

Forma Normal de Chomsky e fechamento de LLC's

Luís Felipe

UFF

17 de Maio de 2023

Forma normal de Chomsky

- Simplificação de GLC's
- Útil em algoritmos que trabalham com gramáticas

Uma GLC está na **forma normal de Chomsky** se toda regra é da forma:

- $A \rightarrow BC$, onde $A \in V$, $B, C \in V \setminus \{S\}$, e S é a variável inicial;
- $A \rightarrow a$, onde $A \in V$, $a \in T$;
- $S \rightarrow \epsilon$, onde S é a variável inicial.

Forma normal de Chomsky e LLC's

Teorema. Qualquer LLC é gerada por uma GLC na forma normal de Chomsky.

Ideia da Prova:

1. Nova variável inicial S_0 e nova regra $S_0 \rightarrow S$, onde S é variável inicial da GLC original; (isto garante que não teremos var. inicial do lado direito)
2. Remova todas regras $A \rightarrow \varepsilon$, $A \neq S$. Para cada regra $R \rightarrow uAv$, adicione $R \rightarrow uv$, onde $u, v \in (V \cup T)^*$. Se A ocorre numa regra mais de uma vez, acrescente uma nova regra para cada ocorrência.

Ex: $R \rightarrow uAvAw$ gera as regras: $R \rightarrow uvAw \mid uAvw \mid uvw$.

Se existe uma regra $R \rightarrow A$, então adicione $R \rightarrow \varepsilon$ e repita o processo (a menos que uma regra $R \rightarrow \varepsilon$ tenha sido removida antes). Assim, eliminamos as regras com ε que não envolvem S .

Continuação

3. Elimine toda regra $A \rightarrow B$. Assim, quando $B \rightarrow u$ aparecer, acrescente $A \rightarrow u$ (a menos que esta regra tenha sido previamente removida), $u \in (V \cup T)^*$.
4. Substitua regras $A \rightarrow u_1 u_2 \dots u_k$, onde $k \geq 3$ e $u_i \in (V \cup T)$ por $A \rightarrow u_1 A_1$, $A_1 \rightarrow u_2 A_2, \dots, A_{k-2} \rightarrow u_{k-1} u_k$. A_i 's são novas variáveis.
Se $k = 2$, em cada regra onde u_i é símbolo terminal o substituímos por uma nova variável U_i e adicionamos a regra $U_i \rightarrow u_i$.

Exemplo

$$S \rightarrow ASA \mid aB$$

$$A \rightarrow B \mid S$$

$$B \rightarrow b \mid \varepsilon$$

1. Nova variável inicial:

$$S_0 \rightarrow S$$

$$S \rightarrow ASA \mid aB$$

$$A \rightarrow B \mid S$$

$$B \rightarrow b \mid \varepsilon$$

2. Regras em ε :

$$S_0 \rightarrow S$$

$$S \rightarrow ASA \mid aB \mid a$$

$$A \rightarrow B \mid S \mid \varepsilon$$

$$B \rightarrow b \mid \varepsilon$$

2. Regras em ε :

$$S_0 \rightarrow S$$

$$S \rightarrow ASA \mid aB \mid a \mid SA \mid AS \mid S$$

$$A \rightarrow B \mid S \mid \varepsilon$$

$$B \rightarrow b$$

3. Regras em $S \rightarrow S$ e $S_0 \rightarrow S$:

$$S_0 \rightarrow S \mid ASA \mid aB \mid a \mid SA \mid AS$$

$$S \rightarrow ASA \mid aB \mid a \mid SA \mid AS \mid S$$

$$A \rightarrow B \mid S$$

$$B \rightarrow b$$

Continuação

3. Regras em $A \rightarrow B$:

$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow B \mid S \mid b$
 $B \rightarrow b$

3. Regras em $A \rightarrow S$:

$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow S \mid b \mid ASA \mid aB \mid a \mid$
 $\quad SA \mid AS$
 $B \rightarrow b$

4. Rotulações não condizentes (SA e aB):

$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $S \rightarrow ASA \mid aB \mid a \mid SA \mid AS$
 $A \rightarrow b \mid ASA \mid aB \mid a \mid SA \mid AS$
 $B \rightarrow b$

4. Novas rotulações (onde for devido):

$S_0 \rightarrow AA_1 \mid UB \mid a \mid SA \mid AS$
 $S \rightarrow AA_1 \mid UB \mid a \mid SA \mid AS$
 $A \rightarrow b \mid AA_1 \mid UB \mid a \mid SA \mid AS$
 $B \rightarrow b$
 $A_1 \rightarrow SA$
 $U \rightarrow a$

Luís Felipe
11/05/23

Continuação

Solução:

$$S_0 \rightarrow AA_1 \mid UB \mid a \mid SA \mid AS$$

$$S \rightarrow AA_1 \mid UB \mid a \mid SA \mid AS$$

$$A \rightarrow b \mid AA_1 \mid UB \mid a \mid SA \mid AS$$

$$B \rightarrow b$$

$$A_1 \rightarrow SA$$

$$U \rightarrow a$$

Propriedades de fechamento em LLC

A classe das LLC é fechada sob:

- **União:** nova variável inicial S' que aponta para S_1 e para S_2 (ou seja, $S' \rightarrow S_1 \mid S_2$).
- **Concatenação:** nova variável inicial S' que aponta para S_1S_2 (ou seja, $S' \rightarrow S_1S_2$).
- **Estrela:** nova variável inicial S' que aponta para S_1S' e para ϵ (ou seja, $S' \rightarrow SS' \mid \epsilon$).

OBS.: As LLC's **não** são fechadas por complemento, diferença e interseção. Veremos isso quando falarmos do Lema do Bombeamento para LLC.

Exemplos:

Considere as gramáticas geradoras das linguagens: L_1 dos palíndromos e $L_2 = \{0^n 1^n, n \geq 0\}$.

1. Construa G que gere $L_1 \cup L_2$
2. Construa G que gere $L_1 L_2$
3. Construa G que gere $L_2 L_1$
4. Construa G que gere L_1^*
5. Construa G que gere L_2^*