

Propriedades das LLC

Objetivos:

- Simplificar GLC para facilitar a prova de propriedades de LLC;
- Propriedades:
 - Lema do Bombeamento;
 - Propriedades de Fechamento.

Forma Normal de Chomsky

Toda LLC (sem ϵ) é gerada por uma GLC cujas produções são da forma:

- $A \rightarrow BC$
- $A \rightarrow a$, onde A, B, C são variáveis e a é um terminal.

Etapas:

1. Eliminar símbolos inúteis: variáveis ou terminais que não aparecem em qualquer derivação de string terminal a partir do símbolo inicial;
2. Eliminar transições vazias: aquelas da forma $A \rightarrow \epsilon$;
3. Eliminar produções unitárias: aquelas da forma $A \rightarrow B$.

FNC - ETAPA 1 Eliminando Símbolos Inúteis

(Símbolo Útil) X é útil para $G = (V, T, P, S)$ se existe alguma derivação da forma

$$S \Rightarrow^* \alpha X \beta \Rightarrow^* w, w \in T^*, X \in V \cup T$$

Se X não é útil, é inútil.

Para identificar símbolos inúteis devemos identificar os símbolos aptos a serem úteis:

- a) X é gerador se $X \Rightarrow^* w$, para alguma string terminal w ;
- b) X é alcançável se existe uma derivação $S \Rightarrow^* \alpha X \beta$ para algum α e β .

Um símbolo útil é gerador e alcançável.

Exemplo: $S \rightarrow AB|a$
 $A \rightarrow b$

Construção de GLC sem Símbolos Inúteis

a) Computando símbolos geradores: seja $G = (V, T, P, S)$. Para computar símbolos geradores, segue-se a seguinte indução:

Base: todo terminal é gerador; desde que ele gera ele próprio;

Indução: Se $A \rightarrow \alpha$ e todo símbolo de α é gerador, então A é gerador. (Note que esta regra inclui $\alpha = \epsilon$).

Exemplo: $S \rightarrow AB|a, A \rightarrow b$.

b) Computando símbolos alcançáveis: seja $G = (V, T, P, S)$. Para computar símbolos alcançáveis, segue-se a seguinte indução:

Base: S é alcançável;

Indução: Se A é alcançável, então qualquer produção a partir de A é formada por símbolos alcançáveis.

Exemplo: $S \rightarrow AB|a, A \rightarrow b$.

FNC - ETAPA 2 Eliminando Produções-ε
<p>(Variável Anulável) A é anulável se $A \Rightarrow^* \epsilon$.</p> <p>Então quando A aparece no corpo de uma produção: $B \rightarrow CAD$, A deve (ou não deve) derivar ϵ. Assim constrói-se duas versões da produção:</p> <ul style="list-style-type: none"> ■ $B \rightarrow CD$ (A foi usado para derivar ϵ); ■ $B \rightarrow CAD$ (A não foi usado para derivar ϵ) e para isso elimina-se $A \rightarrow \epsilon$.

Construção de GLC sem Produções-ε
<p>a) (Computando Símbolos Anuláveis) seja $G = (V, T, P, S)$. Para computar símbolos anuláveis, segue-se o seguinte algoritmo: <u>Base:</u> se $A \rightarrow \epsilon$, então A é anulável; <u>Indução:</u> Se $B \rightarrow C_1 C_2 \dots C_k$, onde cada C_i é anulável, então B é anulável. (Observe que devem ser consideradas produções com variáveis no corpo).</p> <p>b) (Construção de uma nova Gramática) seja $G = (V, T, P, S)$, construa a nova gramática $G_1 = (V, T, P_1, S)$, onde P_1 é determinado:</p> <ul style="list-style-type: none"> ■ Para cada produção $A \rightarrow X_1 X_2 \dots X_k$ de P, onde $k \geq 1$, suponha que m dos X_i's são anuláveis. A nova gramática terá 2^m versões desta produção, onde os X_i's anuláveis estão presentes ou ausentes; ■ Existe uma exceção se $m=k$, isto é, se todos são anuláveis, não se inclui o caso onde todos os X_i's estão ausentes; ■ A produção $A \rightarrow \epsilon$ não aparece em P_1. <p>Exemplo: $S \rightarrow AB, A \rightarrow aAA ε, B \rightarrow bBB ε$</p>

FNC - ETAPA 3 Eliminando Produções Unitárias
<p>(Produção Unitária) é uma produção da forma $A \rightarrow B$ onde A e B são variáveis.</p> <p>(Par Unitário) todo par de variáveis (A, B) tal que $A \Rightarrow^* B$ usando somente seqüência de produções unitárias.</p> <p>Encontrando todos pares unitários, a seqüência $A \Rightarrow B_1 \Rightarrow B_2 \Rightarrow \dots \Rightarrow B_n \Rightarrow \alpha$ Pode ser substituída por $A \rightarrow \alpha$.</p>

Construção de GLC sem Produções Unitárias
<p>Seja $G = (V, T, P, S)$. Construa $G_1 = (V, T, P_1, S)$:</p> <p>a) (Computando Pares Unitários) <u>Base:</u> (A, A) é um par unitário, para todo A. Isto é, $A \Rightarrow^* A$ por zero passos. <u>Indução:</u> Se (A, B) é um par unitário e $B \rightarrow C$ é uma produção onde C é uma variável. Então (A, C) é um par unitário.</p> <p>b) (Construindo P_1) Para cada par unitário (A, B) adicione a P_1 todas produções $A \rightarrow \alpha$, onde $B \rightarrow \alpha$ é uma produção não unitária em P. (Note que se $A=B$, P_1 contém todas as produções não unitárias em P).</p> <p>Exemplo: $I \rightarrow a b a b 0 1, F \rightarrow I (E), T \rightarrow F TxF$ e $E \rightarrow T E+T$</p>

Forma Normal de Chomsky

- PASSO 1) Elimine produções- ϵ
- PASSO 2) Elimine produções unitárias
- PASSO 3) Elimine símbolos inúteis

FNC possui produções nas formas:

- $A \rightarrow BC$, onde A, B e C são variáveis;
- $A \rightarrow a$, onde A é variável e a é terminal.

Computando Forma Normal de Chomsky

Satisfeitos os 3 passos deve-se:

- a) Especificar que todos os corpos de tamanho 2 ou superior consista apenas de variáveis:
 - Para cada terminal a que aparece num corpo de tamanho 2 ou superior, crie uma nova variável A .
 - Adicione a produção $A \rightarrow a$. Agora substitua a por A em todo corpo de tamanho 2 ou superior que a aparece.
- b) Quebrar corpos de tamanho 3 ou superior em uma cascata de produções, cada qual com um corpo de 2 variáveis:
 - Quebram-se as produções $A \rightarrow B_1 B_2 \dots B_k$, $k \geq 3$, em produções com 2 variáveis no corpo;
 - Introduz-se $k-2$ novas variáveis C_1, C_2, \dots, C_{k-2} . A produção original é substituída por $k-1$ produções:

$$A \rightarrow B_1 C_1, C_1 \rightarrow B_2 C_2, C_2 \rightarrow B_3 C_3, \dots, C_{k-2} \rightarrow B_{k-1} B_k$$

Exemplo

$G = (\{E, T, F, I\}, \{a, b, 1, 0, (,), +, x\}, P, E)$

Onde P é o conjunto de produções:

$E \rightarrow E+T \mid TxF \mid (E) \mid a \mid b \mid la \mid lb \mid I0 \mid I1$

$T \rightarrow TxF \mid (E) \mid a \mid b \mid la \mid lb \mid I0 \mid I1$

$F \rightarrow (E) \mid a \mid b \mid la \mid lb \mid I0 \mid I1$

$I \rightarrow a \mid b \mid la \mid lb \mid I0 \mid I1$

Lema do Bombeamento

(Teorema) Seja L uma LLC. Então existe uma constante n tal que se z é qualquer string em L tal que $|z| \leq n$, então podemos escrever $z = uvwxy$, sujeito as seguintes condições:

1. $|vwx| \leq n$
2. $vx \neq \epsilon$. Desde que v e x são as strings a serem bombeadas, esta condição diz que pelo menos uma das strings não deve ser vazia
3. $\forall i \geq 0, uv^iwx^iy \in L$. Isto é, as strings v e x podem ser bombeadas qualquer número de vezes, inclusive 0, e a string resultante ainda é um membro de L.

Exemplos

- (Exemplo 1) $L = \{0^n 1^n 2^n \mid n \geq 1\}$ não é uma LLC.
- (Exemplo 2) $L = \{0^i 1^j 2^i \mid i \geq 1 \text{ e } j \geq 1\}$

(Teorema) As LLC são fechadas para:

1. União
2. Concatenação
3. Fecho
4. Homomorfismo