Teoría de Autómatas y Lenguajes Formales

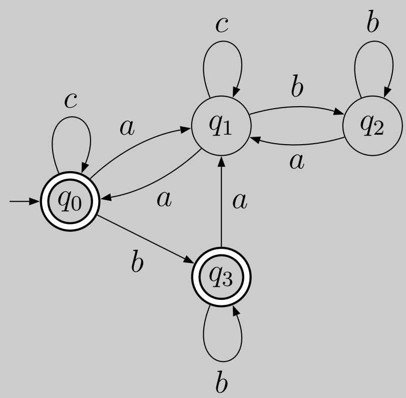
Mayo 2010 **Final**

Normas :

- La duración del examen es de 2 horas y 15 minutos.
- **Todos los ejercicios se entregarán en hojas separadas.**
- El examen tipo test cuenta hasta 2 puntos sobre la nota total.
- Además de obtener un mínimo de 5 puntos en el global del examen, para poder aprobar es necesario obtener un mínimo de 1.5 puntos entre las dos primeras preguntas y un mínimo de 1.5 puntos entre las dos últimas.

1. (a) (1.0 puntos) Defina un autómata finito no determinista para el siguiente lenguaje:
$L = \{x \mid x \in \{a, b, c\}^*,\ n_a(x)\ \textit{es par},\ \textit{y no existe ninguna subcadena bc en } x\}$.
Donde $n_a(x)$ representa el número de a's en x. Se supone que el 0 es un número par.

Solución:
Una posible solución es la siguiente:



(b) (1.0 puntos) Argumenta/demuestra por qué cualquier lenguaje finito sobre el alfabeto $\Sigma = \{a, b, c\}$ es un lenguaje regular.

Solución:
Sea L lenguaje finito qualesquiera sobre el alfabeto $\Sigma = \{a, b, c\} : L = \{w_1, ...w_n | w_i \in \{a, b, c\}^*\}$. Como se puede observar, L tiene n palabras formadas cada una por letras a, b y/o c. Por definición, un lenguaje (y por ello L) es regular si se puede generar o representar mediante un autómata finito (determinista o no), una gramática regular o una expresión regular. Pues, podemos definir facilmente

una expresión regular que representa el lenguaje L: $w_1 + w_2 + ... + w_n$, donde los w_i son las palabras del lenguaje L.

Por ello, podemos concluir que L es un lenguajes regular y, dado que L es un lenguaje finito qualesquiera, todos los leguajes finitos son regulares.

2. (a) (1.6 puntos) Demuestra que la expresión regular

$$RegExp \;=\; [((0+10)(10)^*(11+0)) \;+\; 11]\,(0+1)^*$$

es equivalente/no equivalente al autómata determinista A_{det}:

$$(\{0,1\},\{q_0,q_1,q_2,q_3\},f_{det},q_0,q_3)$$

f_{det}	0	1
$\rightarrow q_0$	q_1	q_2
q_1	q_3	q_2
q_2	q_1	q_3
$*q_3$	q_3	q_3

Todos los pasos para la resolución del ejercicio tienen que estar justificados utilizando algoritmos vistos en clase. No es obligatorio, pero sí conveniente, dar una pequeña explicación de qué se hace y porqué se hace.

(b) (0.4 puntos) Defina el lenguaje generado por la expresión regular en notación conjuntista ($L = \{...|...\}$).

Solución:

Convertimos la expresión regular en un autómata no determinista equivalente

$$(\{0,1\},\{A,B,C,D,E,F,G\},f_{regexp},A,F)$$

f_{regexp}	0	1	λ
→A	C	$\{B,G\}$	
B	C		
C	F	$\{D,E\}$	
D	C		
E		F	
*F	F	F	
G		F	

Convertimos el autómata no determinista en un autómata determinista equivalente

$$(\{0,1\},\{A,B,C,D,E,F,G\},f_{regexp},A,F)$$

f_{regexp}	0	1
→A	C	$\{B,G\}$
C	F	$\{D,E\}$
$\{B,G\}$	C	F
*F	F	F
$\{D,E\}$	C	F

Renombramos los estados

$$(\{0,1\},\{Q_A,Q_B,Q_C,Q_D,Q_F\},f_{regexp},Q_A,Q_F)$$

f_{regexp}	0	1
$\rightarrow Q_A$	Q_C	Q_B
Q_C	Q_F	Q_D
Q_B	Q_C	Q_F
$*Q_F$	Q_F	Q_F
Q_D	Q_C	Q_F

Minimizamos el autómata, calculando el conjunto cociente Q/E:

$$Q/E_0 = \{\{Q_A,Q_C,Q_B,Q_D\},\{Q_F\}\}$$

$$Q/E_1 = \{\{Q_A\},\{Q_C\},\{Q_B,Q_D\},\{Q_F\}\}$$

$$Q/E_2 = \{\{Q_A\},\{Q_C\},\{Q_B,Q_D\},\{Q_F\}\}$$

$$Q/E_2 = Q/E_1 \quad \textit{por lo tanto} \quad Q/E_1 = Q/E$$

$$\underbrace{\{Q_A\}}_{q_0},\underbrace{\{Q_C\}}_{q_1},\underbrace{\{Q_B,Q_D\}}_{q_2},\underbrace{\{Q_F\}}_{q_3}$$

f_{det}	0	1
$\rightarrow q_0$	q_1	q_2
q_1	q_3	q_2
q_2	q_1	q_3
$*q_3$	q_3	q_3

Como se puede ver, el autómata determinista equivalente a la expresión regular *RegExp* coincide con el autómata A_{det}. El lenguaje que definen es

$$L = \{x \in \{0,1\}^* \mid x \textit{ contiene } 00 \textit{ o } x \textit{ contiene } 11\}$$

3. (a) 1.4 puntos

Diseña una gramática independiente de contexto que genere el siguiente lenguaje

$$L = \{a^x b^y c^z \mid x \geq y - z;\ y \geq z;\ x, y, z \geq 0\}$$

Ten en cuenta las siguientes pistas:

- Por la definición del lenguaje, las b's que hay en la palabra están "compensadas" por c's y por a's. Es decir, se puede considerar que las b's se parten en dos grupos: 1) un grupo de b's coincide en número con las c's al final de la palabra y 2) para el otro grupo de b's hay por lo menos el mísmo número de a's al principio de la palabra.
- Aparte de las a's que "compensan" b's, puede haber más a's al principio.

Solución:

- La pista 1 lleva directamente a las producciones:
  $Q ::= bc \mid bQc$
- La pista 2 y la pista 3 llevan directamente a las producciones:
  $P ::= Ab \mid APb$ y $A ::= a \mid aA$
- Se unen estas producciones mediante el axioma S:
  $S ::= A \mid P \mid Q \mid AQ \mid PQ \mid \lambda$

$G = (\{abc\}, \{S, P, Q, A\}, S, P)$, P contiene las producciones de arriba.

(b) 0.1 puntos

Comprueba mediente derivaciones que las siguientes palabras pertenecen a L: aaa, $aabb$, $abbc$, $aaabb$, $aaabbbccc$, $bbbccc$

Solución:

$aaa : S \Rightarrow A \Rightarrow aA \Rightarrow aaA \Rightarrow aaa$
$aabb : S \Rightarrow P \Rightarrow APb \Rightarrow aPb \Rightarrow aAbb \Rightarrow aabb$
$abbc : S \Rightarrow PQ \Rightarrow AbQ \Rightarrow abQ \Rightarrow abbc$
$aaabb : S \Rightarrow P \Rightarrow APb \Rightarrow aAPb \Rightarrow aaPb \Rightarrow aaabb$

$aaabbbccc$: $S \Rightarrow AQ \Rightarrow aAQ \Rightarrow aaAQ \Rightarrow aaaQ \Rightarrow aaabQc \Rightarrow aaabbQcc \Rightarrow aaabbbccc$
$bbbccc$: $S \Rightarrow Q \Rightarrow bQc \Rightarrow bbQcc \Rightarrow bbbccc$

(c) 0.3 puntos
Escribe tu gramática en forma normal de Greibach (FNG).

Solución:

La gramática G no está bien formada (contiene producciones unitarias). La siguiente gramática G' está bien formada.

$G' = (\{abc\}, \{S, P, Q, A\}, S, P)$, P contiene las siguientes producciones:
$Q ::= bc \mid bQc$
$P ::= Ab \mid APb$
$A ::= a \mid aA$
$S ::= a \mid aA \mid Ab \mid APb \mid bc \mid bQc \mid AQ \mid PQ \mid \lambda$

G' no contiene recursiones, por tanto, es suficiente reemplazar las variables en los cuerpos de las producciones, primero en P y, posteriormente, en S.

$G_G = (\{abc\}, \{S, P, Q, A\}, S, P)$, P contiene las siguientes producciones:
$Q ::= bc \mid bQc$
$P ::= ab \mid aAb \mid aPb \mid aAPb$
$A ::= a \mid aA$
$S ::= a \mid aA \mid ab \mid aAb \mid aPb \mid aAPb \mid bc \mid bQc \mid aQ \mid aAQ \mid abQ \mid aAbQ \mid aPbQ \mid aAPbQ \mid \lambda$

(d) 0.2 puntos
Escribe tu gramática en forma normal de Chomsky (FNC).

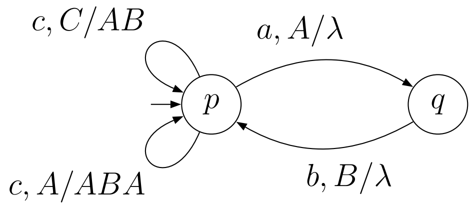
Solución:

La gramática G' del ejercicio anterior está bien formada.

$G_C = (\{abc\}, \{S, P, Q, A, A_1, B, C\}, S, P)$, P contiene las siguientes producciones:
$Q ::= BC \mid BT_2$
$P ::= AB \mid AT_1$
$A ::= a \mid A_1A$
$S ::= a \mid A_1A \mid AB \mid AT_1 \mid BC \mid BT_2 \mid AQ \mid PQ \mid \lambda$
$A_1 ::= a$
$B ::= b$
$C ::= c$
$T_1 ::= PB$
$T_2 ::= QC$

4. Dado el siguiente autómata a pila:

$AP = (\{a, b, c\}, \{A, B, C\}, \{p, q\}, p, C, f)$, donde f según el grafo



(a) 1.5 puntos

Convierte el autómata a pila en una gramática independiente de contexto.

Solución:

Construimos la siguente gramática independiente de contexto
$G = (\Sigma_N, \{a, b, c\}, C, P)$, donde el conjunto de variable Σ_N consta de
$\Sigma_N = \{C \cup \{[pCq]\}|p, q \in Q; X \in \Gamma\}$.
Las producciones P de G son como sigue:

- $C ::= [pCp] \mid [pCq]$.
- $f(p, c, C) = \{p, AB\} \Rightarrow [pCr_2] ::= c[pAr_1][r_1Br_2]$
  $[pCp] ::= c[pAp][pBp]$
  $[pCp] ::= c[pAq][qBp]$
  $[pCq] ::= c[pAp][pBq]$
  $[pCq] ::= c[pAq][qBq]$
- $f(p, c, A) = \{p, ABA\} \Rightarrow [pCr_3] ::= c[pAr_1][r_1Br_2][r_2Ar_3]$
  $[pAp] ::= c[pAp][pBp][pAp]$
  $[pAp] ::= c[pAq][qBp][pAp]$
  $[pAp] ::= c[pAp][pBq][qAp]$
  $[pAp] ::= c[pAq][qBq][qAp]$
  $[pAq] ::= c[pAp][pBp][pAq]$
  $[pAq] ::= c[pAq][qBp][pAq]$
  $[pAq] ::= c[pAp][pBq][qAq]$
  $[pAq] ::= c[pAq][qBq][qAq]$
- $f(p, a, A) = \{q, \lambda\} \Rightarrow [pAr_1] ::= a[q\lambda r_1]$
  $[pAp] ::= a[q\lambda p] \Leftarrow$ No existe
  $[pAq] ::= a[q\lambda q] = a$
- $f(q, b, B) = \{p, \lambda\} \Rightarrow [qBr_1] ::= b[p\lambda r_1]$
  $[qBp] ::= b[p\lambda p] = b$
  $[qBq] ::= b[p\lambda q] \Leftarrow$ No existe

(b) 0.1 puntos

Demuestra el reconocimiento de la palabra *ccabcabab* por el autómata a pila mediante una secuencia de descripciones instantáneas.

Solución:
$(p, ccabcabab, C) \vdash (p, cabcabab, AB) \vdash (p, abcabab, ABAB) \vdash (q, bcabab, BAB) \vdash (p, cabab, AB) \vdash (p, abab, ABAB) \vdash (q, bab, BAB) \vdash (p, ab, AB) \vdash (q, b, B) \vdash (p, \lambda, \lambda)$

(c) 0.3 puntos

Demuestra la generación de la misma palabra *ccabcabab* por la gramática generada en la primera parte de este ejercicio.

Solución:
$C \Rightarrow [pCp] \Rightarrow c[pAq][qBp] \Rightarrow cc[pAq][qBp][pAq][qBp] \Rightarrow cca[qBp][pAq][qBp] \Rightarrow ccab[pAq][qBp] \Rightarrow ccabc[pAq][qBp][pAq][qBp] \Rightarrow ccabca[qBp][pAq][qBp] \Rightarrow ccabcab[pAq][qBp] \Rightarrow ccabcaba[qBp] \Rightarrow ccabcabab$

(d) 0.1 puntos

Convierte el autómata a pila AP en un autómata a pila reconocedor por estado final.

Solución:
$AP_F = (\{a, b, c\}, \{A, B, C, S\}, \{p, q, p_0, p_f\}, p_0, S, f, \{p_f\})$, donde f consta de las siguientes transiciones:

- $f(p_0, \lambda, S) = \{(p, CS)\}$
- Incluir todas las transiciones de AP
- $f(p, \lambda, S) = \{p_f, \lambda\}$