

INF155: Informática Teórica

Clase 14: Lema de Bombeo para Gramáticas de Contexto Libre

Aldo Berrios Valenzuela

Jueves 21 de Abril de 2016

1. Ejemplo CNF

Ejemplo 1.1. Considerando la gramática del ejemplo <> /* Referenciar cuando se integre este archivo al apunte */:

$$E \rightarrow E + T \quad (1)$$

$$E \rightarrow T \quad (2)$$

$$T \rightarrow T * F \quad (3)$$

$$T \rightarrow F \quad (4)$$

$$F \rightarrow (E) \quad (5)$$

$$F \rightarrow a \quad (6)$$

Fase 1: Sólo $A \rightarrow a$ o $A \rightarrow B_1 B_2 \cdots B_n$ con $B_i \in \mathbb{N}$. Inventamos no-terminales:

$$M \rightarrow + \quad (7')$$

$$P \rightarrow * \quad (8')$$

$$A \rightarrow (\quad (9')$$

$$C \rightarrow) \quad (10')$$

y reemplazamos:

$$E \rightarrow EMT \quad (1')$$

$$E \rightarrow T \quad (2')$$

$$T \rightarrow TPF \quad (3')$$

$$T \rightarrow F \quad (4')$$

$$F \rightarrow AEC \quad (5')$$

$$F \rightarrow a \quad (6')$$

Fase 2: Eliminar producciones unitarias:

$$E \rightarrow T \quad (2')$$

$$T \rightarrow F \quad (4')$$

Para (4'), sabemos que las otras producciones son:

$$\begin{aligned} T &\rightarrow TPF & (3') \\ T &\rightarrow AEC & (4' \text{ y } 5') \\ T &\rightarrow a & (4' \text{ y } 6') \end{aligned}$$

Cadena:

$$\begin{aligned} E &\rightarrow T \Rightarrow \dots \\ \rightsquigarrow E &\rightarrow TPF & (\text{reemplazando el lado derecho de } 3') \\ E &\rightarrow AEC \\ E &\rightarrow a \end{aligned}$$

Por lo tanto, resulta:

$$\begin{aligned} E &\rightarrow EMT \\ \left. \begin{aligned} E &\rightarrow TPF \\ E &\rightarrow AEC \\ E &\rightarrow a \end{aligned} \right\} T &\rightarrow F \\ T &\rightarrow TPF \\ \left. \begin{aligned} T &\rightarrow AEC \\ T &\rightarrow a \end{aligned} \right\} T &\rightarrow F \\ F &\rightarrow AEC \\ F &\rightarrow a \\ \\ M &\rightarrow + \\ P &\rightarrow * \\ A &\rightarrow (\\ C &\rightarrow) \end{aligned}$$

Fase 3: Producciones $A \rightarrow B_1 \cdots B_n$, con $n \geq 3$ en CNF. Así:

$$\begin{aligned} E &\rightarrow EMT \\ \rightsquigarrow E &\rightarrow EX_1 \\ X_1 &\rightarrow MT \\ \\ E &\rightarrow TPF \\ \rightsquigarrow E &\rightarrow TX_2 \\ X_2 &\rightarrow PF \\ \\ T &\rightarrow TPF \\ \rightsquigarrow T &\rightarrow TX_3 \\ X_3 &\rightarrow PF \end{aligned}$$

Finalmente, la forma normal de Chomsky de la gramática del ejemplo <> /* Referenciar cuando se integre

este archivo al apunte */ es:

$M \rightarrow +$
 $P \rightarrow *$
 $A \rightarrow ($
 $C \rightarrow)$
 $E \rightarrow EX_1$
 $X_1 \rightarrow MT$
 $E \rightarrow TX_2$
 $X_2 \rightarrow PF$
 $T \rightarrow TX_3$
 $X_3 \rightarrow PF$

2. Lema de Bombeo para Gramáticas de Contexto Libre

Teorema 2.1 (Lema de Bombeo para CFL). Sea L una CFL. Entonces hay una constante $n > 1$ tal que para todo $\sigma \in L$, si $|\sigma| \geq n$ hay $u, v, w, x, y \in \Sigma^*$ con los que podemos escribir:

$$\sigma = uvwxy$$

donde $|vwx| \leq n$, $vx \neq \epsilon$, y para todo $k \geq 0$:

$$uv^kwx^ky \in L$$

Demostración. Sea $L = \mathcal{L}(G)$, con $G = (N, \Sigma, P, S)$ una CFG en CNE. Consideremos un árbol de derivación de σ en G (Figura 1). Es importante destacar que el árbol de la Figura 1 corresponde a un árbol binario con $|\sigma|$ hojas.

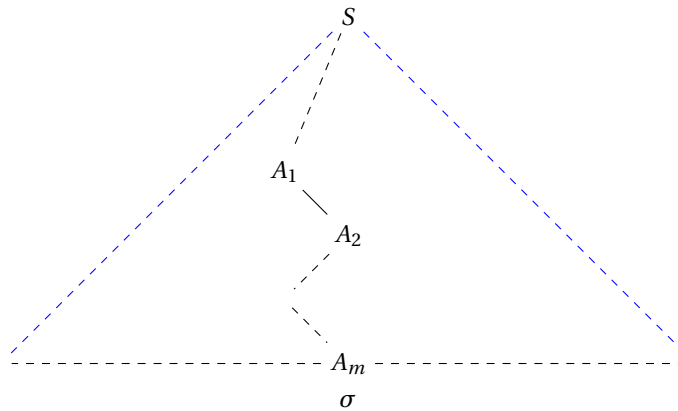


Figura 1: Para cada $\sigma \in L$ tal que $|\sigma| > N$, existe un largo camino en el árbol de derivación. Nótese que este árbol tiene $|\sigma|$ hojas.

Sabemos que la altura del árbol (número de no-terminales en la rama más larga) crece con el número de hojas, correspondiente a $|\sigma|$. Si la altura del árbol es mayor que $|N|$, en la rama más larga un no-terminal se repite (Figura 2).

Como derivación, sabemos que hay $A \in N$ tal que $A \Rightarrow^+ vAx$, y $A \Rightarrow^+ w$ con $vx \neq \epsilon$, $v, w, x \in \Sigma^*$.

$$S \Rightarrow^* uAy \Rightarrow^+ uvAxy \Rightarrow^* uvwxy$$

Pero entonces es una derivación:

$$S \Rightarrow^* uAy \Rightarrow^* uwy$$

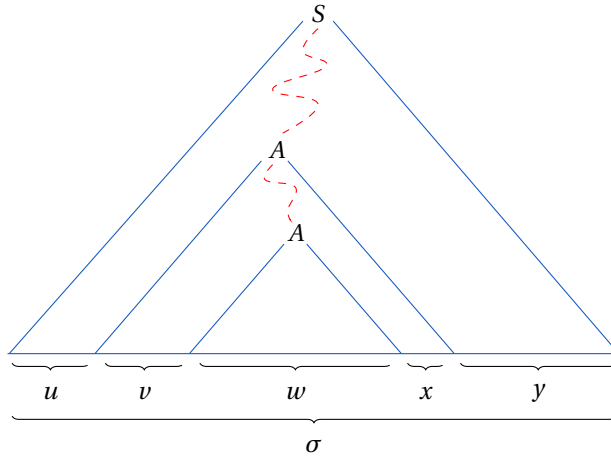


Figura 2: Por lema de bombeo para CFL, se puede descomponer $\sigma = uvwxy$.

y también:

$$\begin{aligned} S &\Rightarrow^* uAy \Rightarrow^+ uvAxy \\ &\Rightarrow^+ uv^k Ax^k y \\ &\Rightarrow^* uv^k wx^k y \end{aligned}$$

O sea, $uv^k wx^k y \in L$ para todo k .

Podemos aplicar la misma idea de muchas hojas \rightsquigarrow gran altura al árbol con raíz en A , y sus hojas (uvx) no pueden ser demasiadas antes de que allí aparezca un no-terminal repetido. Esto da $|vwx| \leq N$. \square

Teorema 2.2. El lenguaje

$$L = \{a^n b^n c^n : n \geq 1\} \quad (2.1)$$

no es CFL.

Demostración. Supongamos que L es CFL, con lo que cumple el lema de bombeo para CFL. Sea N la constante del lema. En seguida, consideremos $\sigma = a^N b^N c^N$, $\sigma \in L$, $|\sigma| = 3N \geq N$. Entonces, por PLCFL¹ hay $u, v, w, x, y \in \Sigma^*$ tales que $|vwx| \leq N$, $vx \neq \epsilon$ tales que para todo $k \geq 0$ se tiene que $uv^k wx^k y \in L$; pero como $|vwx| \leq N$, está formado por a lo más dos tipos de símbolos. Al repetir vx , hay al menos un tipo de símbolo que no se repite, y ese queda corto.

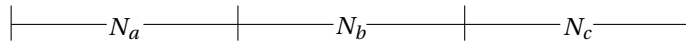


Figura 3: El lenguaje (2.1) exige que la cantidad de a 's sea igual a la cantidad de b 's y a la cantidad de c 's.

El resultado no pertenecen a L . Una contradicción. \square

Ejercicio 2.1. Demuestre que el lenguaje (2.1) no es regular.

Ejercicio 2.2. Sea G en CNF, $\sigma \in \mathcal{L}(G)$. Demuestre que la derivación de σ toma $2|\sigma| - 1$ pasos.

Observación: Un *lenguaje regular* es un caso particular de un CFL. Si demostramos que un lenguaje no es de *contexto libre*, se tiene que tampoco es regular. De la misma forma, si demostramos que un lenguaje es *regular*, entonces es de *contexto libre*.

¹Pumping Lemma for Context Free Language

3. CFG \rightarrow PDA

Teorema 3.1 (CFG \rightarrow PDA). Si $L = \mathcal{L}(G)$, con $G = (N, \Sigma, P, S)$ una CFG, podemos construir un PDA M tal que $L = \mathcal{L}(M)$.

Demostración. La idea es que el PDA tiene en el tope del stack la palabra después de los terminales iniciales en una forma sentencial.

Sea $M = (\{q_0\}, \Sigma, \Sigma \cup N, \delta, q_0, S, \emptyset)$. Acá δ está dado por:

- $(q_0, \alpha) \in \delta(q_0, \epsilon, A)$ si $A \rightarrow \alpha$ es producción.
- $\delta(q_0, a, a) = \{(q_0, \epsilon)\}$ para todo $a \in \Sigma$.

Consideremos como ejemplo, nuestra gramática reglona dada por:

$$E \rightarrow E + T \quad (1)$$

$$E \rightarrow T \quad (2)$$

$$T \rightarrow T * F \quad (3)$$

$$T \rightarrow F \quad (4)$$

$$F \rightarrow (E) \quad (5)$$

$$F \rightarrow a \quad (6)$$

Sabemos $E \Rightarrow^* a * (a + a)$. Podemos describirlo a través del Cuadro 1. Es importante destacar, que el stack del cuadro 1 crece hacia la izquierda (\leftarrow).

Este PDA traza una derivación de extrema izquierda de σ .

□

Entrada	Stack	Operación
$a * (a + a)$	E	$E \rightarrow T$
$a * (a + a)$	T	$T \rightarrow T * F$
$a * (a + a)$	$T * F$	$T \rightarrow F$
$a * (a + a)$	$F * F$	$F \rightarrow a$
$a * (a + a)$	$a * F$	shift a
$* (a + a)$	$* F$	shift $*$
$(a + a)$	F	$F \rightarrow (E)$
$(a + a)$	F	$F \rightarrow (E)$
$(a + a)$	(E)	shift $($
$a + a)$	$E)$	$E \rightarrow E + T$
$a + a)$	$E + T)$	$E \rightarrow T$
$a + a)$	$T + T)$	$T \rightarrow F$
$a + a)$	$T + T)$	$F \rightarrow a$
$a + a)$	$a + T)$	shift a
$+ a)$	$+ T)$	shift $+$
$a)$	$T)$	$T \rightarrow F$
$a)$	$F)$	$F \rightarrow a$
$a)$	$a)$	shift a
$)$	$)$	shift $)$
ϵ	ϵ	

Cuadro 1: Aplicamos producciones en el stack hasta que el tope del stack sea un no-terminal. En ese momento, aplicamos una operación shift la cual hace un pop tanto de la Entrada de no-terminales como del Stack