Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación

IIC2223 — Teoría de Autómatas y Lenguajes Formales — 2' 2021

PAUTA INTERROGACIÓN 1

Pregunta 1

Podemos interpretar el lenguaje buscado como "Las palabras sobre el alfabeto $\{a, b, c\}$ que no contienen la subpalabra abc ni la subpalabra acc".
Una posible expresión regular que satisface lo que buscamos es:

$$(b + c)^* \left(a^+ \cdot (b + c) \cdot (b \cdot (b + c)^*)^?\right)^* a^*$$

La idea de la expresión es primero capturar las palabras que no contengan la letra a (primer paréntesis), luego, en caso de tener (al menos) una a seguida de una b o c, debe haber una letra b antes de la siguiente c para evitar que se formen los patrones (segundo paréntesis) y, finalmente, capturamos las palabras formadas solo por letras a (a^* final).

Dado lo anterior, la distribución de puntaje es la siguiente:

- **(1 punto)** Por manejar correctamente las palabras que no contienen letras a. ($(b + c)^*$)
- **(0.5 puntos)** Por manejar correctamente las palabras que contienen solamente letras a. (a^*)
- **(4.5 puntos)** Por manejar correctamente las palabras que contienen las 3 letras posibles.
  - **(0.5 puntos)** Por considerar que las palabras pueden tener varias letras a. (a^+)
  - **(2 puntos)** Por notar que el patrón se comienza a formar al leer una b o c luego de la última a. ($(b + c)$)
  - **(2 puntos)** Por evitar que se lea una c luego de leer ab y ac. ($b \cdot (b + c)^*$)
- **Nota:** Si la expresión no está completamente correcta, se descontará puntaje por las palabras que están en el lenguaje, pero no están definidas por la expresión y por las palabras definidas por la expresión que no estén en el lenguaje.

Otra forma, igualmente correcta, de resolver el problema era construir el autómata que definía el lenguaje y luego utilizar MNY para obtener la expresión asociada.
Idealmente, luego de completar el proceso se haría un análisis de la expresión encontrada para eliminar redundancias y luego se daría una explicación de cómo funciona la expresión encontrada.

Para este caso, la distribución de puntaje es la siguiente:

- **(6 puntos)** Si la expresión encontrada es correcta.
- **(3-5.5 puntos)** Si la expresión encontrada no produce "abc" ni "acc", pero hay palabras en el lenguaje que no define. El puntaje dependerá de la explicación dada, la legibilidad de la expresión y de cuántos casos no estén cubiertos.
- **(0-4 puntos)** Si la expresión encontrada produce "abc" **o** "acc". El puntaje dependerá de la explicación dada, la legibilidad de la expresión y de si existen palabras en el lenguaje que no están definidas por la expresión.
- **(0-2.5 puntos)** Si la expresión encontrada produce "abc" **y** "acc". El puntaje dependerá de la explicación dada, la legibilidad de la expresión y de si existen palabras en el lenguaje que no están definidas por la expresión.

Pregunta 2

Para demostrar que el lenguaje es regular, basta con demostrar que existe un autómata que lo define.
Como L es regular, sabemos que existe un DFA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ tal que $\mathcal{L}(\mathcal{A}) = L$.
A partir de este, podemos contruir el autómata $\mathcal{A}' = (Q', \Sigma, \Delta, q_0, q_f)$ con:

$$Q' = Q^1 \uplus Q^2 \uplus \{q_f\} \text{ , dónde } Q^i = \{q^i \mid q \in Q\}$$

$$\begin{aligned}\Delta = &\left\{(p^1, a, q^1) \;\mid\; \delta(p,a) = q \,\wedge\, p^1, q^1 \in Q^1\right\}\\ \cup &\left\{(p^1, \varepsilon, q^2) \;\mid\; \exists a.\, \delta(p,a) = q \,\wedge\, p^1 \in Q^1 \,\wedge\, q^2 \in Q^2\right\}\\ \cup &\left\{(p^2, a, q_f) \;\mid\; \delta(p,a) \in F \,\wedge\, p^2 \in Q^2\right\}\end{aligned}$$

Intuitivamente, este autómata es igual al original (pero los estados finales dejan de serlo) y además tiene una copia de todos sus estados y un único estado final. Luego, el autómata intentará, de manera no determinista, adivinar que se encuentra donde debería estar la penúltima letra de la palabra y tomará la épsilon transición simulando que leyó esta letra; si tomó la transición en el momento correcto, debería quedar solamente una letra por leer que lleva a un estado final en el autómata original, por lo que simulamos este comportamiento con el último grupo de transiciones de Δ.
A continuación, debemos demostrar que $\mathcal{L}(\mathcal{A}') = L - 2$.

- $L - 2 \subseteq \mathcal{L}(\mathcal{A}')$:

  Sea $w = a_1 ... a_n \in L - 2$. Sabemos por definición que existe b tal que $a_1 ... a_{n-1} b a_n \in L$, por lo que existirá la siguiente ejecución de aceptación de $\mathcal{A}$:

  $$\rho_{\mathcal{A}} :\; q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} ... \xrightarrow{a_{n-1}} q_{n-1} \xrightarrow{b} q \xrightarrow{a_n} q_n \,\wedge\, q_n \in F$$

  Sabemos por la construcción de $\mathcal{A}'$ que $(q^1_{n-1}, \varepsilon, q^2) \in \Delta$ y que $(q^2, a_n, q_f) \in \Delta$, por lo que podremos formar la siguiente ejecución en $\mathcal{A}'$:

  $$\rho'_{\mathcal{A}} :\; q^1_0 \xrightarrow{a_1} q^1_1 \xrightarrow{a_2} ... \xrightarrow{a_{n-1}} q^1_{n-1} \xrightarrow{\varepsilon} q^2 \xrightarrow{a_n} q_f$$

  Por tanto, $w \in \mathcal{L}(\mathcal{A}')$.

- $\mathcal{L}(\mathcal{A}') \subseteq L - 2$:

  Similarmente, sea $w = a_1 ... a_n \in \mathcal{L}(\mathcal{A}')$. Por construcción sabes que una ejecución de aceptación de $\mathcal{A}'$ sobre w será de la forma:

  $$\rho'_{\mathcal{A}}: \; q_0^1 \xrightarrow{a_1} q_1^1 \xrightarrow{a_2} ... \xrightarrow{a_{n-1}} q_{n-1}^1 \xrightarrow{\varepsilon} q^2 \xrightarrow{a_n} q_f$$

  Sabemos entonces por la construcción de $\mathcal{A}'$ que existe b tal que $\delta(q, b) = q^2$ y que $\delta(q^2, a_n) \in F$. Por lo que podemos construir la siguiente ejecución de $\mathcal{A}$:

  $$\rho_{\mathcal{A}}: \; q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} ... \xrightarrow{a_{n-1}} q_{n-1} \xrightarrow{b} q \xrightarrow{a_n} q_n \;\wedge\; q_n \in F$$

  Entonces $w \in L - 2$. Por tanto $\mathcal{L}(\mathcal{A}') = L - 2$.

Dado lo anterior, la distribución de puntaje es la siguiente:

- **(1 punto)** Por definir correctamente los estados de $\mathcal{A}'$.
- **(3 puntos)** Por definir Δ correctamente:
  - **(1 punto)** Por simular las transiciones de $\mathcal{A}$.
  - **(1 punto)** Por simular la lectura de la penúltima letra de la palabra.
  - **(1 punto)** Por simular la lectura de la última letra de la palabra.
- **(1 punto)** Por demostrar $L - 2 \subseteq \mathcal{L}(\mathcal{A}')$.
- **(1 punto)** Por demostrar $\mathcal{L}(\mathcal{A}') \subseteq L - 2$.

Pregunta 3

El lenguaje Window(L) está definido por:

$$\text{Window}(L) = \{a^k \# w \mid k \geq 1 \wedge w \in \Sigma^* \wedge w|_k \in L\}$$

Un posible lenguaje regular L que hace que Window(L) no sea regular es:

$$L = \mathcal{L}(a^*)$$

Demostraremos que Window(L) no es un lenguaje regular. Sea un N cualquiera. Elegimos la palabra:

$$\underbrace{a^N \# b}_{x} \underbrace{a^N}_{y} \underbrace{\epsilon}_{z}$$

Tal que se cumpla:

$$a^N = \underbrace{a^i}_{u} \underbrace{a^j}_{v} \underbrace{a^k}_{w}$$

y tal que $i + j + k = N$, con $j \neq 0$. Al bombear esta palabra, tenemos lo siguiente:

$$a^N \# b a^i (a^j)^0 a^k = a^N \# b a^i a^k$$

Esta última palabra no pertenece a Window(L), dado que $(ba^{i+k})|_N = ba^{i+k} \notin L$, por lo tanto según el lema de bombeo Window(L) no es un lenguaje regular.

Dado lo anterior, la distribución de puntaje es la siguiente:

- **(2 puntos)** Por encontrar un lenguaje L regular que cumpla con que $\text{Window}(L)$ no es regular.
- **(2 puntos)** Por encontrar una palabra perteneciente a $\text{Window}(L)$ para cada N y dividirla correctamente en x, y, z.
- **(1 punto)** Por bombear la palabra de forma que no pertenezca al lenguaje.
- **(1 punto)** Por utilizar correctamente el lema de bombeo.