

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Facultad de Ciencias

Autor: Adrián Aguilera Moreno



Autómatas y Lenguajes Formales

Tarea 2

1. Demuestra que el lenguaje

$$\{\mathbf{a} \in \{a, b, c\}^* \mid \text{la longitud de } \mathbf{a} \text{ es el cuadrado de un natural.}\}$$

no es regular. Usa tanto el teorema del bombeo como el de Myhill-Nerode.

Demostración: Antes de iniciar con las demostraciones por los teoremas requeridos, analicemos la estructura de las cadenas generadas por

$$L(A) = \{\mathbf{a} \in \{a, b, c\}^* \mid \text{la longitud de } \mathbf{a} \text{ es el cuadrado de un natural.}\}$$

Suponiendo que existe el autómata $A \in DFA$ tal que reconoce el lenguaje anterior.

Nótese que $\Sigma = \{a, b, c\}$ y $L(A) = \Sigma^*$, entonces \mathbf{a} debe ser una cadena finita y supondremos que $L(A)$ es un lenguaje “muy grande” pero finito.

Además, para nuestras demostraciones nos fijaremos en la longitud de las cadenas generadas por nuestro lenguaje y no necesariamente en la estructura de estas cadenas, entonces

$$|abababcbacbac| = |aaaaabbbbbbccc| = |a^5b^6c^3|$$

Lo anterior se puede generalizar usando inducción. Por tanto, en las demostraciones supondré, sin pérdida de generalidad, que

$$\mathbf{a} = a^i b^j c^r$$

para $i, j, r \in \mathbb{N}$.

Obs. El que \mathbf{a} tenga la estructura anterior no indica que b siempre debe estar en esa posición, incluso podríamos prescindir de b o de algún término en la cadena, al suponer esa estructura solo simplificamos el considerar casos (que por ser un conjunto infinito o muy grande, tendríamos muchas combinaciones para concatenar elementos del alfabeto).

Para este ejercicio analicemos 2 versiones en la demostración, *i.e.*,

- **Demostración por Teorema del Bombeo.** Procedamos por reducción al absurdo, supongamos que el lenguaje dado es regular y que $A = \langle Q_A, \Sigma, \delta_A, s, \mathbb{F}_A \rangle$ es un autómata determinista finito [en caso de no serlo, siempre se puede reducir a un DFA] con k estados y $L(A)$ como se definió previamente.

Sea $\mathbf{a} = a^i b^j c^r$ una cadena aceptada por nuestro autómata con

$$i + j + r > k \text{ y al menos } j > k.$$

de lo anterior, podemos deducir que durante la lectura de la cadena \mathbf{a} , al menos en 2 ocasiones se pasa por algún estado q en A (es claro que q debe ser estado de A). Así, para la función δ_A de transición en A , tenemos que

$$\begin{aligned} \delta_A^*(s, a^i b^m) &= q \\ \delta_A^*(q, b^n) &= q \\ \delta_A^*(q, b^l c^r) &= f \in \mathbb{F}_A. \end{aligned}$$

$\forall_{m,n,l} \in \mathbb{N}$. Donde $m + n + l = j$ y en particular $n \neq 0$ [b^n no se colapsa a λ].

Como A es determinista, tenemos que para todo $p \in \mathbb{N}$ se cumple que

$$\delta_A^*(q, b^{np}) = q$$

como consecuencia, directa se tiene que

$$\delta_A^*(s, a^i b^m b^{np} b^l c^r) = f \in \mathbb{F}_A.$$

pero $m + np + l \neq j$ cuando $p \neq 1$. Por tanto $a^i b^m b^{np} b^l c^r \in L(A)$, lo que implica

$$i + m + np + l + r = t^2$$

con $t \in \mathbb{N}$. Sin embargo,

$$i + m + np + l + r = t^2 \not\Rightarrow i + m + n(p + 1) + l + r = t^2 + n = x^2 !!!$$

Contraejemplo: Con $i = 6, m = 3, n = 4, l = 2, r = 1$.

$$16 = 6 + 3 + 4 + 2 + 1 \quad \text{Con } p = 1.$$

$$20 = 6 + 3 + 4(2) + 2 + 1 = 6 + 3 + 8 + 2 + 1 \quad \text{Con } p = 2.$$

Claramente 20 no es el cuadrado de un natural.

lo cual es una contradicción, pues esto debe funcionar para cualquier p natural.

$\therefore \{a \in \{a, b, c\}^* \mid \text{la longitud de } a \text{ es el cuadrado de un natural.}\}$ **NO** es regular.

- **Demostración por Teorema de Myhill-Nerode.** Por refutación, supongamos que L es regular, entonces tenemos que si

$$\alpha = (a^i b^j c^r)^{w^2} \text{ tal que } (i + j + r)^{w^2} = x^{w^2} \Rightarrow \alpha \in L.$$

$$\beta = (a^m b^n c^k)^{v^2} \text{ tal que } (m + n + k)^{v^2} = (y)^{v^2} \Rightarrow \beta \in L.$$

$$\varphi = \alpha.$$

Donde $v \neq w$. Así,

$$\begin{aligned} \alpha\varphi &= (a^i b^j c^r)^{w^2} \cdot (a^i b^j c^r)^{w^2} \\ &= (a^i b^j c^r)^{w^2 + w^2} && \text{Aplicando el supuesto del inicio.} \\ \Rightarrow |\alpha\varphi| &= (i + j + r)^{w^2 + w^2} \\ &= (i + j + r)^{2w^2} \\ &= \left((i + j + r)^{w^2}\right)^2 \end{aligned}$$

Luego, por el Teorema de Myhill-Nerode tenemos que

$$\alpha \equiv_L \beta \Leftrightarrow \forall_{n \in \mathbb{N}} \alpha\varphi^{n^2} \in L \Leftrightarrow \beta\varphi^{n^2} \in L$$

Ahora, verifiquemos que $\beta\varphi \in L$, i.e.,

$$\begin{aligned} \beta\varphi &= (a^m b^n c^k)^{v^2} \cdot (a^i b^j c^r)^{w^2} \\ \Rightarrow &= (y)^{v^2} \cdot (x)^{w^2} \end{aligned}$$

– Si $v > w$, entonces

$$\begin{aligned} y^{v^2 - w^2 + w^2} \cdot x^{w^2} &= y^{v^2 - w^2} \cdot (xy)^{w^2 + w^2} \\ &= y^{v^2 - w^2} \cdot \left((xy)^{w^2}\right)^2 \end{aligned}$$

sabemos que $v^2 - w^2 = z^2$ solo para $\{v, w, z\}$ una terna pitagórica, y no es cierto para cualesquiera naturales, esto contradice que $\beta\varphi \in L$, sin embargo nos falta analizar otro caso.

– Si $w > v$, entonces

$$y^{w^2-v^2+v^2} \cdot x^{w^2} = y^{w^2-v^2} \cdot (xy)^{v^2+w^2}$$

como podemos observar, sucede lo mismo que en el caso anterior, pues $w^2 - v^2 = z^2$ y $v^2 + w^2 = t^2$ solo se cumplen para ternas pitagóricas específicas y no para cualesquiera naturales.

De lo anterior, tenemos que $\alpha \not\equiv_L \beta$ (pues no hay afinación en L), lo que implica que existe una cantidad infinita de clases de equivalencias, esto contradice el punto 3 del Teorema de Myhill-Nerode (el tener índice finito), luego

$\therefore \{a \in \{a, b, c\}^* \mid \text{la longitud de } a \text{ es el cuadrado de un natural}\}$ **NO** es regular.

En todos los casos se obvia que la demostración para $a = \lambda$ [cadena vacía] es por vacuidad.

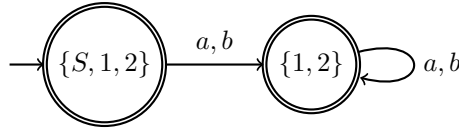
De los analisis anteriores (tanto por Teo. del Bombeo como del Teo. de Myhill-Nerode) concluimos que

$\therefore \{a \in \{a, b, c\}^* \mid \text{la longitud de } a \text{ es el cuadrado de un natural}\}$ **NO** es regular.

□

2. Minimiza el autómata de tu respuesta al ejercicio 2 de la tarea 1.

▽ **Solución:** El autómata del ejercicio 2 en la tarea 1 se presenta a continuación:



Al que llamaremos A . Para la simplicidad de la solución consideremos

$$\mathbf{S} = \{S, 1, 2\}$$

$$\mathbf{q} = \{1, 2\}$$

así, como \mathbf{S} y \mathbf{q} son estados finales, tenemos que la matriz que les representa se ve como la que se muestra a continuación:

$$\begin{matrix} & \mathbf{S} & \mathbf{q} \\ \mathbf{S} & \begin{pmatrix} n & n \end{pmatrix} \\ \mathbf{q} & \begin{pmatrix} n & n \end{pmatrix} \end{matrix}$$

Obs. Podemos omitir analizar la diagonal y la parte (triangular) inferior de la matriz, pues nuestra matriz es simétrica, y por el algoritmo de Minimalización, esto se cumplirá al inicio de las iteraciones y en cuanto termine.

Ahora analicemos la (2,2)-posición en la matriz con respecto al algoritmo de Minimalización, esto es

$$\begin{aligned} (s, q) &= n. \\ (\delta_A(\mathbf{S}, a), \delta_A(\mathbf{q}, a)) &= (\mathbf{q}, \mathbf{q}) \\ &= n. \\ (\delta_A(\mathbf{S}, b), \delta_A(\mathbf{q}, b)) &= (\mathbf{q}, \mathbf{q}) \\ &= n. \end{aligned}$$

por el algoritmo de Minimalización tenemos que si $(q_i, q_j) = n$, entonces $q_i \approx q_j$. Por tanto $\mathbf{S} \approx \mathbf{q}$ [esto es, que formen parte de la misma clase de equivalencia].

Así, el autómata cociente de A es el que a continuación se presenta:



<

3. Da una *CFG* para el lenguaje

$$\{a^n b^{2n} c^k \mid 1 \leq k, n\}.$$

▽ **Solución:**

Sea $G = \langle \{a, b, c\}, \{S, X, Y, Z\}, S, \rightarrow_G \rangle$, donde las reglas de producción están dadas por

$$\begin{aligned} S &\rightarrow_G XYYZ \\ X &\rightarrow_G a \mid aXY Y \\ Y &\rightarrow_G b \\ Z &\rightarrow_G c \mid ZZ \end{aligned}$$

Para mayor practicidad en los ejercicios siguientes, definimos

$$\begin{aligned} \Sigma_G &= \{a, b, c\} \\ \Gamma_G &= \{S, X, Y, Z\} \end{aligned}$$

de lo anterior, se muestra un ejemplo de derivación con base a la gramática propuesta

$$\begin{aligned} S &\Rightarrow_G XYYZ \\ &\Rightarrow_G (aXY Y)YYZ \\ &\Rightarrow_G a(a)YYYYZ \\ &\Rightarrow_G aa(b)YYYZ \\ &\Rightarrow_G aab(b)YYZ \\ &\Rightarrow_G aabb(b)YZ \\ &\Rightarrow_G aabbb(b)Z \\ &\Rightarrow_G aabbbb(ZZ) \\ &\Rightarrow_G aabbbb(ZZZ) \\ &\Rightarrow_G aabbbb(c)ZZ \\ &\Rightarrow_G aabbbbc(c)Z \\ &\Rightarrow_G aabbbbcc(c). \end{aligned}$$

Los paréntesis solo son para ayudar a identificar la sustitución en cada iteración.

<

4. Transforma la gramática de 3 a forma normal de Chomsky.

▽ **Solución:** Sea G' la gramática a construir. Aplicando la regla 1 del algoritmo de transformación a CNF, tenemos que

$$\begin{aligned} A_a &\rightarrow a \\ A_b &\rightarrow b \\ A_c &\rightarrow c \end{aligned}$$

Así, las reglas de producción se ven modificadas

$$\begin{aligned}
 S &\rightarrow XYYZ \\
 X &\rightarrow a \mid A_aXY \\
 Y &\rightarrow b \\
 Z &\rightarrow c \mid ZZ \\
 A_a &\rightarrow a \\
 A_b &\rightarrow b \\
 A_c &\rightarrow c
 \end{aligned}$$

Ahora tenemos símbolos no terminales, pero no cumplen ser CNF, los únicos que de momento cumplen estar en CNF son los símbolos del alfabeto. Luego, por la regla 3 del algoritmo de transformación a CNF, tenemos que

$$\begin{aligned}
 C_1 &\rightarrow XY \\
 C_2 &\rightarrow YZ \\
 C_3 &\rightarrow A_aX \\
 C_4 &\rightarrow YY
 \end{aligned}$$

Ejecutando 4 y 5 en el algoritmo de transformación a CNF, obtenemos

$$\begin{aligned}
 S &\rightarrow_{G'} C_1C_2 \\
 X &\rightarrow_{G'} a \mid C_3C_4 \\
 Y &\rightarrow_{G'} b \\
 Z &\rightarrow_{G'} c \mid ZZ \\
 A_a &\rightarrow_{G'} a \\
 A_b &\rightarrow_{G'} b \\
 A_c &\rightarrow_{G'} c \\
 C_1 &\rightarrow_{G'} XY \\
 C_2 &\rightarrow_{G'} YZ \\
 C_3 &\rightarrow_{G'} A_aX \\
 C_4 &\rightarrow_{G'} YY
 \end{aligned}$$

De lo anterior, es claro que G' está en CNF. ◁

5. Transforma la gramática de 3 a forma normal de Greibach (basta con que presentes algunas reglas de producción representativas).

▽ **Solución:** ◁