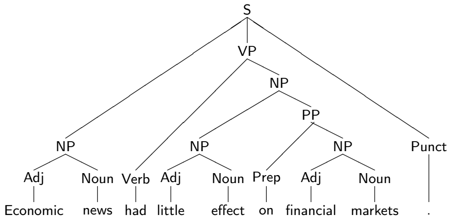
Parent Annotation



Lexicalization



Quiz



- Parent annotation increases the number of symbols.
- How many new NP symbols do we need for the tree above?
  1. 2
  2. 3
  3. 4
  4. 5

Natural Language Processing

PCFG Parsing

Joakim Nivre

Uppsala University
Department of Linguistics and Philology
joakim.nivre@lingfil.uu.se

PCFG Parsing

- Given a sentence S and grammar G
  - derive all possible trees for S according to G
  - find most probable tree T^*
- We can use dynamic programming to solve this efficiently
  - The Cocke-Kasami-Younger (CKY) algorithm
  - Earley's algorithm

Probabilistic CKY

0	I	1	saw	2	her	3	duck	4

- Parse tables:

  $\mathcal{C}[i, j, B]$ = probability of best phrase of type B from i to j
  $\mathcal{B}[i, j, B]$ = back-pointers to retrieve best B from i to j

- CKY assumes Chomsky Normal Form (CNF):

  $A \rightarrow B\ C$
  $A \rightarrow a$

- Any CFG can be converted to CNF

Probabilistic CKY

PARSE(G, x)

for j from 1 to n do # j = end of word

for all $A : A \rightarrow a \in R$ and $a = {}_{j-1}w_j$

$\mathcal{C}[j-1, j, A] := Q(A \rightarrow a)$

Probabilistic CKY

PARSE(G, x)

for j from 1 to n do # j = end of word

  for all $A : A \rightarrow a \in R$ and $a = {}_{j-1}w_j$

    $\mathcal{C}[j-1, j, A] := Q(A \rightarrow a)$

for j from 2 to n do # j = end of phrase

Probabilistic CKY

PARSE(G, x)

for j from 1 to n do # j = end of word

  for all $A : A \rightarrow a \in R$ and $a = {}_{j-1}w_j$

    $\mathcal{C}[j-1, j, A] := Q(A \rightarrow a)$

for j from 2 to n do # j = end of phrase

  for i from $j-2$ downto 0 do # i = start of phrase

Probabilistic CKY

PARSE(G, x)

for j from 1 to n do # j = end of word

    for all $A : A \rightarrow a \in R$ and $a = {}_{j-1}w_j$

        $\mathcal{C}[j-1, j, A] := Q(A \rightarrow a)$

for j from 2 to n do # j = end of phrase

    for i from $j-2$ downto 0 do # i = start of phrase

        for k from $i+1$ to $j-1$ do # k = split point

Probabilistic CKY

PARSE(G, x)

for j from 1 to n do # j = end of word

for all $A : A \rightarrow a \in R$ and $a = {}_{j-1}w_j$

$\mathcal{C}[j-1, j, A] := Q(A \rightarrow a)$

for j from 2 to n do # j = end of phrase

for i from $j-2$ downto 0 do # i = start of phrase

for k from $i+1$ to $j-1$ do # k = split point

for all $A : A \rightarrow BC \in R$ and $\mathcal{C}[i, k, B] > 0$ and $\mathcal{C}[k, j, C] > 0$

Probabilistic CKY

PARSE(G, x)

for j from 1 to n do # j = end of word

for all $A : A \rightarrow a \in R$ and $a = {}_{j-1}w_j$

$\mathcal{C}[j-1, j, A] := Q(A \rightarrow a)$

for j from 2 to n do # j = end of phrase

for i from $j-2$ downto 0 do # i = start of phrase

for k from $i+1$ to $j-1$ do # k = split point

for all $A : A \rightarrow BC \in R$ and $\mathcal{C}[i, k, B] > 0$ and $\mathcal{C}[k, j, C] > 0$

if $(\mathcal{C}[i, j, A] < Q(A \rightarrow BC) \cdot \mathcal{C}[i, k, B] \cdot \mathcal{C}[k, j, C])$ then

Probabilistic CKY

PARSE(G, x)

for j from 1 to n do # j = end of word

  for all $A : A \rightarrow a \in R$ and $a = {}_{j-1}w_j$

    $\mathcal{C}[j-1, j, A] := Q(A \rightarrow a)$

for j from 2 to n do # j = end of phrase

  for i from $j-2$ downto 0 do # i = start of phrase

    for k from $i+1$ to $j-1$ do # k = split point

      for all $A : A \rightarrow BC \in R$ and $\mathcal{C}[i, k, B] > 0$ and $\mathcal{C}[k, j, C] > 0$

        if $(\mathcal{C}[i, j, A] < Q(A \rightarrow BC) \cdot \mathcal{C}[i, k, B] \cdot \mathcal{C}[k, j, C])$ then

          $\mathcal{C}[i, j, A] := Q(A \rightarrow BC) \cdot \mathcal{C}[i, k, B] \cdot \mathcal{C}[k, j, C]$

Probabilistic CKY

PARSE(G, x)

for j from 1 to n do # j = end of word

  for all $A : A \rightarrow a \in R$ and $a = {}_{j-1}w_j$

    $\mathcal{C}[j-1, j, A] := Q(A \rightarrow a)$

for j from 2 to n do # j = end of phrase

  for i from $j-2$ downto 0 do # i = start of phrase

    for k from $i+1$ to $j-1$ do # k = split point

      for all $A : A \rightarrow BC \in R$ and $\mathcal{C}[i,k,B] > 0$ and $\mathcal{C}[k,j,C] > 0$

        if $(\mathcal{C}[i,j,A] < Q(A \rightarrow BC) \cdot \mathcal{C}[i,k,B] \cdot \mathcal{C}[k,j,C])$ then

          $\mathcal{C}[i,j,A] := Q(A \rightarrow BC) \cdot \mathcal{C}[i,k,B] \cdot \mathcal{C}[k,j,C]$

          $\mathcal{B}[i,j,A] := \{k, B, C\}$

Probabilistic CKY

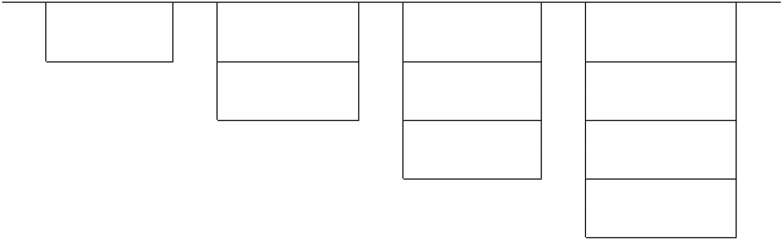
PARSE(G, x)

for j from 1 to n do # j = end of word
    for all $A : A \rightarrow a \in R$ and $a = {}_{j-1}w_j$
        $\mathcal{C}[j-1, j, A] := Q(A \rightarrow a)$
for j from 2 to n do # j = end of phrase
    for i from $j-2$ downto 0 do # i = start of phrase
        for k from $i+1$ to $j-1$ do # k = split point
            for all $A : A \rightarrow BC \in R$ and $\mathcal{C}[i, k, B] > 0$ and $\mathcal{C}[k, j, C] > 0$
                if $(\mathcal{C}[i, j, A] < Q(A \rightarrow BC) \cdot \mathcal{C}[i, k, B] \cdot \mathcal{C}[k, j, C])$ then
                    $\mathcal{C}[i, j, A] := Q(A \rightarrow BC) \cdot \mathcal{C}[i, k, B] \cdot \mathcal{C}[k, j, C]$
                    $\mathcal{B}[i, j, A] := \{k, B, C\}$
return BUILD-TREE($\mathcal{B}[0, n, S]$), $\mathcal{C}[0, n, S]$

Probabilistic CKY

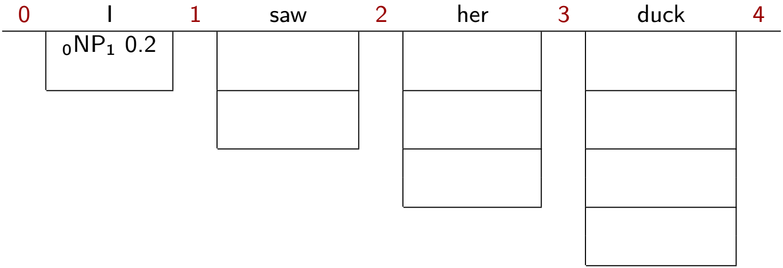
S	→	NP VP	1.0	NP	→	her	0.2
VP	→	Verb NP	0.6	NP	→	I	0.2
VP	→	Verb S	0.3	Verb	→	saw	1.0
VP	→	duck	0.1	Noun	→	duck	1.0
NP	→	Det Noun	0.6	Det	→	her	1.0

0 I 1 saw 2 her 3 duck 4



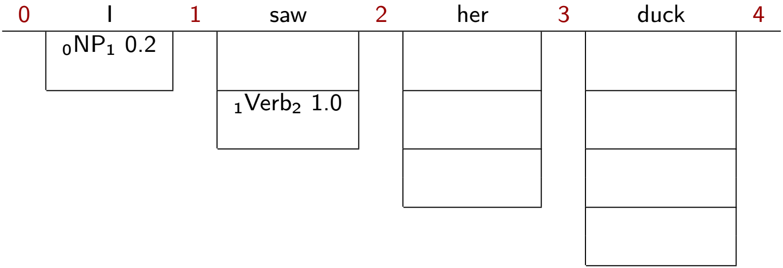
Probabilistic CKY

S	→	NP VP	1.0	NP	→	her	0.2
VP	→	Verb NP	0.6	NP	→	I	0.2
VP	→	Verb S	0.3	Verb	→	saw	1.0
VP	→	duck	0.1	Noun	→	duck	1.0
NP	→	Det Noun	0.6	Det	→	her	1.0



Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0



Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0

0 I 1	saw 2	her 3	duck 4
${}_0\text{NP}_1$ 0.2			
	${}_1\text{Verb}_2$ 1.0		
		${}_2\text{NP}_3$ 0.2	

Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0

0	I	1	saw	2	her	3	duck	4
	${}_0NP_1$ 0.2							
			${}_1Verb_2$ 1.0					
					${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0			

Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0

0 I 1 saw 2 her 3 duck 4

I	saw	her	duck
${}_0NP_1$ 0.2			
	${}_1Verb_2$ 1.0		
		${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0	
			${}_3VP_4$ 0.1

Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0

0	I	1	saw	2	her	3	duck	4
	$_0\text{NP}_1$ 0.2							
			$_1\text{Verb}_2$ 1.0					
					$_2\text{NP}_3$ 0.2 $_2\text{Det}_3$ 1.0			
							$_3\text{VP}_4$ 0.1 $_3\text{Noun}_4$ 1.0	

Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0

0 I 1	1 saw 2	2 her 3	3 duck 4
${}_0NP_1$ 0.2			
	${}_1Verb_2$ 1.0	${}_1VP_3$ 0.12	
		${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0	
			${}_3VP_4$ 0.1 ${}_3Noun_4$ 1.0

Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0

0 I 1	saw 2	her 3	duck 4
${}_0\mathrm{NP}_1$ 0.2		${}_0\mathrm{S}_3$ 0.024	
	${}_1\mathrm{Verb}_2$ 1.0	${}_1\mathrm{VP}_3$ 0.12	
		${}_2\mathrm{NP}_3$ 0.2 ${}_2\mathrm{Det}_3$ 1.0	
			${}_3\mathrm{VP}_4$ 0.1 ${}_3\mathrm{Noun}_4$ 1.0

Probabilistic CKY

S	$\rightarrow$	NP VP	1.0		NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6		NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3		Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1		Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6		Det	$\rightarrow$	her	1.0

0 I 1 saw 2 her 3 duck 4

I	saw	her	duck
${}_0NP_1$ 0.2		${}_0S_3$ 0.024	
	${}_1Verb_2$ 1.0	${}_1VP_3$ 0.12	
		${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0	${}_2NP_4$ 0.6
			${}_3VP_4$ 0.1 ${}_3Noun_4$ 1.0

Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0

0 I 1	saw 2	her 3	duck 4
${}_0NP_1$ 0.2		${}_0S_3$ 0.024	
	${}_1Verb_2$ 1.0	${}_1VP_3$ 0.12	
		${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0	${}_2NP_4$ 0.6 ${}_2S_4$ 0.02
			${}_3VP_4$ 0.1 ${}_3Noun_4$ 1.0

Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0

0	I	1	saw	2	her	3	duck	4
	${}_0NP_1$ 0.2				${}_0S_3$ 0.024			
			${}_1Verb_2$ 1.0		${}_1VP_3$ 0.12		${}_1VP_4$ 0.36	
					${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0		${}_2NP_4$ 0.6 ${}_2S_4$ 0.02	
							${}_3VP_4$ 0.1 ${}_3Noun_4$ 1.0	

Probabilistic CKY

S	→	NP VP	1.0	NP	→	her	0.2
VP	→	Verb NP	0.6	NP	→	I	0.2
VP	→	Verb S	0.3	Verb	→	saw	1.0
VP	→	duck	0.1	Noun	→	duck	1.0
NP	→	Det Noun	0.6	Det	→	her	1.0

0 I 1 saw 2 her 3 duck 4

I	saw	her	duck
${}_0NP_1$ 0.2		${}_0S_3$ 0.024	
	${}_1Verb_2$ 1.0	${}_1VP_3$ 0.12	${}_1VP_4$ 0.36 ${}_1VP_4$ 0.006
		${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0	${}_2NP_4$ 0.6 ${}_2S_4$ 0.02
			${}_3VP_4$ 0.1 ${}_3Noun_4$ 1.0

Probabilistic CKY

S	$\rightarrow$	NP VP	1.0	NP	$\rightarrow$	her	0.2
VP	$\rightarrow$	Verb NP	0.6	NP	$\rightarrow$	I	0.2
VP	$\rightarrow$	Verb S	0.3	Verb	$\rightarrow$	saw	1.0
VP	$\rightarrow$	duck	0.1	Noun	$\rightarrow$	duck	1.0
NP	$\rightarrow$	Det Noun	0.6	Det	$\rightarrow$	her	1.0

0 I 1 saw 2 her 3 duck 4

I	saw	her	duck
${}_0NP_1$ 0.2		${}_0S_3$ 0.024	
	${}_1Verb_2$ 1.0	${}_1VP_3$ 0.12	${}_1VP_4$ 0.36
		${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0	${}_2NP_4$ 0.6 ${}_2S_4$ 0.02
			${}_3VP_4$ 0.1 ${}_3Noun_4$ 1.0

Probabilistic CKY

S	→	NP VP	1.0	NP	→	her	0.2
VP	→	Verb NP	0.6	NP	→	I	0.2
VP	→	Verb S	0.3	Verb	→	saw	1.0
VP	→	duck	0.1	Noun	→	duck	1.0
NP	→	Det Noun	0.6	Det	→	her	1.0

0 I 1	1 saw 2	2 her 3	3 duck 4
${}_0NP_1$ 0.2		${}_0S_3$ 0.024	${}_0S_4$ 0.072
	${}_1Verb_2$ 1.0	${}_1VP_3$ 0.12	${}_1VP_4$ 0.36
		${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0	${}_2NP_4$ 0.6 ${}_2S_4$ 0.02
			${}_3VP_4$ 0.1 ${}_3Noun_4$ 1.0

Probabilistic CKY

S	→	NP VP	1.0	NP	→	her	0.2
VP	→	Verb NP	0.6	NP	→	I	0.2
VP	→	Verb S	0.3	Verb	→	saw	1.0
VP	→	duck	0.1	Noun	→	duck	1.0
NP	→	Det Noun	0.6	Det	→	her	1.0

0 I 1	saw 2	her 3	duck 4
${}_0NP_1$ 0.2		${}_0S_3$ 0.024	${}_0S_4$ 0.072
	${}_1Verb_2$ 1.0	${}_1VP_3$ 0.12	${}_1VP_4$ 0.36
		${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0	${}_2NP_4$ 0.6 ${}_2S_4$ 0.02
			${}_3VP_4$ 0.1 ${}_3Noun_4$ 1.0

Quiz

S	→	NP VP	1.0	NP	→	her	0.2
VP	→	Verb NP	0.6	NP	→	I	0.2
VP	→	Verb S	0.3	Verb	→	saw	1.0
VP	→	duck	0.1	Noun	→	duck	1.0
NP	→	Det Noun	0.6	Det	→	her	1.0

0 I 1	saw 2	her 3	duck 4
${}_0NP_1$ 0.2		${}_0S_3$ 0.024	${}_0S_4$ 0.072
	${}_1Verb_2$ 1.0	${}_1VP_3$ 0.12	${}_1VP_4$ 0.36
		${}_2NP_3$ 0.2 ${}_2Det_3$ 1.0	${}_2NP_4$ 0.6 ${}_2S_4$ 0.02
			${}_3VP_4$ 0.1 ${}_3Noun_4$ 1.0

- Which items are combined into ${}_1VP_3$?
  1. ${}_1Verb_2$ and ${}_2NP_3$
  2. ${}_1Verb_2$ and ${}_2Det_3$
  3. ${}_1Verb_2$ and ${}_2NP_4$