Cocke-Younger-Kasami (CYK) Parser Algorithm

Theory of Computation, L.EIC, 2nd Year

João M. P. Cardoso
Email: jmpc@acm.org

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

DEI DEPARTAMENTO DE ENGENHARIA INFORMÁTICA

Test if String belongs to a CFL

- Cocke-Younger-Kasami (CYK) Algorithm : CFG in CNF
  - X_{ij} – represents the set of variables that produce string i-j
  - $O(n^3)$, using dynamic programming, fill of a table

Test if String belongs to a CFL

- Cocke-Younger-Kasami (CYK) Algorithm
  - X_{ij} – represents the set of variables that produce string i-j
  - $O(n^3)$, using dynamic programming, fill of a table

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

Input string: baaba

{S,A,C}				
	{S,A,C}			
	{B}	{B}		
{S,A}	{B}	{S,C}	{S,A}	
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

X_{12}: $X_{11}X_{22}$; X_{24}: $X_{22}X_{34} \cup X_{23}X_{44}$

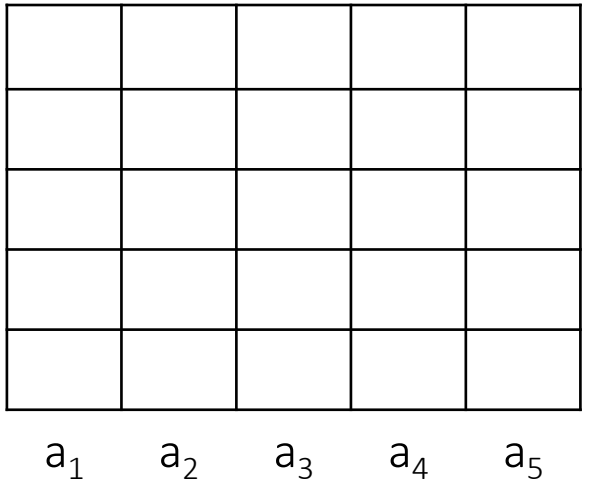
Conclusion: positive if S is in X_{15}; and negative otherwise

CYK algorithm: step by step

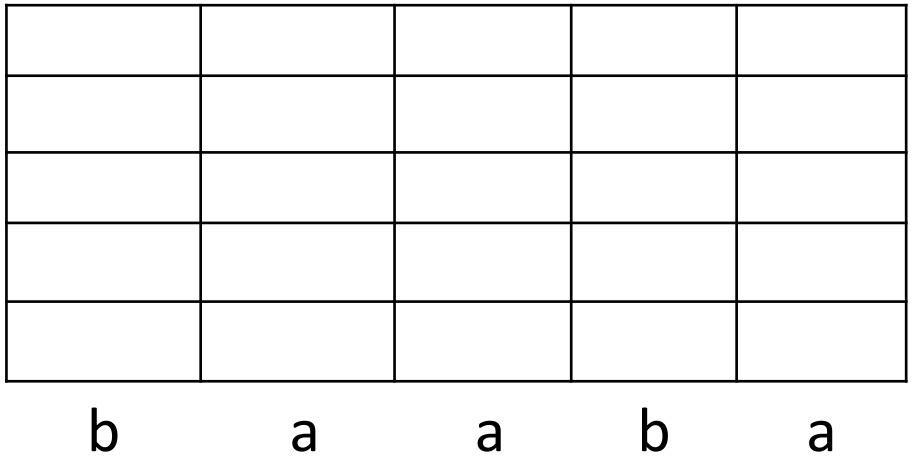
Test if a string belongs to a CFL

Step by step: CYK Algorithm

1. Create a table X of size N×N, where
   - N = length(input string)
   - X_{ij} represents the set of variables that produce string[i:j]



E.g., **Input string: baaba**



S → AB | BC

A → BA | a

B → CC | b

C → AB | a

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the bottom row: all X_{ii} elements
   - $X_{ii} = \{Y: Y \rightarrow a_i\}$

X_{11} and X_{44} …

{B}			{B}	
b	a	a	b	a

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the bottom row: all X_{ii} elements
   - $X_{ii} = \{Y: Y \rightarrow a_i\}$

X_{22}, X_{33} and X_{55} …

{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the rest of the X_{ij} elements, row by row
   - 2nd row: $X_{ij} = \{Y: Y \rightarrow X_{i,\,j-1} \times X_{i+1,\,j}\}$

$X_{12} = \{Y: Y \rightarrow X_{11} \times X_{22} = \{B\} \times \{A,C\} = \{BA, BC\}\}$

{S,A}				
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the rest of the X_{ij} elements, row by row
   - 2nd row: $X_{ij} = \{Y: Y \rightarrow X_{i, j-1} \times X_{i+1, j}\}$

$X_{23} = \{Y: Y \rightarrow \{A,C\} \times \{A,C\} = \{AA, CA, AC, CC\}\}$

{S,A}	{B}			
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the rest of the X_{ij} elements, row by row
   - 2nd row: $X_{ij} = \{Y: Y \rightarrow X_{i, j-1} \times X_{i+1, j}\}$

X_{34} and X_{45} …

{S,A}	{B}	{S,C}	{S,A}	
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the rest of the X_{ij} elements, row by row
   - $X_{i,j} = \{Y: Y \rightarrow U_{k=1}^{j-1\ until j-k=i}(X_{i,j-k} \times X_{j-k-1,j})\}$

$X_{13} = \{Y: Y \rightarrow \{S,A\} \times \{A,C\} = \{SA,SC,AA,AC\} \cup \{B\} \times \{B\} = \{BB\}\} = \{Y: Y \rightarrow \{SA,SC,AA,AC,BB\}$

{}				
{S,A}	{B}	{S,C}	{S,A}	
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the rest of the X_{ij} elements, row by row
   - $X_{i,j} = \{Y: Y \rightarrow U_{k=1}^{j-1\ until j-k=i}(X_{i,j-k} \times X_{j-k-1,j})\}$

$X_{13} = \{Y: Y \rightarrow \{S,A\} \times \{A,C\} = \{SA,SC,AA,AC\} \cup \{B\} \times \{B\} = \{BB\}\} = \{Y: Y \rightarrow \{SA,SC,AA,AC,BB\}$

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

{}				
{S,A}	{B}	{S,C}	{S,A}	
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the rest of the X_{ij} elements, row by row
   - $X_{i,j} = \{Y: Y \rightarrow U_{k=1}^{j-1\ until\ j-k=i}(X_{i,j-k} \times X_{j-k-1,j})\}$

$X_{24} = X_{23} \times X_{44} \cup X_{22} \times X_{34} = \{BB,AS,AC,CS,CC\}: B \rightarrow CC$

{}	{B}			
{S,A}	{B}	{S,C}	{S,A}	
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the rest of the X_{ij} elements, row by row
   - $X_{i,j} = \{Y: Y \rightarrow U_{k=1}^{j-1\ until j-k=i}(X_{i,j-k} \times X_{j-k-1,j})\}$

$X_{35} = X_{34} \times X_{55} \cup X_{33} \times X_{45} = \{SA,SC,CA,CC,AS,AA,CS,CA\}$:
$B \rightarrow CC$

{}	{B}	{B}		
{S,A}	{B}	{S,C}	{S,A}	
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

$S \rightarrow AB \mid BC$

$A \rightarrow BA \mid a$

$B \rightarrow CC \mid b$

$C \rightarrow AB \mid a$

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. Fill the rest of the X_{ij} elements, row by row
   - $X_{i,j} = \{Y: Y \rightarrow U_{k=1}^{j-1\ until j-k=i}(X_{i,j-k} \times X_{j-k-1,j})\}$

$X_{14} = X_{13} \times X_{44} \cup X_{12} \times X_{34} \cup X_{11} \times X_{24} = \{SS,SC,AS,AC,BB\}$
$= \{\}$

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

{}				
{}	{B}	{B}		
{S,A}	{B}	{S,C}	{S,A}	
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

Test if a string belongs to a CFL

Step by step: CYK Algorithm - $O(n^3)$

2. Continue until all rows are filled, and:

{S,A,C}				
{}	{S,A,C}			
{}	{B}	{B}		
{S,A}	{B}	{S,C}	{S,A}	
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

$S \rightarrow AB \mid BC$

$A \rightarrow BA \mid a$

$B \rightarrow CC \mid b$

$C \rightarrow AB \mid a$

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

Conclusion:
Yes, the string belongs to the CFL since $S \in X_{15}$

Test if a string belongs to a CFL

Step by step: CYK Algorithm

2. All groups involved (above bottom row)

- X_{13}: $X_{11} \times X_{23} \cup X_{12} \times X_{33}$ = {BB,SA,SC,AA,AC}: {}
- X_{24}: $X_{23} \times X_{44} \cup X_{22} \times X_{34}$ = {BB,AS,AC,CS,CC}: B → CC
- X_{35} = $X_{34} \times X_{55} \cup X_{33} \times X_{45}$ = {SA,SC,CA,CC,AS,AA,CS,CA} : B → CC

- X_{12}: $X_{11} \times X_{22}$ = {BA, BC}: A → BA, S → BC
- X_{23}: $X_{22} \times X_{33}$ = {AA,AC,CA,CC}: B → CC
- X_{34}: $X_{33} \times X_{44}$ = {AB,CB}: S → AB, C → AB
- X_{45}: $X_{44} \times X_{55}$ = {BA,BC}: A → BA, S → BC

- X_{14} = $X_{13} \times X_{44} \cup X_{12} \times X_{34} \cup X_{11} \times X_{24}$ = {SS,SC,AS,AC,BB} = {}
- X_{25}: $X_{22} \times X_{35} \cup X_{24} \times X_{55} \cup X_{23} \times X_{45}$ = {AB,CB, BA,BC,BS,BA}: S → AB, S → BC, C → AB, A → BA

- X_{15}: $X_{11} \times X_{25} \cup X_{14} \times X_{55} \cup X_{13} \times X_{45} \cup X_{12} \times X_{35}$ = {BS,BA,BC,SB,AB}: A → BA, S → BC, C → AB

S → AB | BC

A → BA | a

B → CC | b

C → AB | a

X_{15}				
X_{14}	X_{25}			
X_{13}	X_{24}	X_{35}		
X_{12}	X_{23}	X_{34}	X_{45}	
X_{11}	X_{22}	X_{33}	X_{44}	X_{55}
a_1	a_2	a_3	a_4	a_5

{S,A,C}				
{}	{S,A,C}			
{}	{B}	{B}		
{S,A}	{B}	{S,C}	{S,A}	
{B}	{A,C}	{A,C}	{B}	{A,C}
b	a	a	b	a

CYK demo:

- http://lxmls.it.pt/2015/cky.html
- Grammar:

  S -> A B
  S -> B C
  A -> B A
  B -> C C
  C -> A B
  A -> a
  B -> b
  C -> a

- Sentence: b a a b a