

## Processament del Llenguatge Humà 8. Anàlisi sintàctica de constituents



UNIVERSITAT POLITÈCNICA DE CATALUNYA  
BARCELONATECH

Facultat d'Informàtica de Barcelona



# Índice

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

- 1 Análisis sintáctico
  - Objetivo y motivación
- 2 Estructuras sintácticas y gramáticas
- 3 Análisis de constituyentes
  - Previos
  - Métodos basados en *charts*
  - Algoritmo CKY probabilístico (PCKY)

# Índice

Análisis  
sintáctico

Objetivo y  
motivación

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

- 1 Análisis sintáctico
  - Objetivo y motivación
- 2 Estructuras sintácticas y gramáticas
- 3 Análisis de constituyentes
  - Previos
  - Métodos basados en *charts*
  - Algoritmo CKY probabilístico (PCKY)

# Objetivo y motivación

- **Objetivo:** encontrar la estructura sintáctica asociada a una oración.  
(Información sobre la combinación de las palabras en una oración)

Análisis  
sintáctico

Objetivo y  
motivación

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

# Objetivo y motivación

- **Objetivo:** encontrar la estructura sintáctica asociada a una oración.  
(Información sobre la combinación de las palabras en una oración)
- Ejemplos de aplicaciones:
  - Reconocimiento de autoría de documentos

# Objetivo y motivación

- **Objetivo:** encontrar la estructura sintáctica asociada a una oración.  
(Información sobre la combinación de las palabras en una oración)
  - Ejemplos de aplicaciones:
    - Reconocimiento de autoría de documentos
    - Detección automática de errores gramaticales
- P.e.: "*People walks*"  $\implies$  error

# Objetivo y motivación

- **Objetivo:** encontrar la estructura sintáctica asociada a una oración.

(Información sobre la combinación de las palabras en una oración)

- Ejemplos de aplicaciones:
  - Reconocimiento de autoría de documentos
  - Detección automática de errores gramaticales  
P.e.: "*People walks*"  $\implies$  error
  - Traducción automática  
P.e.: [es] NN+JJ  $\implies$  [en] JJ+NN

# Objetivo y motivación

- **Objetivo:** encontrar la estructura sintáctica asociada a una oración.  
(Información sobre la combinación de las palabras en una oración)
- Ejemplos de aplicaciones:
  - Reconocimiento de autoría de documentos
  - Detección automática de errores gramaticales  
P.e.: "*People walks*"  $\implies$  error
  - Traducción automática  
P.e.: [es] NN+JJ  $\implies$  [en] JJ+NN
  - Extracción de información  
P.e.:  $X - [subj] \rightarrow \text{visitó} \leftarrow [dobj] - Y \implies \text{visitar}(X,Y)$
  - ...



# Índice

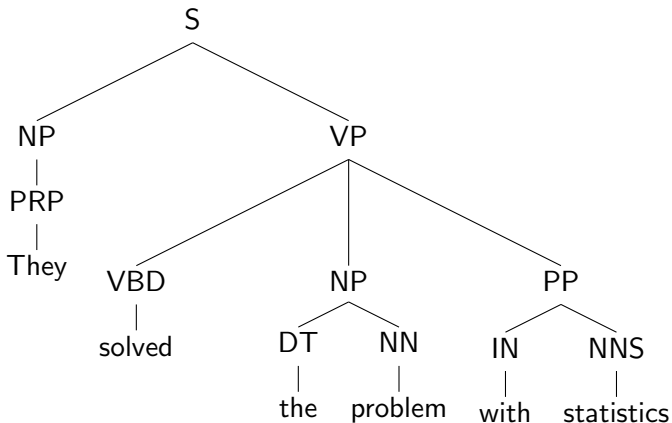
Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

- 1 Análisis sintáctico
  - Objetivo y motivación
- 2 Estructuras sintácticas y gramáticas
- 3 Análisis de constituyentes
  - Previos
  - Métodos basados en *charts*
  - Algoritmo CKY probabilístico (PCKY)

# Un árbol de constituyentes



Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

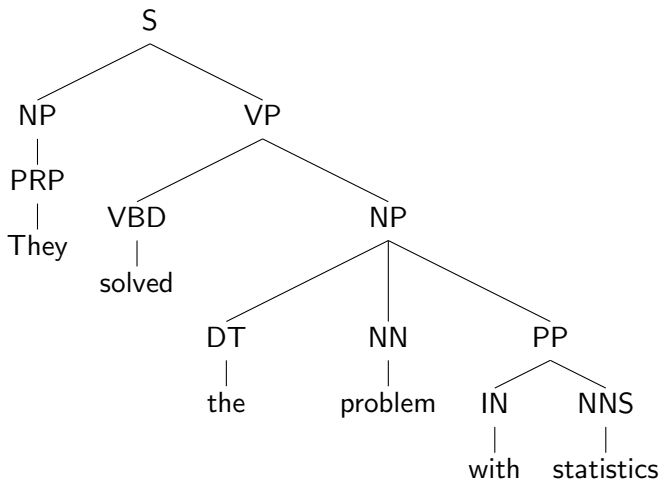
Análisis de  
constituyentes

# Otro

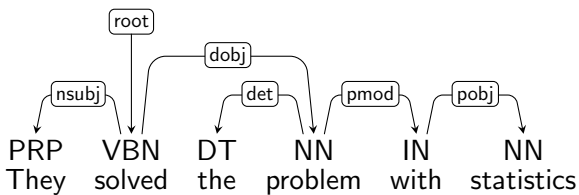
Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes



# Árbol de dependencias

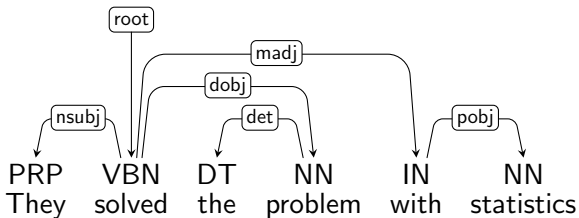
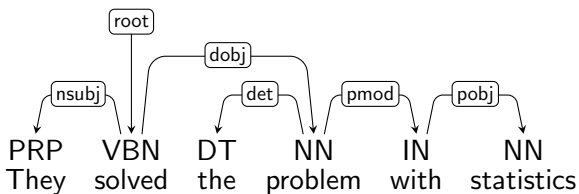


Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

# Árbol de dependencias



Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

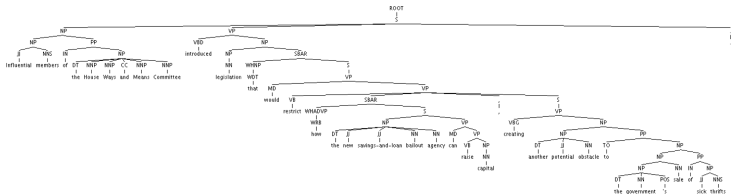
Análisis de  
constituyentes

# Caso "real" de oración en inglés

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

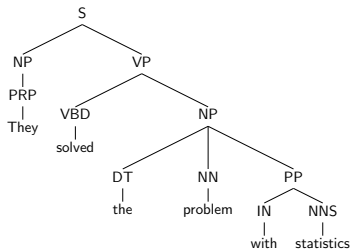
Análisis de  
constituyentes



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.

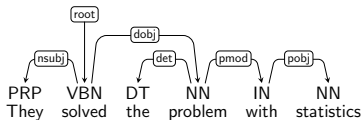
# Teorías de estructuras sintácticas

## Árbol de constituyentes



- Unidad: constituyente
- Constituyente = unidad lingüística abstracta
- Rigidez en la composición

## Árbol de dependencias



- Unidad: dependencia
- Dependencia = relación sintáctica entre dos palabras
- Flexibilidad en el orden

Análisis sintáctico

Estructuras sintácticas y gramáticas

Análisis de constituyentes

# Gramática libre de contexto (CFG)

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Una gramática libre de contexto es una tupla  $G = \langle N, \Sigma, R, S \rangle$  donde:

- $N$ : conjunto de símbolos no terminales
- $S \in N$ : símbolo de inicio
- $\Sigma$ : conjunto de símbolos terminales
- $R$ : conjunto de reglas de la forma  $X \rightarrow Y_1 Y_2 \dots Y_n$  donde  $n \geq 0$ ,  $X \in N$ ,  $Y_i \in N \cup \Sigma$



# Gramática libre de contexto, Ejemplo

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

$$S = \{S\}$$

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

$$R = \left\{ \begin{array}{ll} S \rightarrow NP VP & Vi \rightarrow \text{sleeps} \\ S \rightarrow NP Vi & Vt \rightarrow \text{saw} \\ NP \rightarrow DT NN & NN \rightarrow \text{man} \\ NP \rightarrow NP PP & NN \rightarrow \text{woman} \\ PP \rightarrow IN NP & NN \rightarrow \text{telescope} \\ VP \rightarrow Vt NP & DT \rightarrow \text{the} \\ VP \rightarrow VP PP & IN \rightarrow \text{with} \\ VP \rightarrow Vi PP & IN \rightarrow \text{in} \end{array} \right\}$$

<sup>1</sup>S=sentence, VP=verb phrase, NP=noun phrase, PP=prepositional phrase, DT=determiner, Vi=intransitive verb, Vt=transitive verb, NN=noun, IN=preposition

# Propiedades de una CFG

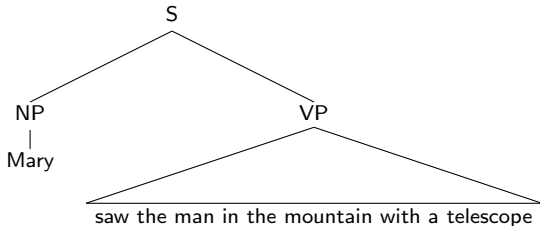
Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

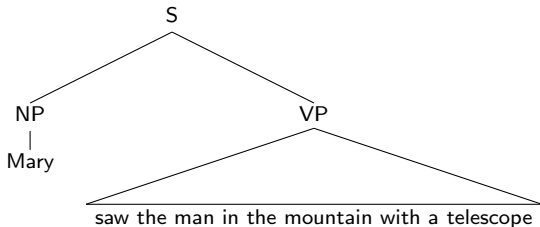
- Una CFG,  $G = \langle N, \Sigma, R, S \rangle$ , define un conjunto de posibles *derivaciones* (árboles  $t \in \mathcal{T}_G$ )
- Una secuencia de terminales,  $s \in \Sigma^*$ , pertenece a  $L_G$  si, al menos, existe una derivación para  $s$
- **Ambigüedad:** Algunas secuencias de  $L_G$  pueden tener más de una derivación.

# Ambigüedad



- *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*

# Ambigüedad



- *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*

Algunos árboles son más probables que otros...

¿Podemos modelar esto?

# Gramática libre de contexto (CFG)

Una gramática libre de contexto es una tupla  $G = \langle N, \Sigma, R, S \rangle$  donde:

- $N$ : conjunto de símbolos no terminales
- $S \in N$ : símbolo de inicio
- $\Sigma$ : conjunto de símbolos terminales
- $R$ : conjunto de reglas de la forma  $X \rightarrow Y_1 Y_2 \dots Y_n$  donde  $n \geq 0$ ,  $X \in N$ ,  $Y_i \in N \cup \Sigma$

# Gramática probabilística libre de contexto (PCFG)

Una gramática **probabilística** libre de contexto es una tupla  $G = \langle N, \Sigma, R, S, q \rangle$  donde:

- $N$ : conjunto de símbolos no terminales
- $S \in N$ : símbolo de inicio
- $\Sigma$ : conjunto de símbolos terminales
- $R$ : conjunto de reglas de la forma  $X \rightarrow Y_1 Y_2 \dots Y_n$  donde  $n \geq 0$ ,  $X \in N$ ,  $Y_i \in N \cup \Sigma$
- $q$ : conjunto de parámetros, uno por regla de  $R$  tal que,

$$\forall X \in N \quad \sum_{\{r \in R \mid \text{antecedente}(r) = X\}} q(r) = 1$$

# Gramática probabilística libre de contexto: Ejemplo

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

$$S = \{S\}$$

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

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

<sup>1</sup>S=sentence, VP=verb phrase, NP=noun phrase, PP=prepositional phrase, DT=determiner, Vi=intransitive verb, Vt=transitive verb, NN=noun, IN=preposition

# Propiedades de una PCFG

- Probabilidad de un árbol  $t \in \mathcal{T}_G$ :

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



# Propiedades de una PCFG

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

- Probabilidad de un árbol  $t \in \mathcal{T}_G$ :

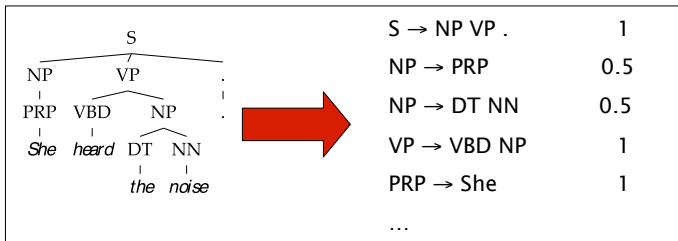
$$p(t) = \prod_{r \in t} q(r)$$

- Árbol más probable para una oración  $s$ ,  $\hat{t} \in \mathcal{T}_G(s)$ :

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

# Aprendizaje de PCFGs

- Inferir las reglas y sus probabilidades a partir de bancos de árboles



Via estimación por máxima verosimilitud:

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

- Se aplica *smoothing*
- La CFG inferida debe ser adecuada para obtener buenos resultados

# Índice

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

- 1 Análisis sintáctico
  - Objetivo y motivación
- 2 Estructuras sintácticas y gramáticas
- 3 Análisis de constituyentes
  - Previos
  - Métodos basados en *charts*
  - Algoritmo CKY probabilístico (PCKY)

# Índice

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Previos

- 1 Análisis sintáctico
  - Objetivo y motivación
- 2 Estructuras sintácticas y gramáticas
- 3 Análisis de constituyentes
  - Previos
  - Métodos basados en *charts*
  - Algoritmo CKY probabilístico (PCKY)

# Objetivos de un analizador sintáctico

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Previos

Dada una oración:

- Encontrar todos sus árboles posibles
- Obtener un ranking de probabilidades de dichos árboles
- Encontrar su árbol más probable

# Aspectos que influyen en el proceso de análisis

Relativos a los recursos:

- Expresividad de la gramática (combinación de símbolos)
- Cobertura (palabras)

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Previos

# Aspectos que influyen en el proceso de análisis

Relativos a los recursos:

- Expresividad de la gramática (combinación de símbolos)
- Cobertura (palabras)

Relativos al analizador:

- Estrategia de análisis (*bottom-up, top-down*)

# Aspectos que influyen en el proceso de análisis

Relativos a los recursos:

- Expresividad de la gramática (combinación de símbolos)
- Cobertura (palabras)

Relativos al analizador:

- Estrategia de análisis (*bottom-up*, *top-down*)
- Orden de aplicación de reglas (la más larga, la más probable)



# Aspectos que influyen en el proceso de análisis

Relativos a los recursos:

- Expresividad de la gramática (combinación de símbolos)
- Cobertura (palabras)

Relativos al analizador:

- Estrategia de análisis (*bottom-up*, *top-down*)
- Orden de aplicación de reglas (la más larga, la más probable)
- Gestión de la ambigüedad (guardar todos, seleccionar uno - probabilidades, semántica, pragmática)
- ...

# Problema del uso de *Backtracking*

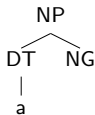
Ambas estrategias (*top-down*, *bottom-up*) pueden causar repetición de derivaciones.

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

NG  $\rightarrow$  NN

NG  $\rightarrow$  NG PP

...



# Problema del uso de *Backtracking*

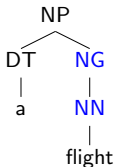
Ambas estrategias (*top-down*, *bottom-up*) pueden causar repetición de derivaciones.

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

▶ NG → NN

NG → NG PP

...



# Problema del uso de *Backtracking*

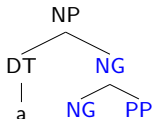
Ambas estrategias (*top-down*, *bottom-up*) pueden repetir derivaciones

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

NG  $\rightarrow$  NN

► NG  $\rightarrow$  NG PP

...



# Problema del uso de *Backtracking*

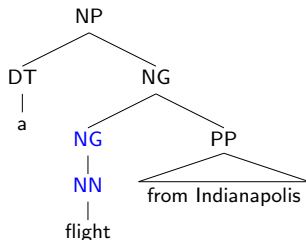
Ambas estrategias (*top-down*, *bottom-up*) pueden repetir derivaciones

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

► **NG → NN**

NG → NG PP

...



# Problema del uso de *Backtracking*

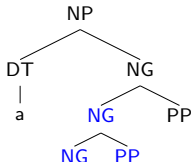
Ambas estrategias (*top-down*, *bottom-up*) pueden repetir derivaciones

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

NG  $\rightarrow$  NN

► NG  $\rightarrow$  NG PP

...



## Problema del uso de *Backtracking*

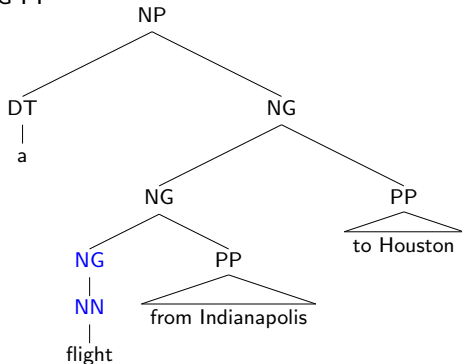
Ambas estrategias (*top-down*, *bottom-up*) pueden repetir derivaciones

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

► NG → NN

NG → NG PP

...



# Índice

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Métodos basados en  
*charts*

- 1 Análisis sintáctico
  - Objetivo y motivación
- 2 Estructuras sintácticas y gramáticas
- 3 Análisis de constituyentes
  - Previos
  - Métodos basados en *charts*
  - Algoritmo CKY probabilístico (PCKY)



# Propiedades

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

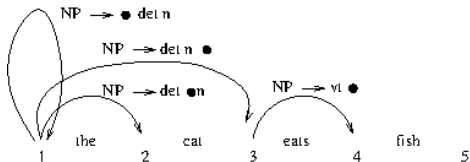
Métodos basados en  
*charts*

- Evitan la repetición de derivaciones mediante programación dinámica
- *Chart*: grafo dirigido que representa las derivaciones
- Representación del *chart*: tabla de programación dinámica.

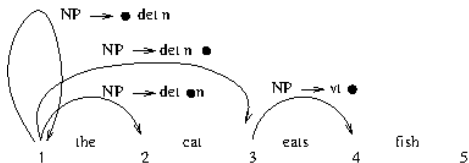
# Chart

- **Nodos:** posiciones entre pares de palabras de la oración a analizar
- **Arcos:** ejecución parcial o total de una regla de la gramática sobre una secuencia de palabras de la oración
- **Etiquetas de arcos:** **reglas punteadas** (estado de ejecución de las reglas)
  - **Pasivas:**  $A \rightarrow B_1 \dots B_k \bullet$
  - **Activas:**  $A \rightarrow B_1 \dots B_i \bullet B_{i+1} \dots B_k$

P.e.:



# Chart como tabla de programación dinámica



					[1,5]
				[1,4]	[2,5]
		[1,3]		[2,4]	[3,5]
	[1,2]	NP → det • n	[2,3]	[3,4]	[4,5]
[1,1]	NP → • det n	[2,2]	[3,3]	[4,4]	[5,5]
	1	2	3	4	5
	the	cat	eats	fish	

Análisis sintáctico

Estructuras sintácticas y gramáticas

Análisis de constituyentes

Métodos basados en charts

# Algoritmos basados en *charts*

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Métodos basados en  
*charts*

- CKY
  - introdujo la programación dinámica
  - limitado a CFGs en Forma Normal de Chomsky (FNC)
  - estrategia *bottom-up* (solo reglas punteadas pasivas)
  - **versión probabilística directa**

# Algoritmos basados en *charts*

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Métodos basados en  
*charts*

- CKY
  - introdujo la programación dinámica
  - limitado a CFGs en Forma Normal de Chomsky (FNC)
  - estrategia *bottom-up* (solo reglas punteadas pasivas)
  - **versión probabilística directa**
- Earley
  - cualquier CFG
  - estrategia *top-down* (reglas punteadas activas/pasivas)
  - versión probabilística no directa

# Algoritmos basados en *charts*

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Métodos basados en  
*charts*

- CKY
  - introdujo la programación dinámica
  - limitado a CFGs en Forma Normal de Chomsky (FNC)
  - estrategia *bottom-up* (solo reglas punteadas pasivas)
  - **versión probabilística directa**
- Earley
  - cualquier CFG
  - estrategia *top-down* (reglas punteadas activas/pasivas)
  - versión probabilística no directa
- Analizador generalizado de charts

# Índice

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

- 1 Análisis sintáctico
  - Objetivo y motivación
- 2 Estructuras sintácticas y gramáticas
- 3 Análisis de constituyentes
  - Previos
  - Métodos basados en *charts*
  - Algoritmo CKY probabilístico (PCKY)

# Propiedades

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

- Bottom-up, reglas punteadas pasivas
- Requiere una PCFG en FNC
- Proporciona el árbol sintáctico más probable para cada oración según la PCFG usada
- Análogo a Viterbi para HMMs.



## Previo: Forma Normal de Chomsky (FNC)

Una PCFG  $G = (N, \Sigma, R, S)$  está expresada en FNC si:

- $N$ : conjunto de símbolos no terminales
- $\Sigma$ : conjunto de símbolos terminales
- $S \in N$ : símbolo de inicio
- $R$ : conjunto de reglas que satisfacen una de las formas:
  - $X \rightarrow Y_1 Y_2$  con  $X, Y_1, Y_2 \in N$
  - $X \rightarrow \alpha$  con  $X \in N$  y  $\alpha \in \Sigma$
- $q$ : conjunto de parámetros, uno por regla de  $R$  tal que,

$$\forall X \in N \quad \sum_{\{r \in R \mid \text{antecedente}(r) = X\}} q(r) = 1$$

## Previo: Forma Normal de Chomsky (FNC)

Una PCFG  $G = (N, \Sigma, R, S)$  está expresada en FNC si:

- $N$ : conjunto de símbolos no terminales
- $\Sigma$ : conjunto de símbolos terminales
- $S \in N$ : símbolo de inicio
- $R$ : conjunto de reglas que satisfacen una de las formas:
  - $X \rightarrow Y_1 Y_2$  con  $X, Y_1, Y_2 \in N$
  - $X \rightarrow \alpha$  con  $X \in N$  y  $\alpha \in \Sigma$
- $q$ : conjunto de parámetros, uno por regla de  $R$  tal que,

$$\forall X \in N \quad \sum_{\{r \in R \mid \text{antecedente}(r) = X\}} q(r) = 1$$

Toda CFG que no genere la secuencia vacía puede convertirse a FNC

# Conversión a FNC

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

- 1 Reglas híbridas: sustituir terminales por nuevos no terminales

P.e.:  $\text{INF\_VP} \rightarrow \text{to VP} (p_1) \implies \begin{array}{l} \text{INF\_VP} \rightarrow \text{TO VP} (p_1) \\ \text{TO} \rightarrow \text{to} (1.0) \end{array}$

- 2 Reglas unitarias:

P.e.:  $\begin{array}{l} \text{NP} \rightarrow \text{N} (p_1) \\ \text{N} \rightarrow \text{dog} (p_2) \end{array} \implies \text{NP} \rightarrow \text{dog} (p_1 * p_2)$

- 3 Reglas no binarias:

P.e.:  $\text{S} \rightarrow \text{VP NP PP} (p_1) \implies \begin{array}{l} \text{S} \rightarrow \text{VP X} (p_1) \\ \text{X} \rightarrow \text{NP PP} (1.0) \end{array}$

# Ejercicio

Transforma la siguiente CFG a FNC

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

# Ejercicio

Transforma la siguiente CFG a FNC

1  $S \rightarrow NP VP$  (1.0)

2  $NP \rightarrow det n$  (0.6)

3  $NP \rightarrow n$  (0.4)

4  $VP \rightarrow vt NP PP$  (0.7)

5  $VP \rightarrow vi$  (0.3)

6  $PP \rightarrow with NP$  (1.0)

7  $det \rightarrow the|a$  (0.6|0.4)

8  $n \rightarrow$   
 $cat|fish|knife$  (0.3|0.5|0.2)

9  $vt \rightarrow eats$  (1.0)

10  $vi \rightarrow eats$  (1.0)

(1) Reglas híbridas

# Ejercicio

Transforma la siguiente CFG a FNC

1  $S \rightarrow NP VP$  (1.0)

2  $NP \rightarrow det n$  (0.6)

3  $NP \rightarrow n$  (0.4)

4  $VP \rightarrow vt NP PP$  (0.7)

5  $VP \rightarrow vi$  (0.3)

6  $PP \rightarrow with NP$  (1.0)

7  $det \rightarrow the|a$  (0.6|0.4)

8  $n \rightarrow$   
 $cat|fish|knife$  (0.3|0.5|0.2)

9  $vt \rightarrow eats$  (1.0)

10  $vi \rightarrow eats$  (1.0)

(1) Reglas híbridas

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

# Ejercicio

Transforma la siguiente CFG a FNC

$$1 \quad S \rightarrow NP VP \quad (1.0)$$

$$2 \quad NP \rightarrow det \ n \quad (0.6)$$

$$3 \quad NP \rightarrow n \quad (0.4)$$

$$4 \quad VP \rightarrow vt \ NP \ PP \quad (0.7)$$

$$5 \quad VP \rightarrow vi \quad (0.3)$$

$$6 \quad PP \rightarrow with \ NP \quad (1.0)$$

$$7 \quad det \rightarrow the|a \quad (0.6|0.4)$$

$$8 \quad n \rightarrow \\ cat|fish|knife \quad (0.3|0.5|0.2)$$

$$9 \quad vt \rightarrow eats \quad (1.0)$$

$$10 \quad vi \rightarrow eats \quad (1.0)$$

(1) Reglas híbridas

$$PP \rightarrow A \ NP \quad (1.0)$$

$$A \rightarrow with \quad (1.0)$$

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

# Ejercicio

Transforma la siguiente CFG a FNC

1  $S \rightarrow NP VP$  (1.0)

2  $NP \rightarrow det n$  (0.6)

3  $NP \rightarrow n$  (0.4)

4  $VP \rightarrow vt NP PP$  (0.7)

5  $VP \rightarrow vi$  (0.3)

6  $PP \rightarrow with NP$  (1.0)

7  $det \rightarrow the|a$  (0.6|0.4)

8  $n \rightarrow$   
 $cat|fish|knife$  (0.3|0.5|0.2)

9  $vt \rightarrow eats$  (1.0)

10  $vi \rightarrow eats$  (1.0)

(1) Reglas híbridas

$PP \rightarrow A NP$  (1.0)

$A \rightarrow with$  (1.0)

(2) Reglas unitarias

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# Ejercicio

Transforma la siguiente CFG a FNC

1  $S \rightarrow NP VP$  (1.0)

2  $NP \rightarrow det n$  (0.6)

3  $NP \rightarrow n$  (0.4)

4  $VP \rightarrow vt NP PP$  (0.7)

5  $VP \rightarrow vi$  (0.3)

6  $PP \rightarrow with NP$  (1.0)

7  $det \rightarrow the|a$  (0.6|0.4)

8  $n \rightarrow$   
 $cat|fish|knife$  (0.3|0.5|0.2)

9  $vt \rightarrow eats$  (1.0)

10  $vi \rightarrow eats$  (1.0)

(1) Reglas híbridas

$PP \rightarrow A NP$  (1.0)

$A \rightarrow with$  (1.0)

(2) Reglas unitarias

# Ejercicio

Transforma la siguiente CFG a FNC

$$1 \quad S \rightarrow NP VP \quad (1.0)$$

$$2 \quad NP \rightarrow det \ n \quad (0.6)$$

$$3 \quad NP \rightarrow n \quad (0.4)$$

$$4 \quad VP \rightarrow vt \ NP \ PP \quad (0.7)$$

$$5 \quad VP \rightarrow vi \quad (0.3)$$

$$6 \quad PP \rightarrow with \ NP \quad (1.0)$$

$$7 \quad det \rightarrow the|a \quad (0.6|0.4)$$

$$8 \quad n \rightarrow \\ cat|fish|knife \quad (0.3|0.5|0.2)$$

$$9 \quad vt \rightarrow eats \quad (1.0)$$

$$10 \quad vi \rightarrow eats \quad (1.0)$$

(1) Reglas híbridas

$$PP \rightarrow A \ NP \quad (1.0)$$

$$A \rightarrow with \quad (1.0)$$

(2) Reglas unitarias

$$NP \rightarrow cat|fish|knife \\ (0.12|0.2|0.08)$$

$$VP \rightarrow eats \quad (0.3)$$

# Ejercicio

Transforma la siguiente CFG a FNC

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

(1) Reglas híbridas

$PP \rightarrow A NP$  (1.0)

$A \rightarrow with$  (1.0)

(2) Reglas unitarias

$NP \rightarrow cat|fish|knife$   
(0.12|0.2|0.08)

$VP \rightarrow eats$  (0.3)

(3) Reglas no binarias

# Ejercicio

Transforma la siguiente CFG a FNC

1  $S \rightarrow NP VP$  (1.0)

2  $NP \rightarrow det n$  (0.6)

3  $NP \rightarrow n$  (0.4)

4  $VP \rightarrow vt NP PP$  (0.7)

5  $VP \rightarrow vi$  (0.3)

6  $PP \rightarrow with NP$  (1.0)

7  $det \rightarrow the|a$  (0.6|0.4)

8  $n \rightarrow$   
 $cat|fish|knife$  (0.3|0.5|0.2)

9  $vt \rightarrow eats$  (1.0)

10  $vi \rightarrow eats$  (1.0)

(1) Reglas híbridas

$PP \rightarrow A NP$  (1.0)

$A \rightarrow with$  (1.0)

(2) Reglas unitarias

$NP \rightarrow cat|fish|knife$   
(0.12|0.2|0.08)

$VP \rightarrow eats$  (0.3)

(3) Reglas no binarias

# Ejercicio

Transforma la siguiente CFG a FNC

$$1 \quad S \rightarrow NP VP \quad (1.0)$$

$$2 \quad NP \rightarrow det \ n \quad (0.6)$$

$$3 \quad NP \rightarrow n \quad (0.4)$$

$$4 \quad VP \rightarrow vt \ NP \ PP \quad (0.7)$$

$$5 \quad VP \rightarrow vi \quad (0.3)$$

$$6 \quad PP \rightarrow with \ NP \quad (1.0)$$

$$7 \quad det \rightarrow the|a \quad (0.6|0.4)$$

$$8 \quad n \rightarrow \\ cat|fish|knife \quad (0.3|0.5|0.2)$$

$$9 \quad vt \rightarrow eats \quad (1.0)$$

$$10 \quad vi \rightarrow eats \quad (1.0)$$

(1) Reglas híbridas

$$PP \rightarrow A \ NP \quad (1.0)$$

$$A \rightarrow with \quad (1.0)$$

(2) Reglas unitarias

$$NP \rightarrow cat|fish|knife \\ (0.12|0.2|0.08)$$

$$VP \rightarrow eats \quad (0.3)$$

(3) Reglas no binarias

$$VP \rightarrow vtB \quad (0.7)$$

$$B \rightarrow NP \ PP \quad (1.0)$$

# Ejercicio

Transforma la siguiente CFG a FNC

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

- 1  $S \rightarrow NP VP$  (1.0)
- 2  $NP \rightarrow det n$  (0.6)
- 3  $NP \rightarrow$   
 $cat|fish|knife$  (0.12|0.2|0.08)
- 4  $VP \rightarrow vt B$  (0.7)
- 5  $B \rightarrow NP PP$  (1.0)
- 6  $VP \rightarrow eats$  (0.3)
- 7  $PP \rightarrow A NP$  (1.0)
- 8  $A \rightarrow with$  (1.0)
- 9  $det \rightarrow the|a$  (0.6|0.4)
- 10  $n \rightarrow$   
 $cat|fish|knife$  (0.3|0.5|0.2)
- 11  $vt \rightarrow eats$  (1.0)

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

# Propiedades

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

- Bottom-up, reglas punteadas pasivas
- Requiere una PCFG en FNC
- Proporciona el árbol sintáctico más probable para cada oración según la PCFG usada
- Análogo a Viterbi para HMMs.

# Algoritmo CKY probabilístico (PCKY)

Aplicar programación dinámica.

Tabla dinámica: (p.e. oración de 4 palabras)

			14
		13	24
	12	23	34
11	22	33	44
1	2	3	4

Cada componente  $(i, j)$  guarda información sobre el subárbol mas probable  $t \in \mathcal{T}_X(s_{ij})$  para cada raíz de árbol  $X$  que subsume la secuencia de palabras  $s_{ij}$

$$\pi(i, j, X) = \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

$$\psi(i, j, X) = \arg \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

Análisis sintáctico

Estructuras sintácticas y gramáticas

Análisis de constituyentes

Algoritmo CKY probabilístico (PCKY)



# Algoritmo CKY probabilístico (PCKY)

Aplicar programación dinámica.

Tabla dinámica: (p.e. oración de 4 palabras)

			14
		13	24
	12	23	34
11	22	33	44
1	2	3	4

Ejemplo para (1,3)

$$\pi(i, j, X) = \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

$$\psi(i, j, X) = \arg \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

Análisis sintáctico

Estructuras sintácticas y gramáticas

Análisis de constituyentes

Algoritmo CKY probabilístico (PCKY)

# Algoritmo CKY probabilístico (PCKY)

Aplicar programación dinámica.

Tabla dinámica: (p.e. oración de 4 palabras)

				14
			13	24
	12	23	34	
11	22	33	44	
	1	2	3	4

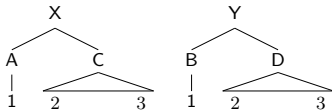
REGLAS APLICADAS EN (1,3)

$X \rightarrow A, C$   
 $Y \rightarrow B, D$

$X \rightarrow A_{11}, C_{23}$   
 $Y \rightarrow B_{11}, D_{23}$   
 $C \rightarrow \dots$   
 $D \rightarrow \dots$

$A \rightarrow w_1$   
 $B \rightarrow w_1$

Ejemplo para (1,3)



$$\pi(i, j, X) = \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

$$\psi(i, j, X) = \arg \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

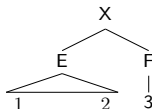
# Algoritmo CKY probabilístico (PCKY)

Aplicar programación dinámica.

Tabla dinámica: (p.e. oración de 4 palabras)

			14
	REGLAS APLICADAS EN (1,3) $X \rightarrow E, F$	13 $X \rightarrow E_{12}, F_{33}$ $Y \rightarrow B_{11}, D_{23}$	24
	12 $E \rightarrow \dots$	23	34
11	22	33 $F \rightarrow w_3$	44
1	2	3	4

Ejemplo para (1,3)



Con mayor probabilidad para la raíz X

$$\pi(i, j, X) = \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

$$\psi(i, j, X) = \arg \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

# Algoritmo CKY probabilístico (PCKY)

Aplicar programación dinámica.

Tabla dinámica: (p.e. oración de 4 palabras)

			14 S → ...
		13	24
	12	23	34
11	22	33	44
1	2	3	4

Objetivo:

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

$$\pi(i, j, X) = \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

$$\psi(i, j, X) = \arg \max_{t \in \mathcal{T}_X(s_{ij})} \prod_{r \in t} q(r)$$

Análisis sintáctico

Estructuras sintácticas y gramáticas

Análisis de constituyentes

Algoritmo CKY probabilístico (PCKY)

# Algoritmo PCKY

Caso base: Nodos terminales

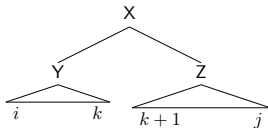
- $\forall i = 1 \dots n, \forall X \rightarrow w_i \in R, \pi(i, i, X) = q(X \rightarrow w_i)$

Caso recursivo: Nodos no terminales

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

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

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



Output:

- Devuelve  $\pi(1, n, S)$  y recupera el árbol via  $\psi(1, n, S)$

# Algoritmo PCKY - Ejemplo -

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

$$S = \{S\}$$

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

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

<sup>1</sup>S=sentence, VP=verb phrase, NP=noun phrase, PP=prepositional phrase, DT=determiner, Vi=intransitive verb, Vt=transitive verb, NN=noun, IN=preposition

# CKY Algorithm - Ejemplo -

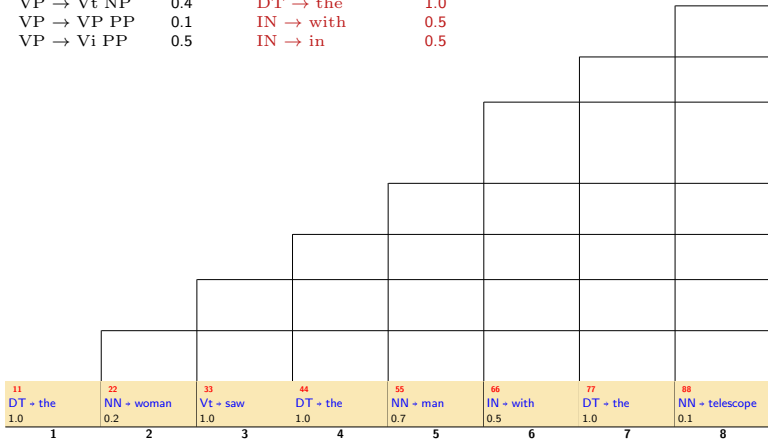
$S \rightarrow NP VP$	0.5	$V_i \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP V_i$	0.5	$V_t \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow V_t NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow V_i PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

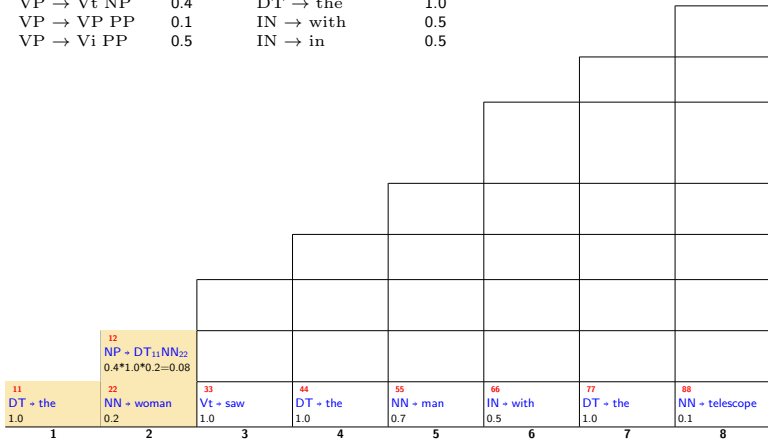
$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)





# CKY Algorithm - Ejemplo -

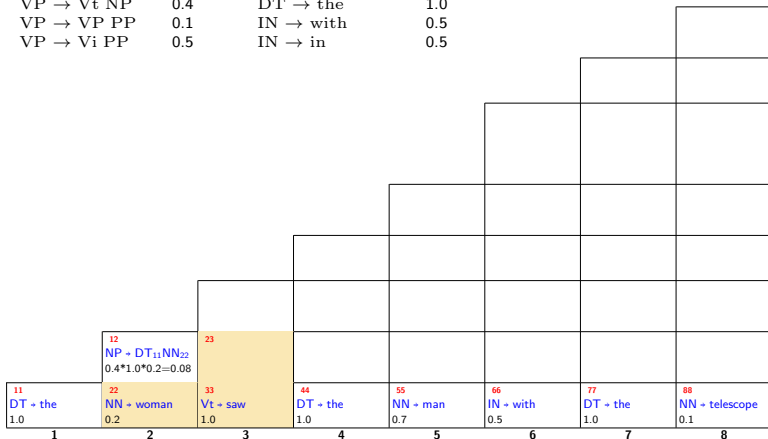
$S \rightarrow NP VP$	0.5	$V_i \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP V_i$	0.5	$V_t \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow V_t NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow V_i PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

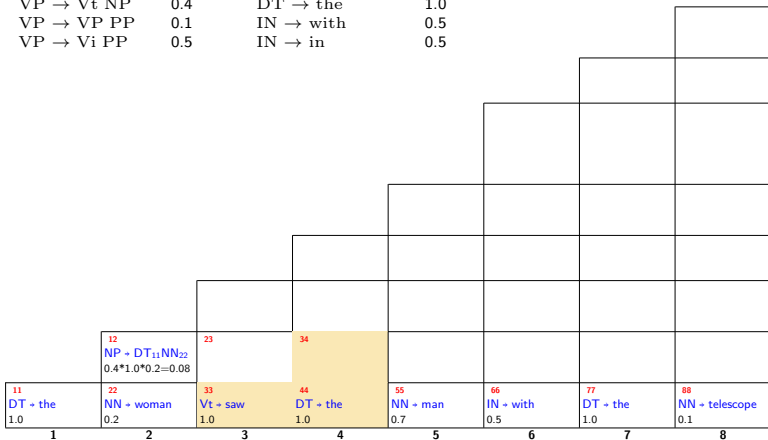
$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)





# CKY Algorithm - Ejemplo -

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

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

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

# CKY Algorithm - Ejemplo -

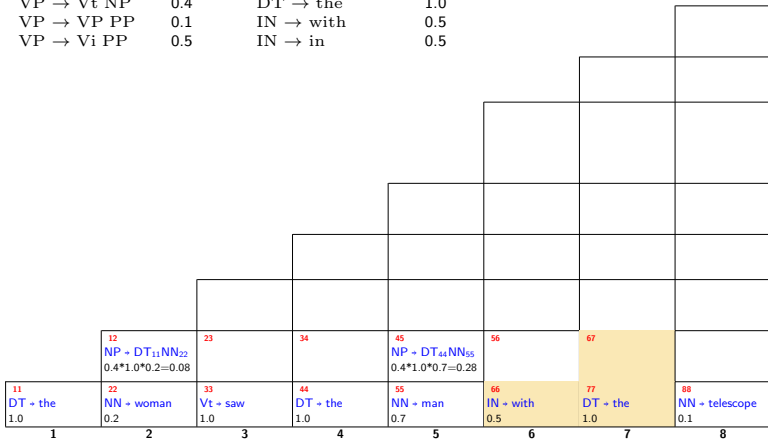
$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis sintáctico

Estructuras sintácticas y gramáticas

Análisis de constituyentes

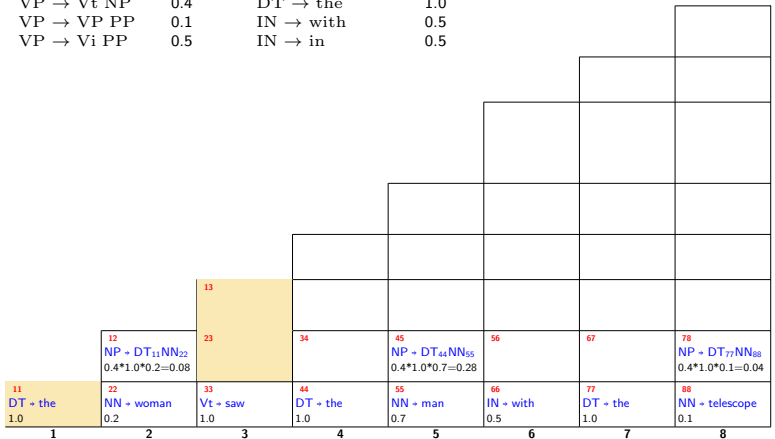
Algoritmo CKY probabilístico (PCKY)





# CKY Algorithm - Ejemplo -

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



Análisis sintáctico

Estructuras sintácticas y gramáticas

Análisis de constituyentes

Algoritmo CKY probabilístico (PCKY)

# CKY Algorithm - Ejemplo -

$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

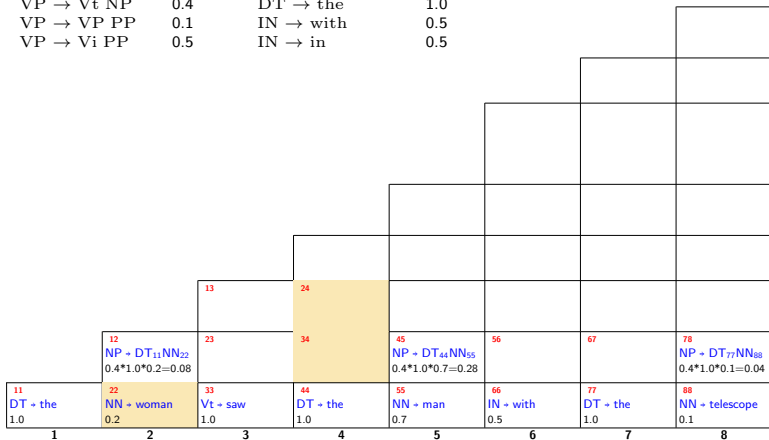
Algoritmo CKY  
probabilístico  
(PCKY)

		<b>12</b>	<b>13</b>					
		NP → DT <sub>11</sub> NN <sub>22</sub> 0.4*1.0*0.2=0.08	23	34	45 NP → DT <sub>44</sub> NN <sub>55</sub> 0.4*1.0*0.7=0.28	56	67	78 NP → DT <sub>77</sub> NN <sub>88</sub> 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 - Ejemplo -

$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5



Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

# CKY Algorithm - Ejemplo -

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

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

		12 NP → DT <sub>11</sub> NN <sub>22</sub> 0.4*1.0*0.2=0.08	13	24						
		23	34	45 NP → DT <sub>44</sub> NN <sub>55</sub> 0.4*1.0*0.7=0.28	56		67	78 NP → DT <sub>77</sub> NN <sub>88</sub> 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 - Ejemplo -

$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)

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

# CKY Algorithm - Ejemplo -

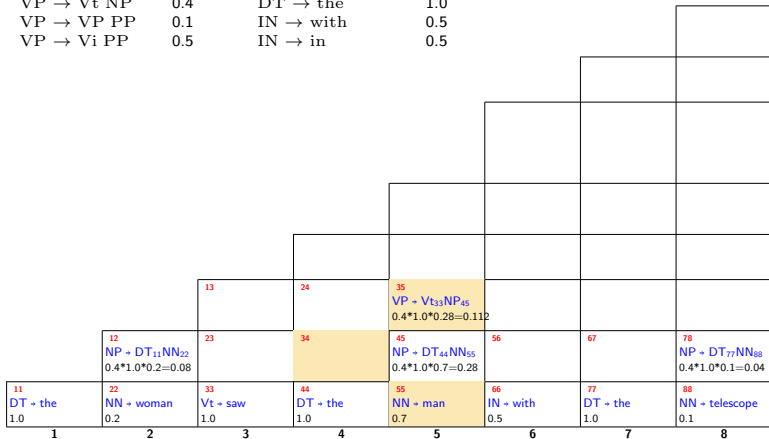
$S \rightarrow NP VP$	0.5	$Vi \rightarrow sleeps$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow saw$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow man$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow woman$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow telescope$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow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

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

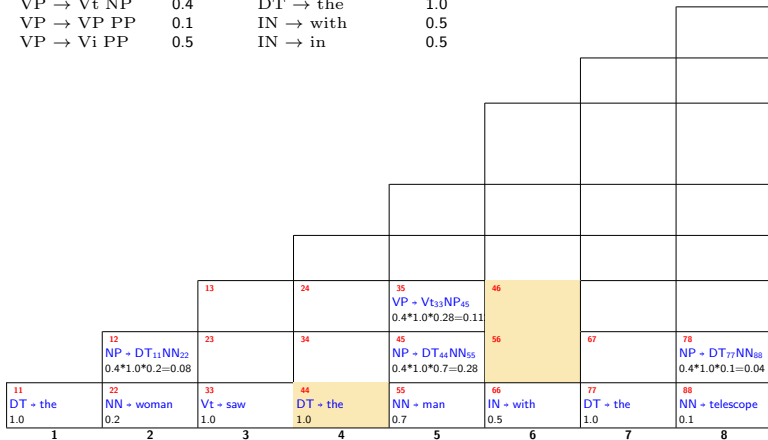
$S \rightarrow NP VP$	0.5	$V_i \rightarrow sleeps$	1.0
$S \rightarrow NP V_i$	0.5	$V_t \rightarrow saw$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow man$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow woman$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow telescope$	0.1
$VP \rightarrow V_t NP$	0.4	$DT \rightarrow the$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow with$	0.5
$VP \rightarrow V_i PP$	0.5	$IN \rightarrow in$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

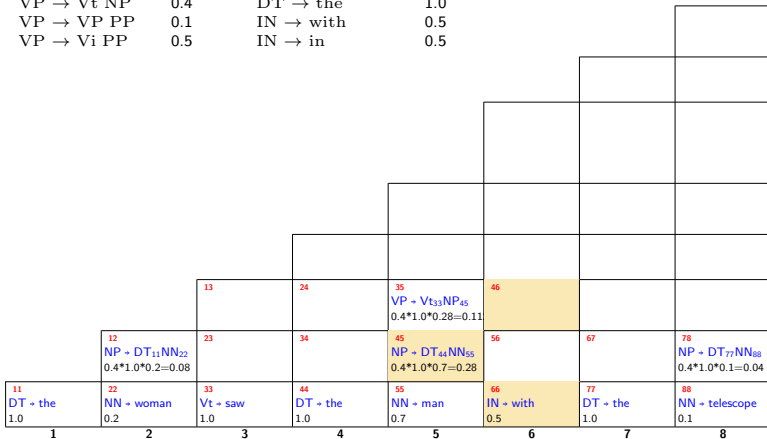
$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

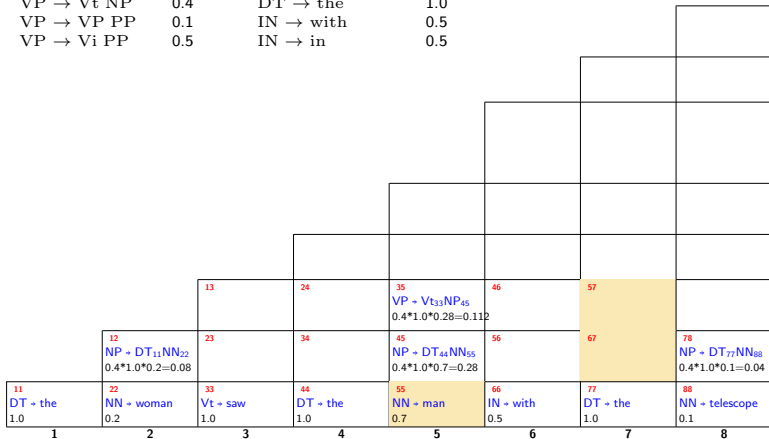
$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

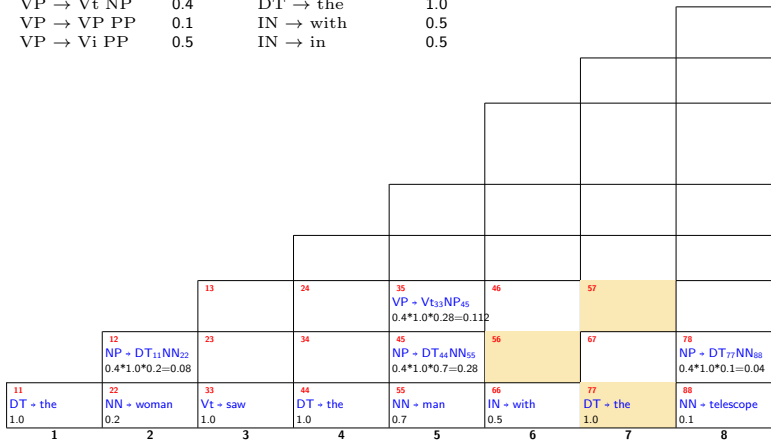
$S \rightarrow NP VP$	0.5	$Vi \rightarrow sleeps$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow saw$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow man$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow woman$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow telescope$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow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

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)





# CKY Algorithm - Ejemplo -

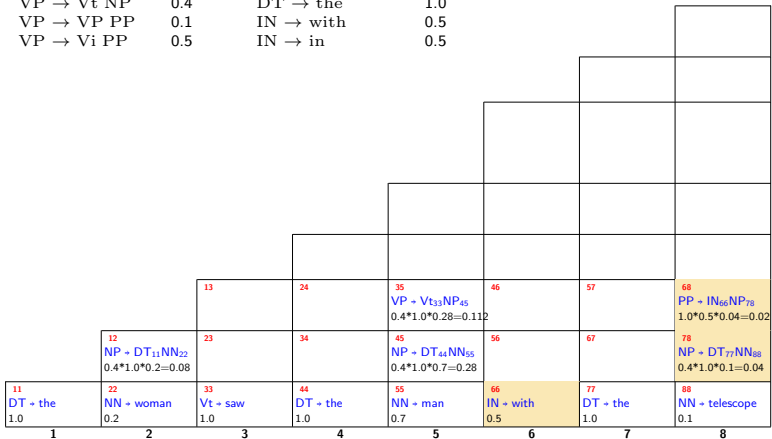
$S \rightarrow NP VP$	0.5	$Vi \rightarrow sleeps$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow saw$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow man$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow woman$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow telescope$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow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

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

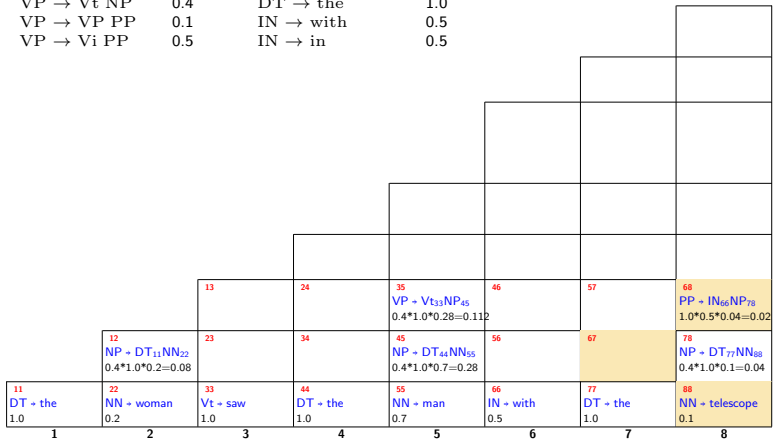
$S \rightarrow NP VP$	0.5	$V_i \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP V_i$	0.5	$V_t \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow V_t NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow V_i PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)





# CKY Algorithm - Ejemplo -

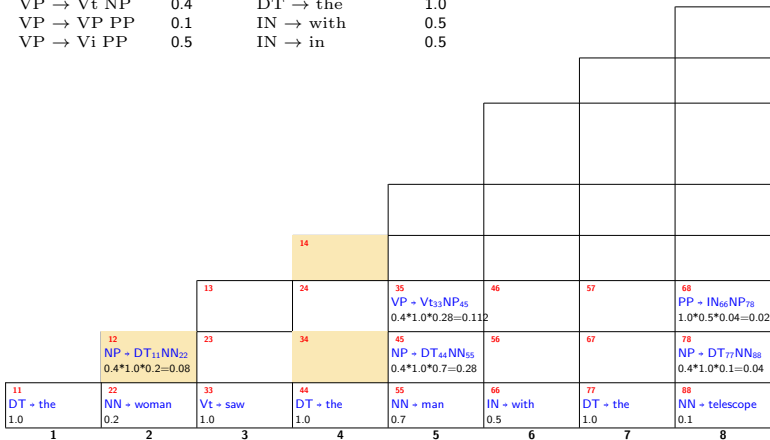
S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)





# CKY Algorithm - Ejemplo -

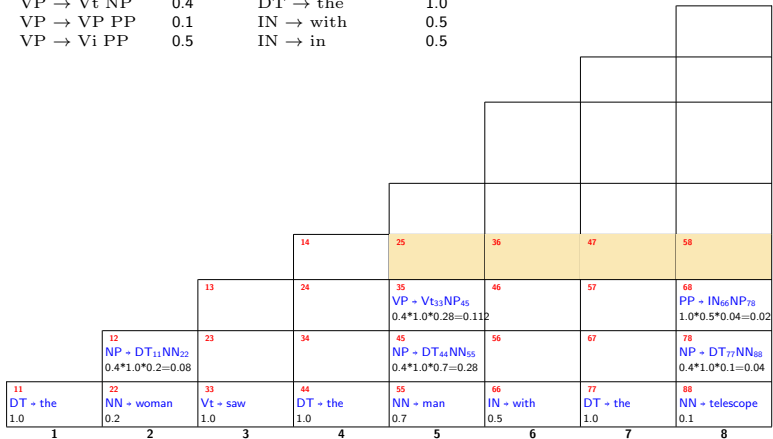
$S \rightarrow NP VP$	0.5	$Vi \rightarrow sleeps$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow saw$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow man$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow woman$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow telescope$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow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

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

$S \rightarrow NP VP$	0.5
$S \rightarrow NP Vi$	0.5
$NP \rightarrow DT NN$	0.4
$NP \rightarrow NP PP$	0.6
$PP \rightarrow IN NP$	1.0
$VP \rightarrow Vt NP$	0.4
$VP \rightarrow VP PP$	0.1
$VP \rightarrow Vi PP$	0.5

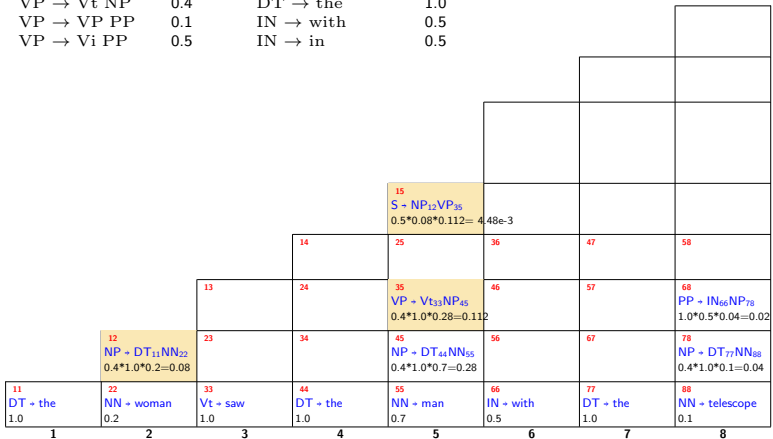
$Vi \rightarrow sleeps$	1.0
$Vt \rightarrow saw$	1.0
$NN \rightarrow man$	0.7
$NN \rightarrow woman$	0.2
$NN \rightarrow telescope$	0.1
$DT \rightarrow the$	1.0
$IN \rightarrow with$	0.5
$IN \rightarrow in$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

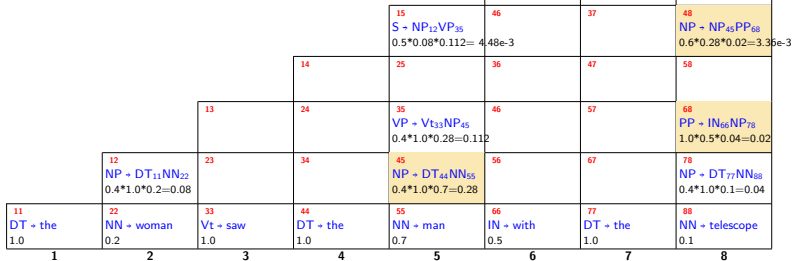
S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)







# CKY Algorithm - Ejemplo -

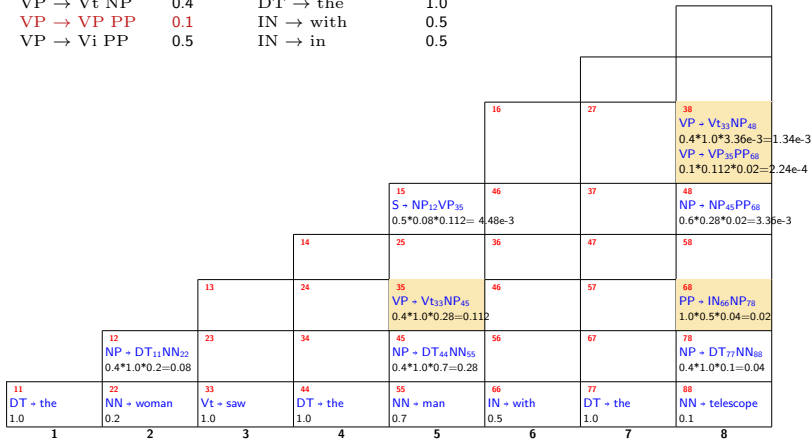
$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

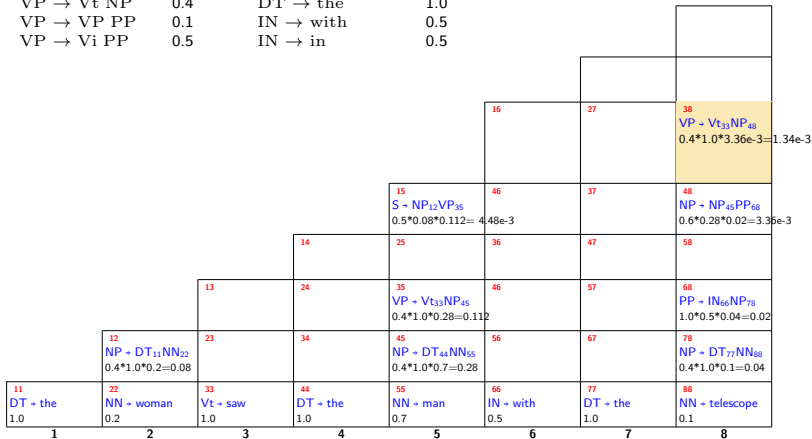
S → NP VP	0.5	Vi → sleeps	1.0
S → NP Vi	0.5	Vt → saw	1.0
NP → DT NN	0.4	NN → man	0.7
NP → NP PP	0.6	NN → woman	0.2
PP → IN NP	1.0	NN → telescope	0.1
VP → Vt NP	0.4	DT → the	1.0
VP → VP PP	0.1	IN → with	0.5
VP → Vi PP	0.5	IN → in	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

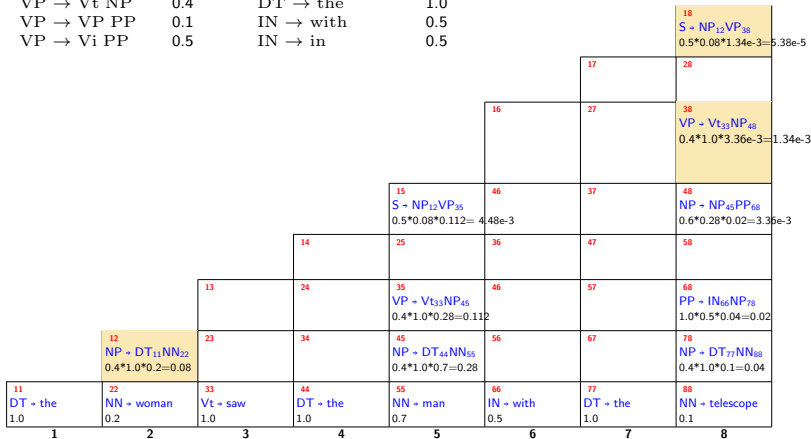
$S \rightarrow NP VP$	0.5	$Vi \rightarrow \text{sleeps}$	1.0
$S \rightarrow NP Vi$	0.5	$Vt \rightarrow \text{saw}$	1.0
$NP \rightarrow DT NN$	0.4	$NN \rightarrow \text{man}$	0.7
$NP \rightarrow NP PP$	0.6	$NN \rightarrow \text{woman}$	0.2
$PP \rightarrow IN NP$	1.0	$NN \rightarrow \text{telescope}$	0.1
$VP \rightarrow Vt NP$	0.4	$DT \rightarrow \text{the}$	1.0
$VP \rightarrow VP PP$	0.1	$IN \rightarrow \text{with}$	0.5
$VP \rightarrow Vi PP$	0.5	$IN \rightarrow \text{in}$	0.5

Análisis  
sintáctico

Estructuras  
sintácticas y  
gramáticas

Análisis de  
constituyentes

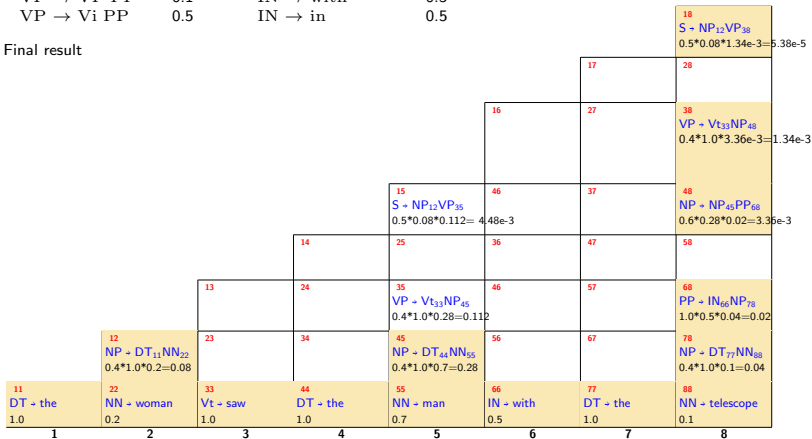
Algoritmo CKY  
probabilístico  
(PCKY)



# CKY Algorithm - Ejemplo -

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



Análisis sintáctico

Estructuras sintácticas y gramáticas

Análisis de constituyentes

Algoritmo CKY probabilístico (PCKY)