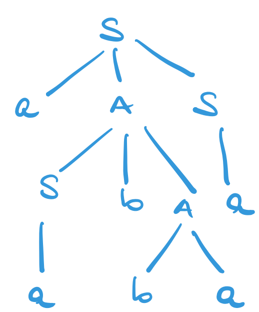
T4: GRAMÁTICAS INDEPENDIENTES DEL CONTEXTO

1. AMBIGÜEDAD DE GRAMÁTICAS.

* Árbol de derivación:
  - cada nodo del árbol contiene un símbolo
  - se efectúa una ramificación por cada producción que se aplique.

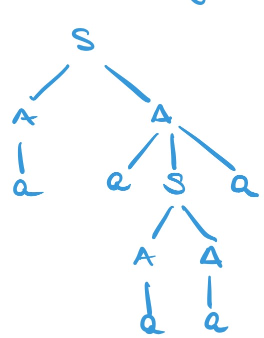
Ejemplo: $S \to aAS$ $\quad S \to a$ $\quad A \to SbA$ $\quad A \to SS$ $\quad A \to ba$

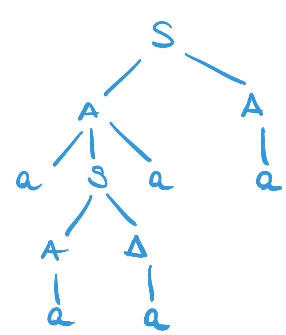
la palabra aabbaa



* Se dice que una gramática es ambigua si existe una palabra con dos árboles de derivación distintos.

Ejemplo: $S \to AA$ $\quad A \to aSa$ $\quad A \to a$, la palabra aaaaa tiene los dos siguientes árboles de derivación:





* Un lenguaje de tipo 2 es inherentemente ambiguo si toda gramática que lo genera es ambigua.

Mecanismos de limpieza de gramáticas.

* Símbolo útil: permiten generar palabras
  - → llegan a símbolos terminales.
  - → se pueden alcanzar desde S.

Algoritmo de eliminación:

1. Eliminar variables que no llegan a símbolos terminales.
2. Eliminar símbolos inalcanzables y producciones donde aparezcan.

PASO 1:

1. $V_t \neq 0$ Conjunto de variables que generan palabras.
2. Para cada $A \rightarrow w$, A se introduce en V_t.
3. Mientras V_t cambie
   4. Para cada producción $B \rightarrow \alpha$
      5. Si todas las variables de $\alpha \in V_t$ → Se mete B en V_t.
6. Eliminar variables que estén en V y no en V_t.
7. Eliminar todas las producciones donde aparezcan.

PASO 2 :

1. $V_S = \{S\}$ Conjunto de símbolos alcanzables.
2. Iterativamente, se añaden todos los símbolos y variables que estén en las producciones de V_S.
3. Se eliminan las variables y símbolos que no estén en V_S.

Ejemplo: R46.1.

a) $S \to aAb \mid cEB \mid CE \quad A \to dBE \mid \underline{ccC} \quad B \to \underline{ff} \mid D$

$C \to gFB \mid \underline{ae} \qquad D \to \underline{h}$

1. $\{C, D, B, A, S\}$ Eliminamos E, F y todos los sitios donde aparezcan.

$S \to aAb \quad A \to ccC \quad B \to ff \mid D$

$C \to ae \quad D \to h$

2. $\{S, A, C\}$ $\{a, b, c, e\}$ $\cancel{B}$ $\cancel{D}$ $\cancel{h}$ $\cancel{f}$

$S \to aAb \quad A \to ccC \quad C \to ae$

b) $S \to aB \quad A \to bcCCC \mid dA \quad B \to e \quad C \to fA \quad D \to Dgh$

1. $\{B, S\}$ $\quad S \to aB \quad B \to e$

2. $\{S, B\}$ $\{a, e\}$ ✓

c) $S \to a \mid aA \mid B \mid C \quad A \to aB \mid \varepsilon \quad B \to Aa \quad C \to bCD \quad D \to ccc$

1. $\{S, A, D, B\}$ $\quad S \to a \mid aA \mid B \quad A \to aB \mid \varepsilon \quad D \to ccc$

$B \to Aa$

2. $\{S, A, B\}$ $\cancel{D}$ $\quad S \to a \mid aA \mid B$

$\{a, \varepsilon\}$ $\cancel{c}$ $\quad A \to aB \mid \varepsilon \quad B \to Aa$

* Si S es inútil $\to$ el lenguaje generado es $\emptyset$ pero **NO** se puede eliminar la variable S.

2. FORMAS NORMALES

Definen características que deben verificar las producciones de una gramática.

- Producciones nulas
- Producciones unitarias
- Forma normal de Chomsky
- Forma normal de Greibach

Producciones nulas: $A \to \varepsilon$

Para eliminar todas las producciones nulas de la forma $A \to \varepsilon$:

- Buscamos todas las producciones $B \to \alpha A \beta$ y añadimos $B \to \alpha\beta$.

* Eliminamos las variables anulables (solo producciones nulas).

→ Tras esto la palabra vacía deja de generarse.

Ejemplo: Rel. 2.

a. $S \to aA \mid bA \mid a \qquad A \to aA \mid bAb \mid \varepsilon$

$S \to aA \mid a \mid bA \mid b \qquad A \to aA \mid a \mid bAb \mid bb$

b. $S \to ABaC \qquad A \to AB \qquad B \to b \mid \varepsilon \qquad C \to D \mid E \qquad D \to d$

$S \to ABaC \mid AaC \qquad A \to AB \mid A \qquad B \to b \qquad C \to D \mid E \qquad D \to d$

c. $S \to AB \qquad A \to aA \mid abB \mid aCa \qquad B \to bA \mid BB \mid \varepsilon$

$\cancel{C \to \varepsilon} \qquad D \to dB \mid BCB$

$S \to AB \mid A \qquad A \to aA \mid abB \mid ab \mid aa \qquad B \to bA \mid BB \mid B$

"aCa" no hace falta

$D \to dB \mid BCB \mid B \mid C \mid BC \mid CB \mid BB$

Producciones unitarias: $A \to B$

Para eliminar las producciones unitarias: añadimos todas las producciones de B a A.

Ejemplo: 2.4b.3.

a) $S \to CBa \mid D$ $\quad A \to bbC$ $\quad B \to Sc \mid ddd$ $\quad C \to eA \mid f \mid C$

$D \to E \mid SABC$ $\quad E \to gh$ $\qquad S \to D \quad C \to C \quad D \to E$

$S \to cBa \mid SABC \mid gh$ $\quad A \to bbC$ $\quad B \to Sc \mid ddd$ $\quad C \to eA \mid f$

$D \to gh \mid SABC$ $\quad E \to gh$

b) $S \to Aa \mid Ba \mid B$ $\quad A \to Aa \mid \underline{\varepsilon}$ $\quad B \to aA \mid BB \mid \underline{\varepsilon}$

Primero eliminamos nulas:

$\underline{S} \to Aa \mid a \mid Ba \mid \underline{B} \mid \cancel{\varepsilon}$ $\quad A \to Aa \mid a$ $\quad \underline{B} \to aA \mid a \mid BB \mid \cancel{B}$

Lo quitamos por quitar nulas.

Añadimos $\underline{S_0} \to \underline{S} \mid \varepsilon$

$S_0 \to \varepsilon \mid Aa \mid a \mid Ba \mid aA \mid BB$ $\quad A \to Aa \mid a$

$S \to Aa \mid a \mid Ba \mid aA \mid BB$ $\quad B \to aA \mid a \mid BB$

FORMA NORMAL DE CHOMSKY

Una gramática está en forma normal de Chomsky si verifica que todas las producciones son:

$$A \to BC \quad ó \quad A \to a \quad ó \quad S \to \varepsilon$$

0. La gramática tiene que estar limpia.
1. No debe tener producciones nulas ni unitarias.
2. Eliminar todos los terminales salvo $A \to a$ creando nuevas variables.
3. Desglosar producciones de longitud > 2.

Ejemplo: R4.10.

$S \to AAA \mid B \quad A \to aA \mid B \quad B \to \varepsilon$

0. Todos los símbolos son terminables y alcanzables.

1. Eliminamos nulas:

$S \to AAA \mid \varepsilon \quad A \to \quad \mid \varepsilon \quad (B \to \varepsilon)$

Repetimos: $(A \to \varepsilon)$

$S \to AAA \mid AA \mid \underline{A} \mid \varepsilon \quad A \to aA \mid a$

Eliminamos unitarios:

$S \to AAA \mid AA \mid aA \mid a \mid \varepsilon \quad A \to aA \mid a$

2. $S \to AAA \mid AA \mid C_aA \mid a \mid \varepsilon \quad A \to C_aA \mid a \quad C_a \to a$

3. Desglosamos producciones de longitud > 2

$S \to AD \mid AA \mid C_aA \mid a \mid \varepsilon \quad A \to C_aA \mid a \quad C_a \to a$

$D \to AA$