# Entrega 5

José Antonio Álvarez

28 de diciembre de 2017

**Ejercicio 13.** Encuentra una gramática libre de contexto en forma normal de Chomsky que genere el siguiente lenguaje:

$$L = \{ucv : u, v \in \{0,1\} \text{ y n}^{\text{o}} \text{ de subcadenas 01 en } u \text{ es igual al n}^{\text{o}} \text{ de subcadenas 10 en } v\}$$

Comprueba con el algoritmo **CYK** si la cadena *010c101* pertenece al lenguaje generado por la gramática.

1. Gramática libre del contexto que genera el lenguaje:

$$S \to S_1AS_1|S_1BS_0|S_0CS_1|S_0DS_0|YXY$$
$$Y \to S_1|S_0|\varepsilon$$
$$S_0 \to 0|0S_0$$
$$S_1 \to 1|1S_1$$
$$A \to 1A|A1|B0|0C|X|c$$
$$B \to 1B|B0|0D|X|c$$
$$C \to 0C|C1|C0|X|c$$
$$D \to 0D|D0|X|c$$
$$X \to 01A10$$

   Donde $A$ siempre tendrá 1 a ambos lados; $B$, 1 a la izquierda y 0 a la derecha; $C$, 0 a la izquierda y 1 a la derecha; y por último $D$, 0 a ambos lados.

   $S$ y sus transiciones son introducidos (en vez de empezar directamente en $A$) para que tanto $u$ como $v$ no sean vacías.

   $Y$ ha sido incluida para aquellas palabras que del tipo $0001c10$, donde $v$ (lo mismo para $u$ con otras palabras) no está vacía pero no incluye ningún elemento además de la subcadena 10. Estas palabras no están contempladas si no incluimos este estado.

2. Gramática en forma normal de Chomsky que genere el lenguaje:

$$S \to S_1E_2|S_0E_3|YX|XY|S_0X_0$$
$$E_2 \to AS_1|BS_0$$
$$E_3 \to CS_1|DS_0$$
$$S_0 \to 0|S_0S_0$$

$$S_1 \to 1|S_1S_1$$
$$A \to S_1A|AS_1|BS_0|S_0C|S_0X_0|c$$
$$B \to S_1B|BS_0|S_0D|S_0X_0|c$$
$$C \to S_0C|CS_1|DS_0|S_0X_0|c$$
$$D \to S_0D|DS_0|S_0X_0|c$$
$$X \to S_0X_0$$
$$X_0 \to S_1X_1$$
$$X_1 \to AX_2$$
$$X_2 \to S_1S_0$$

3. Apliquemos ahora **CYK** para la cadena *010c010*:

Tabla 1: **CYK** - ejercicio 13

| 0 | 1 | 0 | c | 0 | 1 | 0 |
|---|---|---|---|---|---|---|
| $S_0$ | $S_1$ | $S_0$ | $A, B, C, D$ | $S_0$ | $S_1$ | $S_0$ |
| | $X_2$ | $A, B, C, D$ | $A, B, C, D, E_2, E_3$ | | $X_2$ | |
| | $A, B$ | $S, A, B, C, D, E_2, E_3$ | $A, C, E_2, E_3$ | | | |
| | $S, A, B, E_2$ | $S, A, C, E_2, E_3$ | $X_1$ | | | |
| | $S, A$ | $X_1$ | | | | |
| | $X_0, X_1$ | | | | | |
| $S$ | | | | | | |

**Ejercicio 15.** Encuentra una gramática libre de contexto en forma normal de Chomsky que genere los siguientes lenguaje definidos sobre el alfabeto $\{a, 0, 1\}$ :

$$L_1 = \{auava : u, v \in \{0,1\}^+ \text{ y } u^{-1} = v\}$$
$$L_2 = \{uvu : u, v \in \{0,1\}^+ \text{ y } u^{-1} = v\}$$

Comprueba con el algoritmo **CYK** si la cadena *a0a0a* pertenece a $L_1$ y la cadena *011001* pertenece al lenguaje $L_2$ .

1. Gramática libre del contexto que genera $L_1$:

$$S \to aAa$$
$$A \to XAX|a$$
$$X \to 0|1$$

2. Gramática en forma normal de Chomsky que genera $L_1$:

$$S \to AE_1$$
$$E_1 \to BA$$
$$A \to a$$
$$B \to XE_2|a$$
$$E_2 \to BX$$
$$X \to 0|1$$

Tabla 2: **CYK** - ejercicio 15

| a | 0 | a | 0 | a |
|---|---|---|---|---|
| $A, B$ | $X$ | $A, B$ | $X$ | $A, B$ |
| $E_2$ | | $E_2$ | | |
| | $B$ | | | |
| | $E_1$ | | | |
| $S$ | | | | |

3. Apliquemos ahora **CYK** para la cadena *a0a0a*:

4. $L_2$ no es libre del contexto. Lo demostraremos utilizando el lema de bombeo y la cadena $f = 0^n 1^n 1^n 0^n 0^n 1^n$ (obviamente del lenguaje tomando $u = 0^n 1^n$) Reescribamos el lenguaje para que la prueba sea más sencilla:

$$L_2 = \{u_1 v u_2 : u_1, u_2, v \in \{0,1\}^+,\ u_1^{-1} = v\ ,\ u_1 = u_2\}$$

Aplicando el lema de bombeo, sea una descomposición de $f$ de la forma $f = abcde$ donde $|bcd| \leq n$, $|b| \geq 1$ y $|d| \geq 1$. Llamemos $x = bcd$.

*a*) Si $x$ unicamente toma elementos de $u_1$, al bombear no se cumple ni que $v = u_1$ ni que $u_1 = u_2$.

*b*) Si $x$ toma elementos tanto de $u_1$ como de $v$, entonces $x = 1^k, k \in \{1, ..., n\}$. Bombeando unicamente unos, $u_1 \neq u_2$.

*c*) Si $x$ toma elementos tanto de $v$ como de $u_2$, entonces $x = 0^k, k \in \{1, ..., n\}$. Bombeando unicamente ceros, $u_1 \neq u_2$.

*d*) Finalmente, si $x$ unicamente toma elementos de $u_2$, al bombear no se cumple ni que $v = u_1$ ni que $u_1 = u_2$.

Por tanto el lenguaje no es libre del contexto. Es obvio entonces que no podemos encontrar una gramática en forma normal de Chomsky que lo genere. Como el algoritmo **CYK** unicamente se aplica a gramáticas de esta forma, no puedo aplicarlo. Aún así es obvio que la cadena *011001* pertenece al lenguaje (de hecho es $f$ para $n = 1$).

**Ejercicio 21.** Si $L_1$ y $L_2$ son lenguajes sobre el alfabeto $A$, entonces se define el cociente:

$$L_1/L_2 = \{u \in A^* : \exists w \in L_2 \text{ tal que } uw \in L_1\}$$

Demostrar que si $L_1$ es independiente del contexto y $L_2$ regular, entonces $L_1/L_2$ es independiente del contexto.

Para probar que $L_1/L_2$ es independiente del contexto encontraremos $A_3$, un autómata con pila por el criterio de estados finales que lo genere. Sea $A_1 = (Q_1, A, B, \delta_1, q_0^1, Z_0, F_1)$ un autómata con pila por el criterio de estamos finales que acepta $L_1$ y $A_2 = (Q_2, A.\delta_2, q_0^2, F_2)$ un autómata determinista que acepta $L_2$, entonces:

$$A_3 = (Q_1 \times Q_2, A, B, \delta_3, (q_0^3 = (q_0^1, q_0^2)), Z_0, F_3)$$

Definamos ahora el conjunto $F_3$ y la función de transición $\delta_3$:

$$(q_i, q_j) \in F_3 \leftrightarrow q_i \in F_1 \text{ y } q_j \in F_2$$

Si $L_2 \neq \{0,1\}^*$ existirá un estado $q_x \in Q_2$ a partir del cual no se puede alcanzar ningún estado final. Llamaremos a este estado el estado de error. Sea $a \in A$, $X \in B$ y $C = B^* \cup \{\varepsilon\}$. Se define la función de transición como:

$$\delta_3^*((q_i,q_0),X,a) = \{(q_k,q_0) : (q_k,Y) \in \delta_1(q_i,X,a), Y \in C\}$$

$$\delta_3^*((q_i,q_j),X,\varepsilon) = \{(q_k,q_z) : \exists b \in A \text{ tal que } (q_k,Y) \in \delta_1^*(q_i,X,b), Y \in C, q_z \in \delta_2^*(q_j,b), q_z \neq q_x\}$$

La idea intuitiva reside en recorrer los dos autómatas al mismo tiempo. En primer lugar consumiremos la palabra $u$ moviéndonos unicamente por el autómata $A_1$. Cuando lleguemos a dicho estado, para comprobar si es posible leer una palabra de $L_2$ llegando a un estado final de $A_1$ utilizamos las transiciones del segundo tipo, donde consumimos una cadena vacía y avanzamos de igual foma por $A_1$ que por $A_2$.

Como $A_3$ acepta $L_1/L_2$, es independiente del contexto.

**Ejercicio 22.** Si $L$ es un lenguaje sobre $\{0,1\}$, sea $SUF(L)$ el conjunto de los sufijos de palabras de $L$ :

$$SUF(L) = \{u \in \{0,1\}^* : \exists v \in \{0,1\}^*, \text{ tal que } vu \in L\}$$

Demostrar que si $L$ es independiente del contexto, entonces $SUF(L)$ también es independiente del contexto.

Aunque podría utilizar la misma solución que en el ejercicio anterior cambiando un par de detalles en la definición del autómata lo resolveré de otra forma más interesante.
Sea $L$ un lenguaje sobre un alfabeto $A$. Defino el *lenguaje inverso de* $L$ como:

$$INV(L) = \{u \in A : u^{-1} \in L\}$$

Veamos que si $L$ es independiente del contexto, $INV(L)$ también lo es. Utilizando la misma idea que al pasar de un autómata con pila a una gramática independiente del contexto con una ligera variación, a partir de las transiciones de la forma:

$$\delta(q_i,X_0,a) = (q_j,X_1X_2...X_N)$$

Obtenemos producciones de la forma:

$$[q_j,X_0,q_i] \to a[q_j,X_1,p_1][p_1,X_2,p_2]...[p_{N-2},X_{N-1},p_{N-1}][p_{N-1},X_N,q_i], p_i \in Q, \forall i \in \{1,...,N-1\}$$

Intuitivamente, invirtiendo el autómata como haríamos en lenguajes regulares. Hemos obtenido una gramática que genera $INV(L)$ a partir de un automáta que genera $L$. Por lo tanto $INV(L)$ es independiente del contexto si y solo si $L$ lo es.

Es sencillo darse cuenta, utilizando la definición del ejercicio anterior de que $L/\{0,1\}^* = PREF(L)$ es el conjunto de los prefijos de las palabras de $L$, definido análogamente al de los sufijos. También es sencillo comprobar que $INV(SUF(L)) = PREF(L)$.

Sabemos por el ejercicio anterior que $PREF(L)$ es independiente del contexto. Como es el inverso de $SUF(L)$, $SUF(L)$ también es independiente del contexto.