Pontificia Universidad Católica de Chile
Departamento de Ciencia de la Computación
IIC2223 – Teoría de Autómatas y Lenguajes Formales
Segundo semestre de 2024
Profesor: Cristian Riveros
Ayudante: Amaranta Salas

Ayudantia 4

Autómatas con ϵ-transiciones y Teorema de Kleene

Problema 1

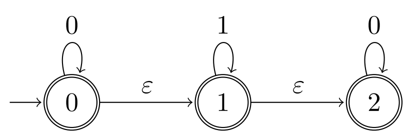
Sea $\Sigma = \{0, 1\}^*$. Construya expresiones regulares y ε-NFAs para los siguientes lenguajes:

- $L = \{w = a_1 a_2 \dots a_n \in \Sigma^* \mid \exists i, j.\, 0 \leq i \leq j \leq n : a_1 \dots a_i = 0 \dots 0 \wedge a_{i+1} \dots a_j = 1 \dots 1 \wedge a_{j+1} \dots a_n = 0 \dots 0\}$
- $L = \{w = a_1 a_2 \dots a_n \in \Sigma^* \mid \forall i \text{ impar} : a_i a_{i+1} = 01 \vee a_i a_{i+1} = 10\}$

Solución

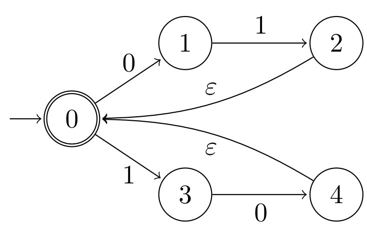
- $L = \{w = a_1 a_2 \dots a_n \in \Sigma^* \mid \exists i, j.\, 0 \leq i \leq j \leq n : a_1 \dots a_i = 0 \dots 0 \wedge a_{i+1} \dots a_j = 1 \dots 1 \wedge a_{j+1} \dots a_n = 0 \dots 0\}$

$$R = 0^* 1^* 0^*$$



- $L = \{w = a_1 a_2 \dots a_n \in \Sigma^* \mid \forall i \text{ impar} : a_i a_{i+1} = 01 \vee a_i a_{i+1} = 10\}$

$$R = (01 + 10)^*$$



Problema 2

Sea Σ un alfabeto finito y R una expresión regular sobre Σ^*. Se define el operador:

$$\mathcal{L}(R^{\downarrow\downarrow}) = \{u_1 u_2 \dots u_k \mid \exists k \geq 1. \exists w' \in \mathcal{L}(R).\, w' = u_1 v_1 u_2 v_2 \dots u_k v_k \wedge u_i, v_i \in \Sigma^*\}$$

Demuestre que si $\mathcal{L}(R)$ es regular, entonces $\mathcal{L}(R^{\downarrow\downarrow})$ es regular.

Solución

Como $\mathcal{R}$ es una expresión regular, existe un DFA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ tal que $\mathcal{L}(A) = \mathcal{L}(\mathcal{R})$. Además, $\mathcal{L}(\mathcal{R}^{\downarrow\downarrow})$ corresponde al lenguaje de las palabras que son una subpalabra (no necesariamente contigua o distinta de ϵ) de alguna palabra de $\mathcal{R}$.
Ahora, podemos comenzar desde el autómata $\mathcal{A}$ y modificarlo de forma tal que "salte" algunos segmentos de las palabras al momento de leerlas, advinando de esta forma las partes que se deben añadir a w para obtener palabras en el lenguaje original $\mathcal{L}(\mathcal{R})$.
Construiremos el siguiente ε-NFA:

$$\mathcal{A}' \;=\; (Q, \Sigma, \Delta, \{q_0\}, F)$$

La relación de transición Δ será como sigue:

- Si $\delta(p, a) = q$, con $p, q \in Q, a \in \Sigma$, entonces $(p, a, q) \in \Delta$ (replicamos las transiciones originales de $\mathcal{A}$).
- Si existe una palabra $v \in \Sigma^*$ tal que al ser leída desde el estado p se llega al estado q (i.e $\hat{\delta}(p, v) = q$), entonces podemos saltar directamente al estado q, es decir $(p, \epsilon, q) \in \Delta$.

Ahora debemos demostrar que $\mathcal{L}(\mathcal{A}') = \mathcal{L}(R^{\downarrow\downarrow})$.

$\mathcal{L}(\mathcal{R}^{\downarrow\downarrow}) \subseteq \mathcal{L}(\mathcal{A}')$

Sea $w = u_1 u_2 \cdots u_k$ tal que existe una palabra $w' = u_1 v_1 \cdots u_k v_k$ en $\mathcal{L}(\mathcal{R})$. Como $w' \in \mathcal{L}(\mathcal{R})$ entonces es aceptada por el autómata $\mathcal{A}$ y existe una ejecución ρ de $\mathcal{A}$ sobre w' que tiene la siguiente forma:

$$\rho :\; q_0 \overset{u_1}{\to} p_0 \overset{v_1}{\to} q_1 \overset{u_2}{\to} p_1 \overset{v_2}{\to} q_2 \overset{u_3}{\to} p_2 \cdots \overset{v_k}{\to} q_k$$

donde q_k es un estado final.
Ahora notemos en esta ejecución que $\hat{\delta}(q_{i-1}, u_i) = p_{i-1}$ y $\hat{\delta}(p_{i-1}, v_i) = q_i$ y, por como definimos el $\mathcal{A}'$, deducimos que una ejecución de $\mathcal{A}'$ sobre la palabra w será:

$$\rho' :\; q_0 \overset{u_1}{\to} p_0 \overset{\epsilon}{\to} q_1 \overset{u_2}{\to} p_1 \overset{\epsilon}{\to} q_2 \overset{u_3}{\to} p_2 \cdots \overset{\epsilon}{\to} q_k$$

Como el estado q_k también es final en $\mathcal{A}'$, ρ' será una ejecución de aceptación para w en $\mathcal{A}'$ y por tanto $\mathcal{L}(\mathcal{R}^{\downarrow\downarrow}) \subseteq \mathcal{L}(\mathcal{A}')$.

$\mathcal{L}(\mathcal{A}') \subseteq \mathcal{L}(\mathcal{R}^{\downarrow\downarrow})$

Sea $w \in \mathcal{L}(\mathcal{A}')$. Podemos demostrar que existe una ejecución de aceptación, ρ, de $\mathcal{A}'$ sobre la palabra w tal que ρ no tiene dos ϵ-transiciones seguidas.
En caso de existir un fragmento de ρ de la forma $q_i \overset{\epsilon}{\to} p_i \overset{\epsilon}{\to} q_{i+1}$, es claro que $\hat{\delta}(q_i, \epsilon) = q_{i+1}$ y, por la forma en que se añadieron ϵ-transiciones a Δ, existirá la transición $(q_i, \epsilon, q_{i+1}) \in \Delta$, y por tanto la ejecución que es idéntica a ρ, pero que realiza directamente $q_i \overset{\epsilon}{\to} q_{i+1}$.
Por lo anterior, podemos considerar que ρ tiene la siguiente forma:

$$\rho :\; q_0 \overset{u_1}{\to} p_0 \overset{\epsilon}{\to} q_1 \overset{u_2}{\to} p_1 \overset{\epsilon}{\to} q_2 \overset{u_3}{\to} p_2 \cdots \overset{\epsilon}{\to} q_k$$

Donde q_k es un estado final, pues lo único que hicimos fue reescribir ρ, pero "separándola" según sus epsilon transiciones.
Luego, mirando la definición de $\mathcal{A}'$, sabemos que para cada ϵ-transición debe haber una palabra v_i tal que $\hat{\delta}(p_{i-1}, v_i) = q_i$, por lo que podemos construir la ejecución ρ' de $\mathcal{A}$ sobre una palabra $w' = u_1 v_1 \cdots u_k v_k$ dada por:

$$\rho' :\; q_0 \overset{u_1}{\to} p_0 \overset{v_1}{\to} q_1 \overset{u_2}{\to} p_1 \overset{v_2}{\to} q_2 \overset{u_3}{\to} p_2 \cdots \overset{v_k}{\to} q_k$$

Donde q_k es un estado final, lo que significa que $w' \in \mathcal{L}(\mathcal{R})$ y por lo tanto $w \in \mathcal{L}(\mathcal{R}^{\downarrow\downarrow})$.

Problema 3

Escriba una expresión regular para el siguiente lenguaje sobre el alfabeto $\{a, b, c\}$:

$$L = \{w \in \{a, b, c\}^* \mid w \text{ no contiene tres } c \text{ seguidas}\}$$

Explique por qué su expresión regular define a L.

Solución

Para resolver la pregunta anterior, vamos a proponer una posible expresión regular que defina a L:

$$(c + cc)^?(a + b + (a + b)c + (a + b)cc)^*$$

En primer lugar, explicaremos la construcción del término de la derecha. La expresión regular no debe definir palabras en las que después de dos letras 'c' haya otra 'c'. Para ello, se incluye la expresión que tiene una 'a' o 'b' y después dos letras 'c' y luego puede haber una 'a' o 'b' o 'ac' o 'bc'. Pero nunca habrán tres letras 'c' seguidas. Finalmente, se añade el término de la izquierda, porque el lenguaje puede incluir la palabra 'c' o 'cc'. Notemos, que la expresión regular define correctamente a L, ya que incluye todas las palabras sobre el alfabeto $\{a, b, c\}$ que no contiene tres letras 'c' seguidas.