IIC2223 – Teoría de Autómatas y Lenguajes Formales – 2' 2024
IIC2224 – Autómatas y Compiladores

Pauta Interrogación 1

Preguntas en blanco: Preguntas o items entregados en blanco se evaluarán con 0.5 de 6 puntos.

Pregunta 1

Sean L_1 y L_2 lenguajes regulares. Demuestre que $L_1 \cdot L_2$ es un lenguaje regular.

Solución

Sean $A_1 = (Q_1, \Sigma, \Delta, I_1, F_1)$ y $A_2 = (Q_2, \Sigma, \Delta, I_2, F_2)$ dos NFAs tal que $L_1 = \mathcal{L}(A_1),\ L_2 = \mathcal{L}(A_2)$

Sea el NFA $A = (Q, \Sigma, \Delta, I, F)$, tal que,

$$Q = Q_1 \uplus Q_2$$
$$I = I_1$$
$$F = F_2$$
$$\Delta = \Delta_1 \uplus \Delta_2 \uplus \{(p, \varepsilon, q).\ p \in F_1, q \in I_2\}$$

Por demostrar. $L_1 \cdot L_2 = \mathcal{L}(A)$, es decir: $w \in L_1 \cdot L_2 \Leftrightarrow w \in \mathcal{L}(A)$

- $w \in L_1 \cdot L_2 \Rightarrow w \in \mathcal{L}(A)$
  Sea $w = w_1 w_2 \in L_1 \cdot L_2$ tal que $w_1 \in L_1, w_2 \in L_2$. Como $\mathcal{L}(A_1) = L_1$ y $\mathcal{L}(A_2) = L_2$, se tiene, necesariamente, que:

  $$(q_0^1, w_1) \vdash^*_{\Delta_1} (q_f^1, \varepsilon) \wedge (q_0^2, w_2) \vdash^*_{\Delta_2} (q_f^2, \varepsilon)$$

  Con $q_0^1 \in I_1, q_f^1 \in F_1, q_0^2 \in I_2, q_f^2 \in F_2$. Por construcción y dado que $(q_f^1, \epsilon, q_0^2) \in \Delta$, se tiene que:

  $$(q_0^1, w_1 w_2) \vdash^*_{\Delta} (q_f^1, w_2) \vdash_{\Delta} (q_0^2, w_2) \vdash^*_{\Delta} (q_f^2, \varepsilon)$$

  Así, se concluye que $(q_0^1, w_1 w_2) \vdash^*_{\Delta} (q_f^2, \varepsilon) \Rightarrow w_1 w_2 \in \mathcal{L}(A)$, con $q_0^1 \in I, q_f^2 \in F$.

- $w \in \mathcal{L}(A) \Rightarrow w \in L_1 \cdot L_2$
  Sea $w \in \mathcal{L}(A)$, se tiene que:

  $$(q_0, w) \vdash^*_{\Delta} (q_f, \varepsilon)$$

  Con, $q_0 \in I = I_1, q_f \in F = F_2$. Por construcción, se tiene que en algún momento se debe saltar de Q_1 a Q_2, esto es,

  $$(q_0, w_1 w_2) \vdash^*_{\Delta} (q_1, w_2) \vdash_{\Delta} (q_2, w_2) \vdash^*_{\Delta} (q_f, \varepsilon)$$

  Por la existencia de la ε-transición, $q_1 \in F_1, q_2 \in I_2$. Como $q_0, q_1 \in Q_1$ y $q_2, q_f \in Q_2$, se tiene que:

  $$(q_0, w_1) \vdash^*_{\Delta_1} (q_1, \varepsilon)\ \wedge\ (q_2, w_2) \vdash^*_{\Delta_2} (q_f, \varepsilon)$$

  Por lo tanto, $w_1 \in L_1, w_2 \in L_2$.

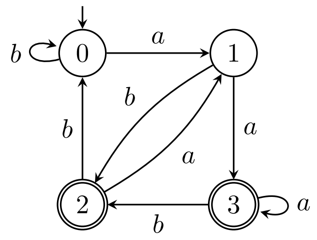
Distribución de puntajes

Los 6 puntos de la pregunta se encuentran distribuidos de la siguiente forma:

- **(1 pto.)** Por definir los estados del autómata A, -0.5 si no considera la unión disjunta.
- **(1 pto.)** Por definir los conjuntos de estados iniciales I y finales F del autómata A.
- **(1 pto.)** Por definir el conjunto Δ de transiciones del autómata A, -0.5 si no considera la unión disjunta.
- **(1.5 ptos)** Por la demostración hacia el lado derecho de la doble implicancia.
  - **(0.5 ptos)** Por el planteamiento y escribir las transiciones de estados.
  - **(0.5 ptos)** Por utilizar la construcción del autómata A.
  - **(0.5 ptos)** Por lograr concluir la demostración utilizando lo anterior.
- **(1.5 ptos)** Por la demostración hacia el lado izquierdo de la doble implicancia. (una buena argumentación de por qué los pasos anteriores son reversibles también se considerará correcta).
  - **(0.5 ptos)** Por el planteamiento y escribir las transiciones de estados.
  - **(0.5 ptos)** Por utilizar la construcción del autómata A.
  - **(0.5 ptos)** Por lograr concluir la demostración utilizando lo anterior.

Pregunta 2

Considere el siguiente autómata finito determinista $\mathcal{A}$ sobre el alfabeto $\Sigma = \{a, b\}$.



1. Minimice el número de estados de $\mathcal{A}$ utilizando el algoritmo de minimización de estados.

2. Construya una expresión regular equivalente a $\mathcal{L}(\mathcal{A})$. Justifique su respuesta.

Solución

1. Como se solicita en el enunciado, seguiremos el algoritmo de minimización de estados. Realizamos los siguientes pasos:

   (a) Paso 1 del algoritmo: Construimos la tabla con los pares $\{p, q\}$ sin marcar (dejamos con negrita los estados finales, es decir, 2 y 3):

	0	1	**2**	**3**
0	-			
1	-	-		
2	-	-	-	
3	-	-	-	-

   (b) Paso 2 del algoritmo: Marcamos $\{p, q\}$ si $p \in F$ y $q \notin F$ o viceversa:
   Debido a lo anterior, tenemos que marcar los siguientes pares $\{0, 2\}$, $\{0, 3\}$, $\{1, 2\}$, $\{1, 3\}$. Obtenemos la siguiente tabla:

	0	1	**2**	**3**
0	-		✓	✓
1	-	-	✓	✓
2	-	-	-	
3	-	-	-	-

   (c) Paso 3 del algoritmo (Lo repetimos hasta que no hayan cambios): Si $\{p, q\}$ no está marcado y $\{\delta(p, a), \delta(q, a)\}$ está marcado para algún $a \in \Sigma$, entonces marcamos $\{p, q\}$.
   Siguiendo lo anterior, notemos que debemos marcar los siguientes pares:

   - El par $\{0, 1\}$ debido a que $\{\delta(0, a), \delta(1, a)\} \to \{1, 3\}$, que resulta en un par que está marcado. Notemos que si seguimos $\{\delta(0, b), \delta(1, b)\} \to \{0, 2\}$ que también está marcado. Por lo que marcamos el par $\{0, 1\}$.
   - El par $\{2, 3\}$ debido a que $\{\delta(2, a), \delta(3, a)\} \to \{1, 3\}$, que que resulta en un par que está marcado. Notemos que si seguimos $\{\delta(2, b), \delta(3, b)\} \to \{0, 2\}$ que también está marcado. Por lo que marcamos el par $\{2, 3\}$.
     Obtenemos la siguiente tabla:

	0	1	**2**	**3**
0	-	✓	✓	✓
1	-	-	✓	✓
2	-	-	-	✓
3	-	-	-	-

   Notemos que la tabla ya esta completa, por lo que no es necesario repetir el paso 3. Ahora, notemos que todos los pares de estados $\{p, q\}$ de la tabla están marcados, por lo que todos los pares de estados son distinguibles. Debido a lo anterior podemos concluir que el autómata ya se encuentra minimizado, es decir, el resultado del algoritmo es el autómata inicial.

Distribución de puntajes

Los 6 puntos de la pregunta se encuentran distribuidos de la siguiente forma:

- **(1 pto.)** Por completar la tabla o explicar el proceso del llenado de esta siguiendo el algoritmo.
- **(1 pto.)** Por marcar los pares de estados $\{p, q\}$ si $p \in F$ y $q \notin F$ o viceversa, es decir, por realizar correctamente el paso 1.
- **(1.5 pto.)** Por marcar el par $\{0, 1\}$ debido a $\{\delta(0, a), \delta(1, a)\} \to \{1, 3\}$ (o siguiendo la otra transición).
- **(1.5 ptos.)** Por marcar el par $\{2, 3\}$ debido a que $\{\delta(2, a), \delta(3, a)\} \to \{1, 3\}$ (o siguiendo la otra transición).
- **(1 pto.)** Por concluir correctamente luego de finalizar el algoritmo.

2. Para esta pregunta, existen diversas formas de llegar a una expresión regular equivalente a $\mathcal{L}(\mathcal{A})$, por lo que desarrollaremos 2 de ellas a continuación.

   (a) La primera estrategia consiste en analizar el autómata:

   - Si vemos el estado 0, podemos notar que su función es recordar que las últimas dos letras son bb o que el largo de la palabra es < 2.
   - Luego, el estado 1 recuerda que las últimas dos letras son ba o a.
   - A continuación, el estado 2 recuerda que las últimas dos letras son ab.
   - Por último, el estado 3 recuerda que las últimas dos letras son aa.

   Si juntamos todo lo anterior podemos concluir que el autómata acepta todas las palabras donde la **penúltima** letra es una a. Eso nos da la siguiente expresión regular:

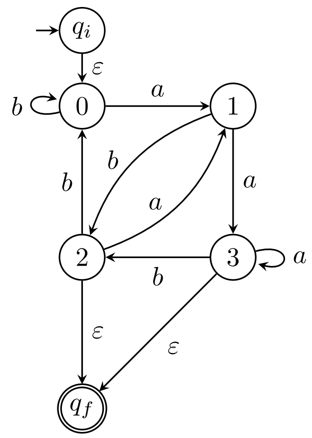
$$(a + b)^* \cdot a \cdot (a + b)$$

   Distribución de puntajes

   Si se siguió esta estrategia, los 6 puntos de la pregunta se encuentran distribuidos de la siguiente forma:

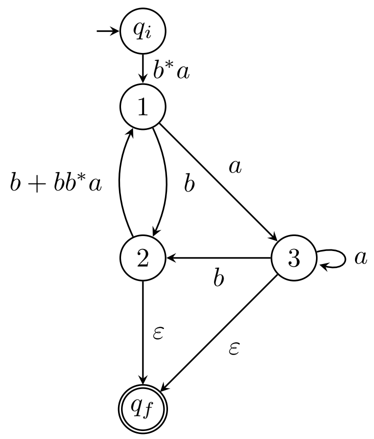
   - **(2 ptos.)** Por explicar que hace cada estado y/o explicar cuales son las palabras que acepta el autómata.
   - **(4 ptos.)** Por construir la expresión regular correctamente relacionándolo con lo anterior.

   (b) La segunda estrategia es utilizando el algoritmo de eliminación de estados:
   Antes de comenzar con el método de eliminación de estados, primero se debe convertir el DFA a un GNFA. Para ello, debemos agregar 2 nuevos estados q_i y q_f. q_i debemos conectarlo al estado 0 y los estados 2 y 3 deben ser conectados a q_f mediante ε-transiciones, resultando en el siguiente GNFA:

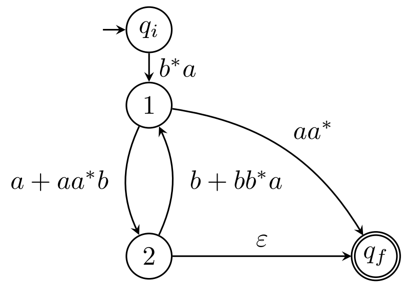


   A continuación, iniciamos el método de eliminación de estados. Por simplicidad, las expresiones regulares se irán simplificando a medida que se avanza en el método, pero no se mostrará explícitamente. Por ejemplo, (εa) se simplificará a (a) y $(\varnothing + a)$ se simplificará a (a).

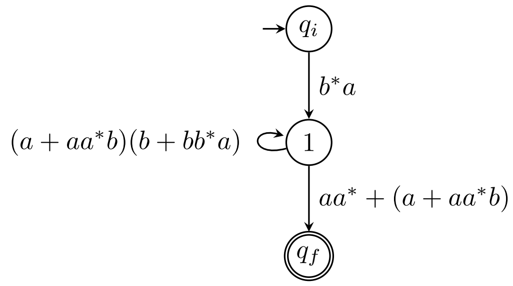
   Paso 1: Eliminar el estado 0.



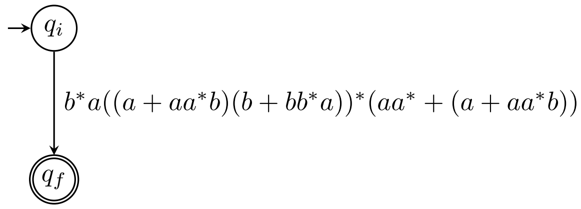
Paso 2: Eliminar el estado 3.



Paso 3: Eliminar el estado 2.



Paso 4: Eliminar el estado 1.



Dado que solo tenemos 2 estados, el inicial y el final, la expresión regular resultante es la expresión correspondiente a la transición entre estos dos estados, es decir, $b^*a((a + aa^*b)(b + bb^*a))^*(aa^* + (a + aa^*b))$.

Distribución de puntajes

Si se siguió esta estrategia, los 6 puntos de la pregunta se encuentran distribuidos de la siguiente forma:

- **(1 pto.)** Por convertir el NFA/DFA a GNFA.
- **(1 pto.)** Por eliminar el estado 3.
- **(1 pto.)** Por eliminar el estado 0.
- **(1 pto.)** Por eliminar el estado 2.
- **(1 pto.)** Por eliminar el estado 1.
- **(1 pto.)** Por concluir con la eliminación de estados y obtener la expresión regular.

Problema 3

Sea $\Sigma = \{a, b, c\}$. Demuestre que los siguientes lenguajes no son regulares:

1. $\text{Prod} = \left\{a^i b^j c^{i \cdot j} \in \Sigma^+ \mid i \geq 0 \land j \geq 0\right\}$
2. $\text{Mod} = \left\{a^i b^j c^{i \bmod j} \in \Sigma^+ \mid i \geq 0 \land j \geq 2\right\}$

Solución

1. Necesitamos demostrar que el lenguaje $\text{Prod} = \{a^i b^j c^{i \cdot j} \in \Sigma^+ \mid i \geq 0 \land j \geq 0\}$ no es regular. Para ello, podemos utilizar el contra positivo del lema de bombeo.
   - Para todo $N > 0$, elegimos la palabra $P = a^N b^N c^{N^2}$.
   - Elegimos x, y, z tal que $P = xyz$, $x = a^N b^N$, $y = c^{N^2}$ $(|y| \geq N)$ y $z = \varepsilon$.
   - Para toda descomposición $y = c^k c^l c^m$ con $k + l + m = N^2$ y $l > 0$.
   - Tenemos nuestra nueva palabra $P' = a^N b^N c^{k+i \cdot l+m}$.
   - Elegimos bombear con $i = 0$, obteniendo que $N^2 = k + 0 \cdot l + m$, pero $N^2 > k + m$, ya que $l > 0$.
   - Por lo tanto, $N^2 \neq k + m$ y por ende, Prod no es regular.
2. Similar a antes, necesitamos demostrar que el lenguaje $\text{Mod} = \{a^i b^j c^{i \bmod j} \in \Sigma^+ \mid i \geq 0 \land j \geq 2\}$ no es regular. Para ello, nuevamente utilizamos el contra positivo del lema de bombeo.
   - Para todo $N > 0$, elegimos la palabra $M = a^N b^{2N} c^N$.
   - Elegimos x, y, z tal que $M = xyz$, $x = a^N b^{2N}$, $y = c^N$ $(|y| \geq N)$ y $z = \varepsilon$.
   - Para toda descomposición $y = c^k c^l c^m$ con $k + l + m = N$ y $l > 0$.
   - Tenemos nuestra nueva palabra $M' = a^N b^{2N} c^{k+i \cdot l+m}$.
   - Elegimos bombear con $i = 0$, obteniendo que $N = k + 0 \cdot l + m$, pero $N > k + m$, ya que $l > 0$.
   - Por lo tanto, $N \neq k + m$ y por ende, Mod no es regular.

Distribución de puntajes

Ambos ítems valen 6 puntos cada uno, distribuidos de la siguiente forma:

- **(1 pto.)** Por el correcto uso del lema de bombeo.
- **(1 pto.)** Por la correcta elección de la palabra P o M, esta debe estar bien definida, dadas las restricciones del lenguaje.
- **(1 pto.)** Por la correcta elección de x, y y z.
- **(1 pto.)** Por la correcta elección de i al bombear.
- **(2 ptos)** Por la correcta demostración de que la palabra bombeada no pertenece al lenguaje.

Problema 4

Sea $\mathcal{A} = (Q, \Sigma, \Delta, I, F)$ un autómata finito no-determinista (NFA). Se define el lenguaje universal de $\mathcal{A}$ como:

$$\mathcal{L}_\forall(\mathcal{A}) = \{w \in \Sigma^* \mid \text{ toda ejecución de } \mathcal{A} \text{ sobre } w \text{ es de aceptación } \}$$

Demuestre que para todo NFA $\mathcal{A}$, el lenguaje $\mathcal{L}_\forall(\mathcal{A})$ es siempre regular.

Solución

Sea $\mathcal{A} = (Q, \Sigma, \Delta, I, F)$ un NFA. Tenemos el siguiente lenguaje:

$$\mathcal{L}_\forall(\mathcal{A}) = \{w | \text{ todas las ejecuciones de } \mathcal{A} \text{ sobre w que son de aceptación}\}$$

Para demostrar que el lenguaje es regular, definimos el siguiente autómata:

$$\mathcal{A}_\forall^{det} = (2^Q, \Sigma, \Delta^{det}, q_0^{det}, F_\forall^{det})$$

- $2^Q = \{S | S \subseteq Q\}$
- $q_0^{det} = I$
- $\Delta^{det}(S, a) = \{\exists p \in S.(p, a, q) \in \Delta\}$
- $F_\forall^{det} = \{S \subseteq Q | S \subseteq F\} = 2^F$

En palabras simples, como funciona el autómata es igual que cuando se determiniza un NFA, solamente que lo que cambia es que ahora los estados finales son todas las combinaciones que se pueden hacer con los estados finales de $\mathcal{A}$, con esto nos aseguramos de que todas las ejecuciones sean aceptadas por dicho autómata.

Es importante ver que aca $F_\forall^{det}$ contiene al vacio. Esto, dado que si es que no existen ejecuciones de $\mathcal{A}$ sobre w, entonces automáticamente se cumple la propiedad (dado que es un para todo).

Sea $w = a_1 a_2 a_3 ... a_n \in \Sigma^*$ cualquiera. Como $\mathcal{A}_\forall^{det}$ es determinista. Sea ρ la ejcución de $\mathcal{A}_\forall^{det}$ sobre w, tal que:

$$\rho : S_0 \xrightarrow{a_1} S_1 \xrightarrow{a_2} ... \xrightarrow{a_n} S_n$$

En clases se vio en la demostración de determinización que:

Lema 1. Si $\rho : q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} ... \xrightarrow{a_n} q_n$ es una ejecución de $\mathcal{A}$ sobre w, entonces $\forall i.\ p_i \in S_i$.

Lema 2. $\forall i \leq n.\ \forall p \in S_i$, existe una ejecución $\rho : q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} ... \xrightarrow{a_i} q_i$ de $\mathcal{A}$ sobre w tal que $p_i = p$.

Usaremos ambos lemas demostrados en clases para demostrar la correctitud de la construcción.

Por demostrar. $w \in \mathcal{L}(\mathcal{A}_\forall^{det}) \longleftrightarrow w \in \mathcal{L}_\forall(\mathcal{A})$

Como sabemos, si $w \in \mathcal{L}(\mathcal{A}_\forall^{det})$ entonces $S_n \subseteq 2^F$, por lo tanto, la siguiente demostración es equivalente:

$$S_n \subseteq 2^F \longleftrightarrow w \in \mathcal{L}_\forall(\mathcal{A})$$

($\rightarrow$) Suponga que $S_n \subseteq 2^F$. Sea $\rho : q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} ... \xrightarrow{a_n} q_n$ una ejecución de aceptación de $\mathcal{A}$ sobre w cualquiera. Por lema 1: $p_n \in S_n$, por lo tanto $p_n \in F$. Con esto tenemos que ρ es de aceptación. Como esto sirve para cualquier ejecución ρ, entonces todas las ejecuciones son de aceptación. Con esto concluimos que $w \in \mathcal{L}_\forall(\mathcal{A})$.

($\leftarrow$) Vamos a demostrar por contrapositivo que $S_n \not\subseteq 2^F \rightarrow w \notin \mathcal{L}_\forall(\mathcal{A})$.

Suponga que existe $p \in S_n$ tal que $p \notin F$. Por lema 2: $\exists \rho : p_0 \xrightarrow{a_1} ... \xrightarrow{a_n} p_n$ de $\mathcal{A}$ sobre w tal que $p_n = p$. Como $p_n = p \notin F$, entonces ρ no es de aceptación. Por lo tanto $w \notin \mathcal{L}(\mathcal{A}_\forall^{det})$.

Distribución de puntajes

- **(2 ptos.)** Por definir correctamente la determinización de $\mathcal{A}$.
- **(2 ptos.)** Por definir correctamente $F_\forall^{det}$.
- **(1 pto.)** Por demostrar que $w \in \mathcal{L}(\mathcal{A}_\forall^{det}) \rightarrow w \in \mathcal{L}_\forall(\mathcal{A})$.
- **(1 pto.)** Por demostrar que $w \in \mathcal{L}(\mathcal{A}_\forall^{det}) \leftarrow w \in \mathcal{L}_\forall(\mathcal{A})$.