Part IX.
General Parsing Methods

Chomsky Normal Form (CNF)

Definition: Let $G = (N, T, P, S)$ be a CFG. G is in *Chomsky normal form* if every rule in P has one of these forms

- $A \rightarrow BC$, where $A, B, C \in N$;
- $A \rightarrow a$, where $A \in N, a \in T$;

Example:

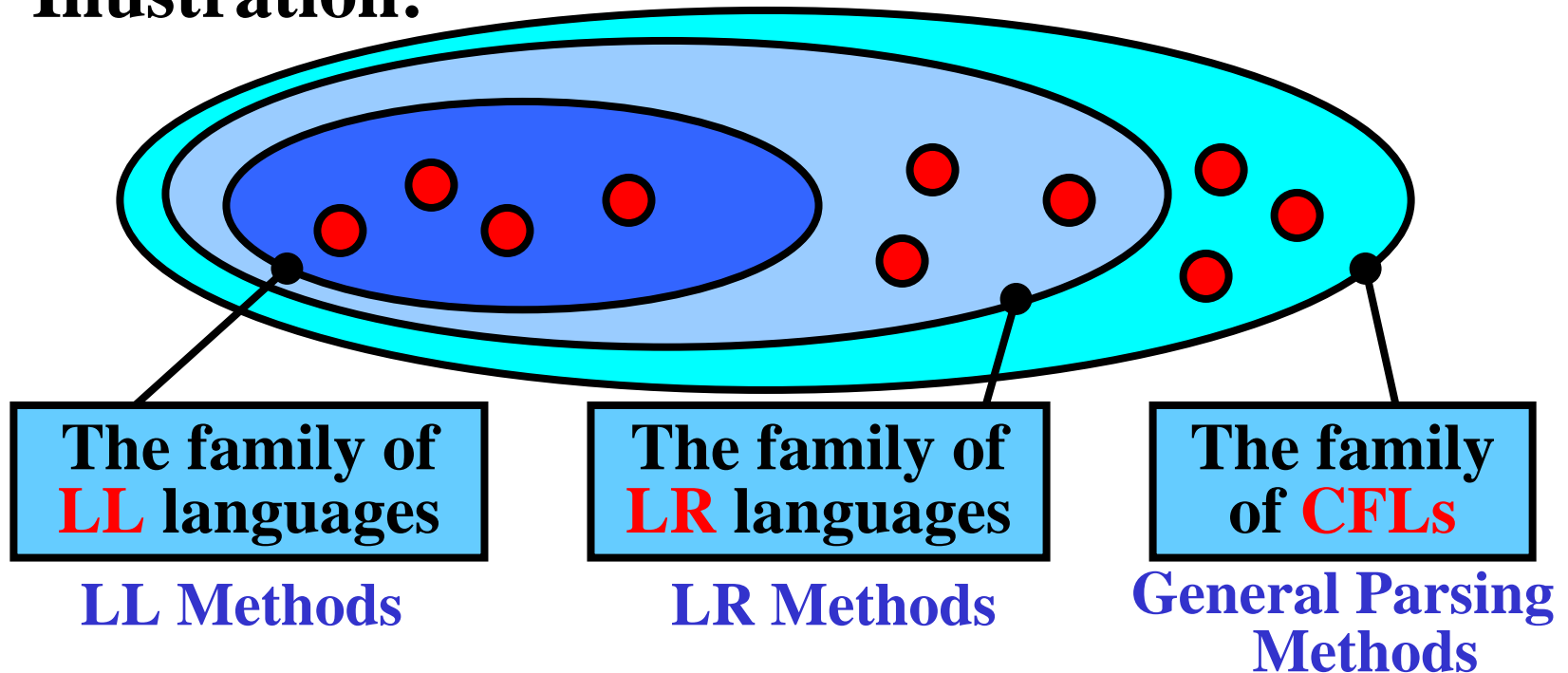$G = (N, T, P, S)$, where $N = \{A, B, C, S\}$, $T = \{a, b\}$, $P = \{S \rightarrow CB, C \rightarrow AS, S \rightarrow AB, A \rightarrow a, B \rightarrow b\}$ is in Chomsky normal form.

Note: $L(G) = \{a^n b^n : n \geq 1\}$

General Parsing Methods

• **General *Parsing* methods** (GP) are applicable to all context-free languages (CFLs)

Illustration:

| The family of **LL** languages | The family of **LR** languages | The family of **CFLs** |

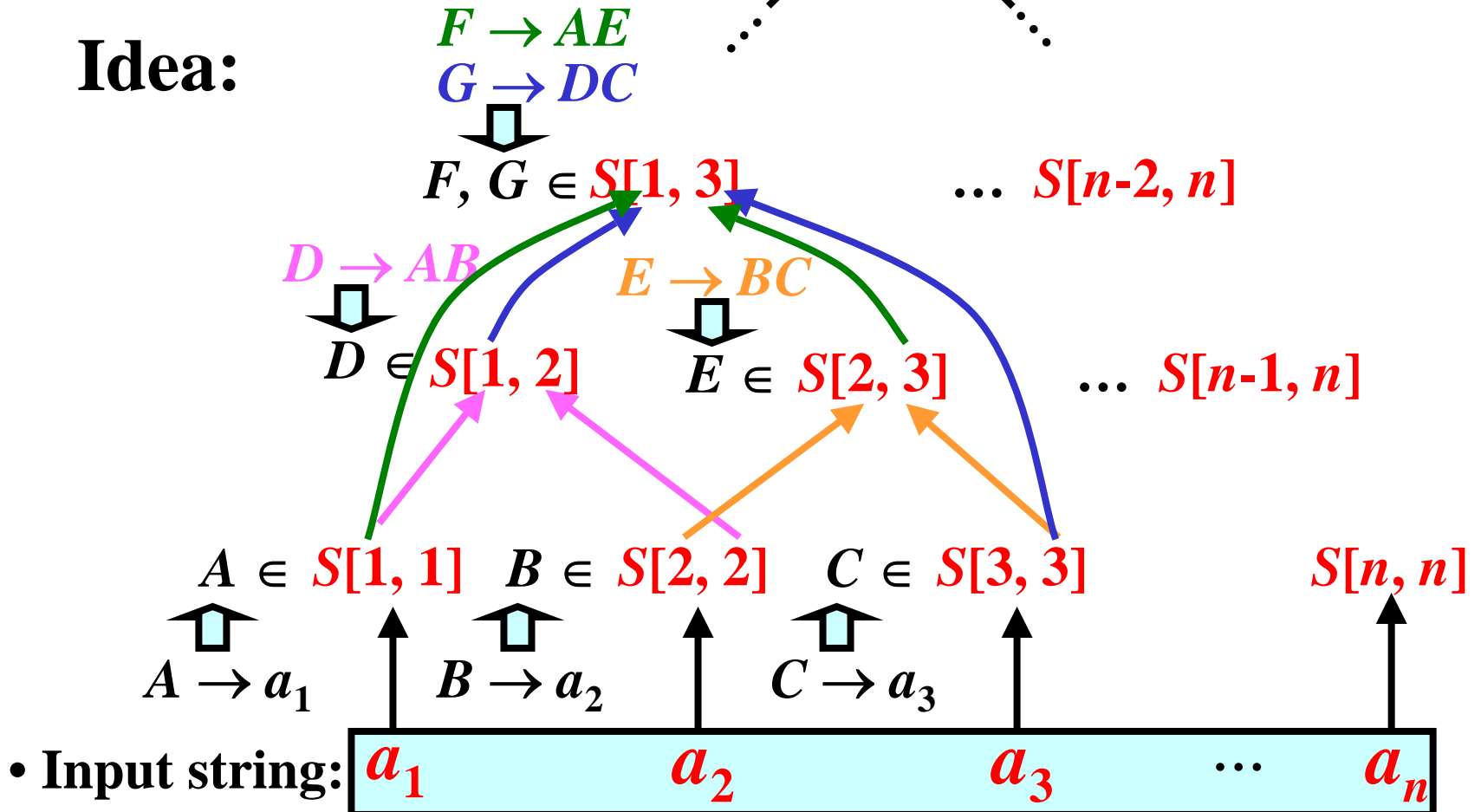LL Methods   LR Methods   General Parsing Methods

• **Note:** The family of **LR** languages = the family of a **deterministic CFL**

GP Based on Chomsky Normal Form

if $S \in S[1, n]$ then
$$S \Rightarrow^* a_1 \ldots a_n$$

$S[1, n]$

Idea:

$F \rightarrow AE$

$G \rightarrow DC$

$F, G \in S[1, 3]$ $\quad \ldots \quad S[n\text{-}2, n]$

$D \rightarrow AB$

$E \rightarrow BC$

$D \in S[1, 2]$ $\qquad E \in S[2, 3]$ $\quad \ldots \quad S[n\text{-}1, n]$

$A \in S[1, 1]$ $\quad B \in S[2, 2]$ $\quad C \in S[3, 3]$ $\qquad S[n, n]$

$A \rightarrow a_1$ $\qquad B \rightarrow a_2$ $\qquad C \rightarrow a_3$

• **Input string:** $a_1 \qquad a_2 \qquad a_3 \quad \cdots \quad a_n$

Algorithm: GP Based on CNF

- **Input:** $G = (N, T, P, S)$ in CNF, $w = a_1 \ldots a_n$
- **Output:** **YES** if $w \in L(G)$

  NO if $w \notin L(G)$

- **Method:**

- **for each** a_i, $i = 1, \ldots, n$ **do**

  $S[i, i] := \{A : A \to a_i \in P\}$

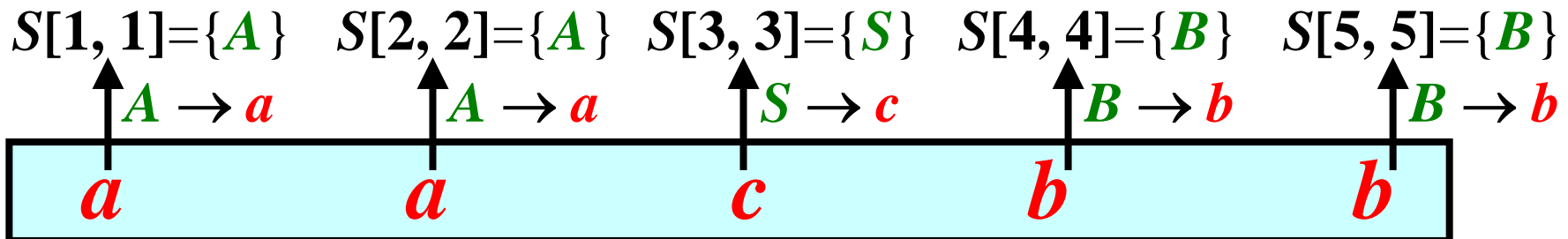- **Apply the following rule until no $S[i, k]$ can be changed:**

  if $A \to BC \in P$, $B \in S[i, j]$, $C \in S[j+1, k]$,

  where $1 \leq i \leq j < k \leq n$ **then** add A to $S[i, k]$

- **if** $S \in S[1, n]$ **then** write ('**YES**')

  else write ('**NO**')

GP Based on CNF: Example 1/5

$G = (N, T, P, S)$, where $N = \{A, B, C, S\}$, $T = \{a, b\}$,
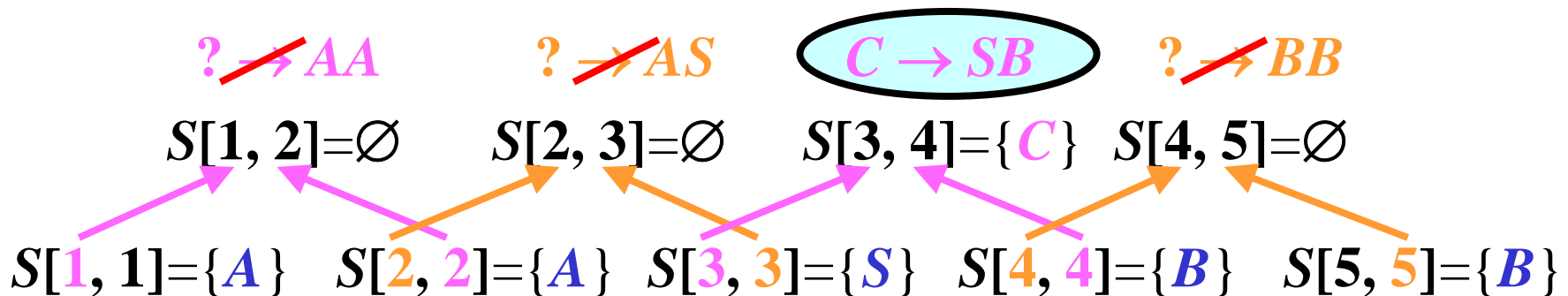$P = \{S \rightarrow AC, C \rightarrow SB, A \rightarrow a, B \rightarrow b, S \rightarrow c\}$
Question: $aacbb \in L(G)$?

$S[1, 1] = \{A\}$　　$S[2, 2] = \{A\}$　　$S[3, 3] = \{S\}$　　$S[4, 4] = \{B\}$　　$S[5, 5] = \{B\}$

$A \rightarrow a$　　　$A \rightarrow a$　　　$S \rightarrow c$　　　$B \rightarrow b$　　　$B \rightarrow b$

| a | a | c | b | b |

GP Based on CNF: Example 2/5

$G = (N, T, P, S)$, where $N = \{A, B, C, S\}$, $T = \{a, b\}$,
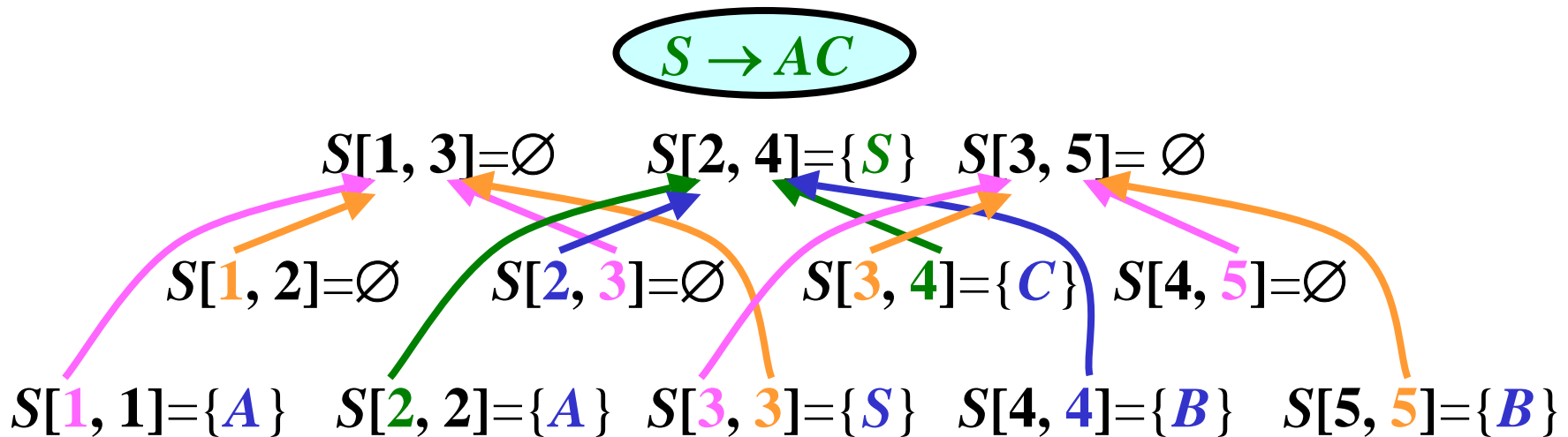$P = \{S \rightarrow AC, C \rightarrow SB, A \rightarrow a, B \rightarrow b, S \rightarrow c\}$
Question: $aacbb \in L(G)$?

$? \not\rightarrow AA \qquad ? \not\rightarrow AS \qquad \boxed{C \rightarrow SB} \qquad ? \not\rightarrow BB$

$S[1, 2]=\varnothing \qquad S[2, 3]=\varnothing \qquad S[3, 4]=\{C\} \quad S[4, 5]=\varnothing$

$S[1, 1]=\{A\} \quad S[2, 2]=\{A\} \quad S[3, 3]=\{S\} \quad S[4, 4]=\{B\} \quad S[5, 5]=\{B\}$

| a | a | c | b | b |

GP Based on CNF: Example 3/5

$G = (N, T, P, S)$, where $N = \{A, B, C, S\}$, $T = \{a, b\}$,
$P = \{S \to AC, C \to SB, A \to a, B \to b, S \to c\}$
Question: $aacbb \in L(G)$?

GP Based on CNF: Example 4/5

$G = (N, T, P, S)$, where $N = \{A, B, C, S\}$, $T = \{a, b\}$,
$P = \{S \rightarrow AC, C \rightarrow SB, A \rightarrow a, B \rightarrow b, S \rightarrow c\}$
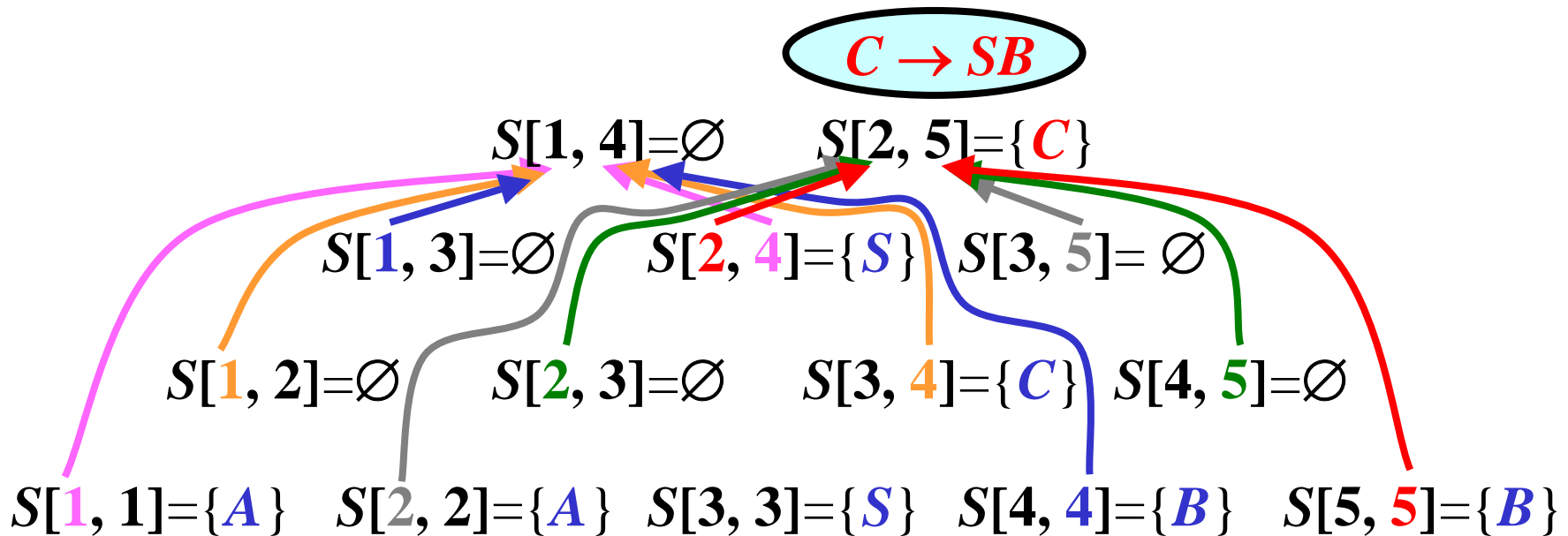**Question:** $aacbb \in L(G)$?

GP Based on CNF: Example 5/5

$G = (N, T, P, S)$, where $N = \{A, B, C, S\}$, $T = \{a, b\}$,
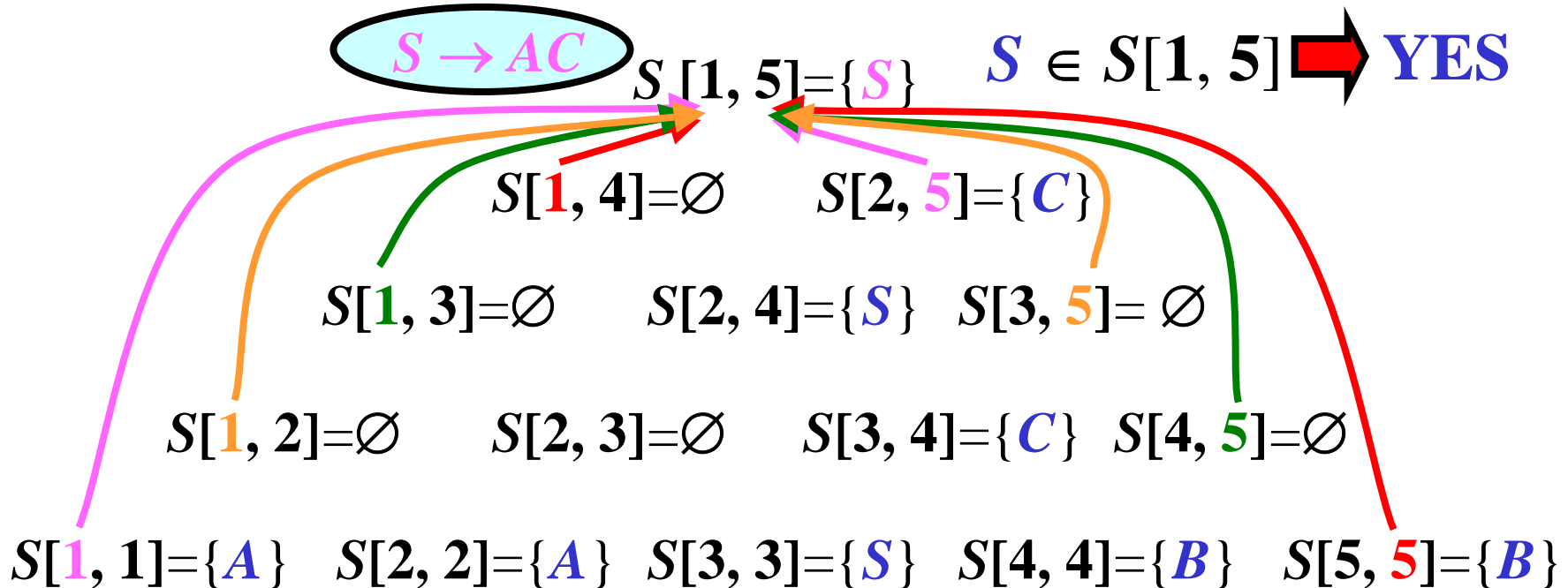$P = \{S \rightarrow AC, C \rightarrow SB, A \rightarrow a, B \rightarrow b, S \rightarrow c\}$
Question: $aacbb \in L(G)$?

$S \rightarrow AC$ $S[1, 5] = \{S\}$ $S \in S[1, 5]$ ➡ **YES**

$S[1, 4] = \varnothing$ $S[2, 5] = \{C\}$

$S[1, 3] = \varnothing$ $S[2, 4] = \{S\}$ $S[3, 5] = \varnothing$

$S[1, 2] = \varnothing$ $S[2, 3] = \varnothing$ $S[3, 4] = \{C\}$ $S[4, 5] = \varnothing$

$S[1, 1] = \{A\}$ $S[2, 2] = \{A\}$ $S[3, 3] = \{S\}$ $S[4, 4] = \{B\}$ $S[5, 5] = \{B\}$

| *a* | *a* | *c* | *b* | *b* |