Teoría de Lenguajes

Soluciones
1er. Parcial – Curso 2023

Ejercicio 1.- [Evaluación individual del obligatorio]

a) Describa en qué consisten los archivos *json* utilizados en el laboratorio. Mencione sus componentes y el contenido de cada uno. No tiene por qué escribir el nombre exacto de cada campo pero sí una breve descripción que haga referencia a su contenido.

R: Son archivos que podrían usarse para un chatbot. Cada archivo contiene una lista de objetos con tres campos principales: *tag*, *patterns* y *responses*. El campo *tag* tiene como valor una string que es el tópico del chat (ej. *greeting*). El campo *patterns* tiene una lista de *strings* con posibles mensajes que activan una de las respuestas del campo *responses*. El campo *responses* es una lista de strings con posibles respuestas a las entradas contenidas en el campo *patterns*.

b) Responda en función de lo realizado en su entrega de laboratorio.

i) ¿Cómo resolvieron en su grupo los requerimientos de contar? Por ejemplo, cuando se pide la cantidad de patrones de una intención. Este requerimiento era necesario en los programas 3,4 y 5.

R: Esta parte es dependiente de cada entrega pero una forma de hacerlo es obtener por cada response el valor como una string, partirla por el separador de lista y obtener su largo.

ii) ¿Cómo resolvieron en su grupo las simplificaciones de los archivos *json*?
Más precisamente en el programa 6, que se pide dejar únicamente el primer patrón y la primera respuesta de cada intención en caso de haber más de una; y en el programa 7, que se pide sustituir todas las intenciones, patrones y respuestas por T, P y R, respectivamente.

R: Esta parte es dependiente de cada entrega. Una forma de hacerlo es con **re.sub**, utilizando una expresión regular por campo. Encapsular en un grupo lo anterior y lo posterior al campo para replicarlos en la substitución.

c) ¿Qué funciones del módulo *re* de *Python* utilizó para resolver el laboratorio (ej. *re.sub*)? Describa brevemente el funcionamiento de cada una.

R: Esta parte es dependiente de cada entrega. Tres funciones habituales son re.search, re.findall y re.sub.

Ejercicio 2.-
Indique si las siguientes afirmaciones son verdaderas o falsas, justificando su respuesta en cada caso.

a) Si L es un lenguaje no Regular, entonces $\exists$ L′ lenguaje regular no vacío / L′ $\subseteq$ L

Verdadero. Como L no es Regular, es distinto de $\emptyset$. Considerando cualquier tira *w* de L, se puede tomar el lenguaje L´ = $\{w\}$ (formado por esa única tira w) el cual está incluido en L.

b) El lenguaje L_{2b} = {x / x es de la forma $a^{2p}b^{r+p}c^{r+j}$ con j≥p, r>0, 0<p≤2 } no es regular

Verdadero. L_{2b} NO es Regular.
Se verificará utilizando el contrarrecíproco del Pumping Lema.
Sea N la constante del PL, consideramos la tira $z = aab^{N+1}c^{N+1}$ que pertenece a L_{2b} (se considera p=1)

Se toman todas las descomposiciones z = uvw que cumplen: |v|>0 y |uv|≤N y se muestra que en todas ellas existe algún i, para el que la tira $uv^iw \notin L_{2b}$.

Familia 1:
u = ε
v = a
$w = ab^{N+1}c^{N+1}$

$z_i = (a)^i ab^{N+1}c^{N+1}$
Tomando i=0; $z_o = ab^{N+1}c^{N+1}$ que $\notin L_{2b}$ porque la cantidad de **a**´s al comienzo debe ser par (2 o 4 por la restricción de p).

Familia 2:
u = ε
$v = aab^p$ p≥0 p+2 ≤N
$w = b^{N+1-p}c^{N+1}$

$z_i = (aab^p)^i b^{N+1-p}c^{N+1}$
Nuevamente tomando i=0; $z_o = b^{N+1-p}c^{N+1}$ que $\notin L_a$ porque no comienza con al menos dos **a**'s.

Familia 3:
u = a p≥0
$v = ab^p$ 1+p≤N
$w = b^{N+1-p}c^N$

$z_i = a(ab^p)^i b^{N+1-p} c^{N+1}$
Tomando i=0; $z_o = ab^{N+1-q} c^{N+1}$ que $\notin L_{2b}$ porque z_0 comienza con una sola **a.**

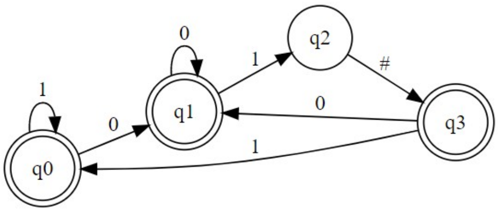
Familia 4:
$u = aab^p$ p≥0 q>0
$v = b^q$ 2+p+q≤N
$w = b^{N-p-q}c^{N+1}$

$z_i = aab^p(b^q)^i b^{N-p-q} c^{N+1}$
Tomando i=3; $z_3 = aab^pb^{3q} b^{N-p-q} c^{N+1} = aa\, b^{N+2q} c^{N+1} \notin L_{2b}$ porque $|z_3|_b > |z_3|_c$ lo cual viola la restricción de j≥p al ser q>0

Estas son las posibles descomposiciones de z que cumplen |v|>0 y |uv|≤ N
Entonces por el contrarecíproco del Pumping Lema para lenguajes regulares se puede afirmar que L_{2b} NO es Regular.

c) El lenguaje L_{2c} = { x / x $\in$ {0,1,#} y donde cada secuencia de 01 viene seguida de un # } es regular

Verdadero. Se puede construir el siguiente AFD M_{2c} / L_{2c} = L(M_{2c})



Ejercicio 3.-

Sea L_3 el lenguaje reconocido por el siguiente autómata finito M_3=(Q, Σ,δ,q_0, F) donde:
Q={q_0, q_1, q_2, q_3, q_4}   Σ={a,b}   F={q_1, q_3} y δ dada por:

	a	**b**
q_0	q_1	q_2
q_1	q_3	q_0
q_2	q_1	q_4
q_3	q_1	q_0
q_4	q_3	-

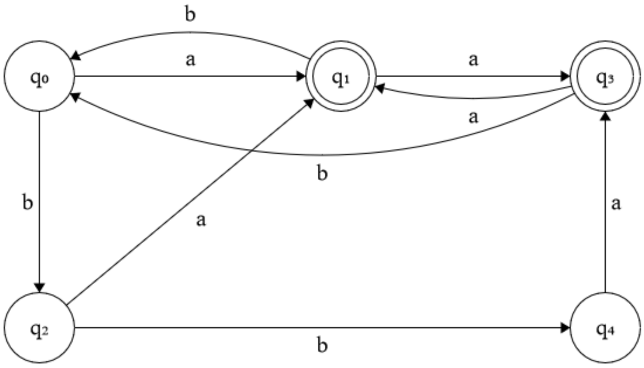
a) Defina la relación R_L
b) ¿Cuántas clases de equivalencia define la relación R_{L3}?
c) Halle las expresiones regulares que definen las clases según R_{L3} y dé una expresión regular *r* tal que L(*r*) = L(M_3). Justifique

Solución

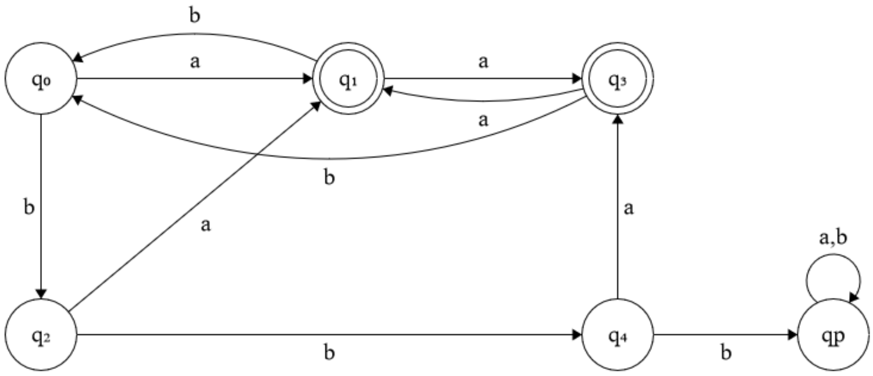
a) Sea un lenguaje L $\subseteq$ Σ * , x, y $\in$ Σ *
x RL y ⇔ ∀ z $\in$ Σ * se cumple que: xz $\in$ L ∧ yz $\in$ L   o
xz $\notin$ L ∧ yz $\notin$ L

b) Al tener el autómata mínimo (en la parte (c)), el número de clases será la respuesta a esta parte del ejercicio.

c)



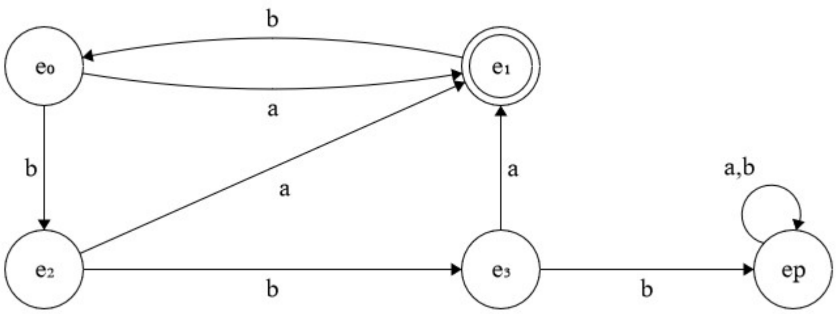
Lo primero que debemos hacer es minimizar el autómata finito M_{3a}. Para ello completaremos la función de transición para los casos en que no está definida, agregando un estado pozo q_p.



Ahora corremos el algoritmo de minimización.

Π_0 [q_0 q_2 q_4 q_p] [q_1 q_3]
Π_1 [q_0 q_2 q_4] [q_p] [$q_1 q_3$]
Π_2 [q_0 q_2] [q_4] [q_p] [$q_1 q_3$]
Π_3 [q_0] [q_2] [q_4] [q_p] [$q_1 q_3$]
Π_3 [q_0] [q_2] [q_4] [q_p] [$q_1 q_3$]

El autómata mínimo es entonces:



Ahora usamos el método de clases de equivalencia para hallar las expresiones regulares asociadas. Planteamos el sistema de ecuaciones basándonos en las transiciones entrantes de cada estado:

$X_0 = \varepsilon \mid X_1b$
$X_1 = X_0a \mid X_2a \mid X_3a$
$X_2 = X_0b$
$X_3 = X_2b$
$X_p = X_3b \mid X_p(a|b)$

Sustituimos X_0, X_2 y X_3 en X_1
$X_1 = (\varepsilon \mid X_1b)a \mid X_0ba \mid X_2ba = a \mid X_1ba \mid X_0ba \mid X_2ba$

Como queremos usar el Lema de Arden en su versión $X = X.r \mid s \Rightarrow X = s.r^*$, entonces intentamos sustituir hasta que X_1 esté escrita en términos de ella misma.

$X_1 = a \mid X_1ba \mid (\varepsilon \mid X_1b)ba \mid (X_0b)ba = a \mid X_1ba \mid ba \mid X_1bba \mid (\varepsilon \mid X_1b)bba = a \mid X_1ba \mid ba \mid X_1bba \mid bba \mid X_1bbba$

Reordenando la expresión: $X_1 = X_1(ba|bba|bbba) \mid a|ba|bba$

Aplicamos Arden: X_1 = (a|ba|bba).(ba|bba|bbba)*

Entonces ahora podemos hallar X_0 y por lo tanto X_2 y X_3:
X_0 = ε | X_1b = ε | (a|ba|bba).(ba|bba|bbba)*.b
X_2 = X_0b = (ε | (a|ba|bba).(ba|bba|bbba)*.b).b = b | (a|ba|bba).(ba|bba|bbba)*.bb
X_3 = X_2b = (b | (a|ba|bba).(ba|bba|bbba)*.bb).b = bb | (a|ba|bba).(ba|bba|bbba)*.bbb
Finalmente estamos en condiciones de hallar X_p

X_p = X_3b | X_p(a|b) = (bb | (a|ba|bba).(ba|bba|bbba)*.bbb)b | X_p(a|b)
= bbb | (a|ba|bba).(ba|bba|bbba)*.bbbb | X_p(a|b)

Y usando el Lema de Arden: X_P = (bbb | (a|ba|bba).(ba|bba|bbba)*.bbbb).(a|b)*
⇒ X_P = bbb(a|b)* | (a|ba|bba).(ba|bba|bbba)*.bbbb)(a|b)*

Como el autómata es mínimo y completo, entonces por el corolario del Teorema de Myhill-Nerode, las clases de equivalencia de R_M y las de R_L coinciden.

Finalmente, la expresión X_1 es la que define las tiras del lenguaje aceptado por el autómata, ya que es la asociada al único estado final.

Entonces L(M) = L((a|ba|bba).(ba|bba|bbba)*)

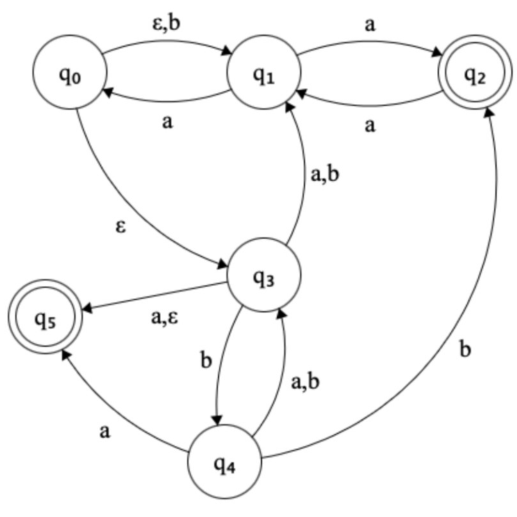
Ejercicio 4.-

Sea L_4 el lenguaje reconocido por el siguiente autómata finito M_4=(Q, Σ,δ,q_0, F) donde:
Q={q_0, q_1, q_2, q_3, q_4, q_5} Σ={a,b} F={q_2, q_5} y δ dada por:

	a	**b**	**ε**
q_0	∅	{q_1}	{q_1,q_3}
q_1	{q_0,q_2}	∅	∅
q_2	{q_1}	∅	∅
q_3	{q_1,q_5}	{q_1,q_4}	{q_5}
q_4	{q_3,q_5}	{q_2,q_3}	∅
q_5	∅	∅	∅

Construya un autómata finito determinista M′ / L(M_4)=L(M′)

Solución



Pasaje de AFND-eps → AFND

$\varepsilon_{cl}(q_0) = \{q_0, q_1, q_3, q_5\}$
$\varepsilon_{cl}(q_1) = \{q_1\}$
$\varepsilon_{cl}(q_2) = \{q_2\}$
$\varepsilon_{cl}(q_3) = \{q_3, q_5\}$
$\varepsilon_{cl}(q_4) = \{q_4\}$
$\varepsilon_{cl}(q_5) = \{q_5\}$

$\delta'(q_0, a) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_0), a)) = \varepsilon_{cl}(\tilde{\delta}(\{q_0, q_1, q_3, q_5\}, a)) = \varepsilon_{cl}(\{q_0, q_2, q_1, q_5\}) = \{q_0, q_1, q_2, q_3, q_5\}$
$\delta'(q_0, b) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_0), b)) = \varepsilon_{cl}(\tilde{\delta}(\{q_0, q_1, q_3, q_5\}, b)) = \varepsilon_{cl}(\{q_1, q_4\}) = \{q_1, q_4\}$
$\delta'(q_1, a) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_1), a)) = \varepsilon_{cl}(\tilde{\delta}(\{q_1\}, a)) = \varepsilon_{cl}(\{q_0, q_2\}) = \{q_0, q_1, q_2, q_3, q_5\}$
$\delta'(q_0, b) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_1), b)) = \varepsilon_{cl}(\tilde{\delta}(\{q_1\}, b)) = \varepsilon_{cl}(\{\}) = \{\}$
$\delta'(q_2, a) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_2), a)) = \varepsilon_{cl}(\tilde{\delta}(\{q_2\}, a)) = \varepsilon_{cl}(\{q_1\}) = \{q_1\}$
$\delta'(q_2, b) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_2), b)) = \varepsilon_{cl}(\tilde{\delta}(\{q_2\}, b)) = \varepsilon_{cl}(\{\}) = \{\}$
$\delta'(q_3, a) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_3), a)) = \varepsilon_{cl}(\tilde{\delta}(\{q_3, q_5\}, a)) = \varepsilon_{cl}(\{q_1, q_5\}) = \{q_1, q_5\}$
$\delta'(q_3, b) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_3), b)) = \varepsilon_{cl}(\tilde{\delta}(\{q_3, q_5\}, b)) = \varepsilon_{cl}(\{q_1, q_4\}) = \{q_1, q_4\}$
$\delta'(q_4, a) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_4), a)) = \varepsilon_{cl}(\tilde{\delta}(\{q_4\}, a)) = \varepsilon_{cl}(\{q_3, q_5\}) = \{q_3, q_5\}$
$\delta'(q_4, b) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_4), b)) = \varepsilon_{cl}(\tilde{\delta}(\{q_4\}, b)) = \varepsilon_{cl}(\{q_2, q_3\}) = \{q_2, q_3, q_5\}$
$\delta'(q_5, a) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_5), a)) = \varepsilon_{cl}(\tilde{\delta}(\{q_5\}, a)) = \varepsilon_{cl}(\{\}) = \{\}$
$\delta'(q_5, b) = \varepsilon_{cl}(\tilde{\delta}(\varepsilon_{cl}(q_5), b)) = \varepsilon_{cl}(\tilde{\delta}(\{q_5\}, b)) = \varepsilon_{cl}(\{\}) = \{\}$

Entonces obtenemos el siguiente AFND:

δ'	a	b
q_0	$\{q_0, q_1, q_2, q_3, q_5\}$	$\{q_1, q_4\}$
q_1	$\{q_0, q_1, q_2, q_3, q_5\}$	$\{\}$
q_2	$\{q_1\}$	$\{\}$
q_3	$\{q_1, q_5\}$	$\{q_1, q_4\}$
q_4	$\{q_3, q_5\}$	$\{q_2, q_3, q_5\}$
q_5	$\{\}$	$\{\}$

con $F' = \{q_0, q_2, q_5\}$

Pasaje de AFND → AFD

δ''	a	b
$[q_0]$	$[q_0q_1q_2q_3q_5]$	$[q_1q_4]$
$[q_0q_1q_2q_3q_5]$	$[q_0q_1q_2q_3q_5]$	$[q_1q_4]$
$[q_1q_4]$	$[q_0q_1q_2q_3q_5]$	$[q_2q_3q_5]$
$[q_2q_3q_5]$	$[q_1q_5]$	$[q_1q_4]$
$[q_1q_5]$	$[q_0q_1q_2q_3q_5]$	$[]$

con $F'' = \{ q_0,\ q_0q_1q_2q_3q_5, q_2q_3q_5,\ q_1q_5\}$

Ejercicio 5.-

Se construye un autómata de dos cintas que acepta $L_4 = \{ (a^k b^n, b^n a^{(n \bmod 3)}) \ / \ n>0, k\geq 0, k \text{ par} \}$