Teoría de Lenguajes

2do. Parcial – Curso 2019
Soluciones

Ejercicio 1 [Evaluación individual del obligatorio]

a) Considere la siguiente gramática con símbolo inicial N:

N -> N D | D
D -> '0' | '1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9'

y la función:

```
def printT(t):
    print(t.label())
    for n in t:
        if type(n) is nltk.Tree:
            printT(n)
        else:
             print(n)
        break
```

Escriba la salida en pantalla al invocar `printT` con el árbol correspondiente a cada una de las tiras presentadas a continuación.

i) 7654321
ii) 123
iii) 1

Nota: la instrucción `break` termina la ejecución de la iteración.

b) Escriba una gramática lineal derecha – con la sintaxis *nltk* - que reconozca números binarios divisibles entre 3 (programa1.py del Laboratorio).

c) Explique cómo resolvió la detección del sublenguaje de los strings en el programa3.py del Laboratorio.

Solución

a)

i) NNNNNNND7 (separados por \n pero ambas se toman bien)
ii) NNND1
iii) ND1

b)

$S \rightarrow$ '0' S | '1' S_1 | '0'
$S_1 \rightarrow$ '0' S_2 | '1' S | '1'
$S_2 \rightarrow$ '0' S_1 | '1' S_2

c)

S → Letra S | Letra
Letra → '_' | 'a' | 'b' | 'c' |.... | 'z' | 'A' | 'B' | 'C' | .... | 'Z'

Ejercicio 2

Dado el siguiente lenguaje:

$L_2 = \{ xyz \ / \ y \in L(0^*1^*2^*) \ ; \ x,z \in \{a,b\}^* \ ; \ |x|_a=|z|_b \ ; \ |y|_0=|y|_2>0 \ ; \ |y|_1=2^*|x|_a \}$

a) Clasifique L_2 según la Jerarquía de Chomsky. Justifique.
b) Construya una gramática G_2 / $L(G_2) = L_2$.
c) Construya un autómata M_2 / $L(M_2) = L_2$.

Solución

a) Clasifique L_2 según la Jerarquía de Chomsky. Justifique.

L_2 es R.E. por la GI de la parte b) y por la M.T. de la parte c)
Y demostraremos que no es L.C. usando el Contrareciproco del Pumping Lema para Lenguajes L.C.

Dado n, elegimos $z = a^n 0\ 1^{2n}\ 2\ b^n$ pertenece a L_2 y $|z| = 4n+2$
Estudiamos todas las descomposiciones de z=uvwxy, que cumplen: |vx|=>1 y |vwx|<=n

casos	a^n	0	1^{2n}	2	b^n
1	v x				
2	v x	x	x		
3	v		x		
4	v	v	v x		
5			v x		
6			vx	x	x
7			v		x
8			v	v	v x
9					v x

Caso 1)
$u = a^p$
$v = a^q \quad q+l=>1$
$w = a^k \quad q+k+l<=n$
$x = a^l$
$y = a^{n-p-q-k-l} 0\ 1^{2n}\ 2\ b^n$

$z_i = a^p (a^q)^i a^k (a^l)^i a^{n-p-q-k-l} 0\ 1^{2n}\ 2\ b^n$
$= a^{p+qi+k+l.i+n-p-q-k-l} 0\ 1^{2n}\ 2\ b^n$
$= a^{n + (q+l).(i-1)} 0\ 1^{2n}\ 2\ b^n$

Elegimos i=0, $z_0 = a^{n-(q+l)} 0\ 1^{2n}\ 2\ b^n$ no pertenece a L_2, dado que la cantidad de a_s en la primera parte es distinta a la cantidad de b_s del final.

Caso 9) Es análogo al caso 1), elegimos i=0, z_0 no pertenece a L_2, dado que la cantidad de a_s en la primera parte es distinta a la cantidad de b_s del final.

Caso 5) Es análogo al caso 1), elegimos i=0, z_0 no pertenece a L_2, dado que la cantidad de 1s no es el doble de la cantidad de a_s de la primera parte.

Nota: Las gramáticas y los autómatas deben corresponderse con el tipo del lenguaje considerado en cada caso, según la Jerarquía de Chomsky. Se valora positivamente la simplicidad de las soluciones propuestas así como una breve explicación de éstas. Todas las respuestas deben estar debidamente justificadas.

Caso 3)
$u = a^{n-p-q}$
$v = a^p$ $p+l=>1$
$w = a^q\ 0\ 1^k$ $p+q+1+k+l<=n$
$x = 1^l$
$y = 1^{2n-k-l}\ 2\ b^n$

$z_i = a^{n-p-q}\ (a^p)^i\ a^q\ 0\ 1^k\ (1^l)^i\ 1^{2n-k-l}\ 2\ b^n$
$= a^{n+p.(i-1)}\ 0\ 1^{2n+l.(i-1)}\ 2\ b^n$

Si p=0, elegimos i=0 y nos queda que $z_0 = a^n\ 0\ 1^{2n-l}\ 2\ b^n$ no pertenece a L_2, dado que la cantidad de 1_s no es el doble de la cantidad de a_s de la primera parte.
Si p≠0, elegimos i=0 y nos queda que $z_0 = a^{n-p}\ 0\ 1^{2n-l}\ 2\ b^n$ no pertenece a L_2, dado que la cantidad a_s de la primera parte no es igual a la cantidad de b_s de la segunda parte.

Caso 7)
$u = a^n\ 0\ 1^{2n-p-q}$
$v = 1^p$ $p+l=>1$
$w = 1^q\ 2\ b^k$ $p+q+1+k+l<=n$
$x = b^l$
$y = b^{n-k-l}$

$z_i = a^n\ 0\ 1^{2n-p-q}\ (1^p)^i\ 1^q\ 2\ b^k\ (b^l)^i\ b^{n-k-l}$
$= a^n\ 0\ 1^{2n+p.(i-1)}\ 2\ b^{n+l.(i-1)}\ b^n$

Si p=0, elegimos i=0 y nos queda que $z_0 = a^n\ 0\ 1^{2n}\ 2\ b^{n-l}\ b^n$ no pertenece a L_2, dado que la cantidad a_s de la primera parte no es igual a la cantidad de b_s de la segunda parte.
Si p≠0, elegimos i=0 y nos queda que $z_0 = a^n\ 0\ 1^{2n+p.(i-1)}\ 2\ b^{n+l.(i-1)}\ b^n$ no pertenece a L_2, dado que la cantidad de 1s no es el doble de la cantidad de a_s de la primera parte.

Caso 2)
$u = a^{n-p-q-k}$
$v = a^p$ $p+k+1+l=>1$
$w = a^q$ $q+k+l<=n$
$x = a^k\ 0\ 1^l$ $k=>0$
$y = 1^{2n-l}\ 2\ b^n$ $l=>0$

$z_i = a^{n-p-q-k}\ (a^p)^i\ a^q\ (a^k\ 0\ 1^l)^i\ 0\ 1^{2n-l}\ 2\ b^n$

Elegimos i=0, y nos queda $z_0 = a^{n-p-k}\ a^q\ 1^{2n-l}\ 2\ b^n$ que no pertenece a L_2 dado que z_0 no tiene 0_s.

Caso 4) Es análogo al caso 2), elegimos i=0, z_0 no pertenece a L_2, dado que z_0 no tiene 0_s.

Caso 6) y 8) Son análogos al caso 2), elegimos i=0, z_0 no pertenece a L_2, dado que z_0 no tiene 1_s.

Como estudiamos todas las descomposiciones posibles de z, y para cada uno encontramos un i talque z_i no pertenece a L_2, concluimos por el C.R. del P.L.2 que L_2 no es un Lenguaje Libre de Contexto.

b) Construya una gramática G_2 / $L(G_2) = L_2$

S → BS'A

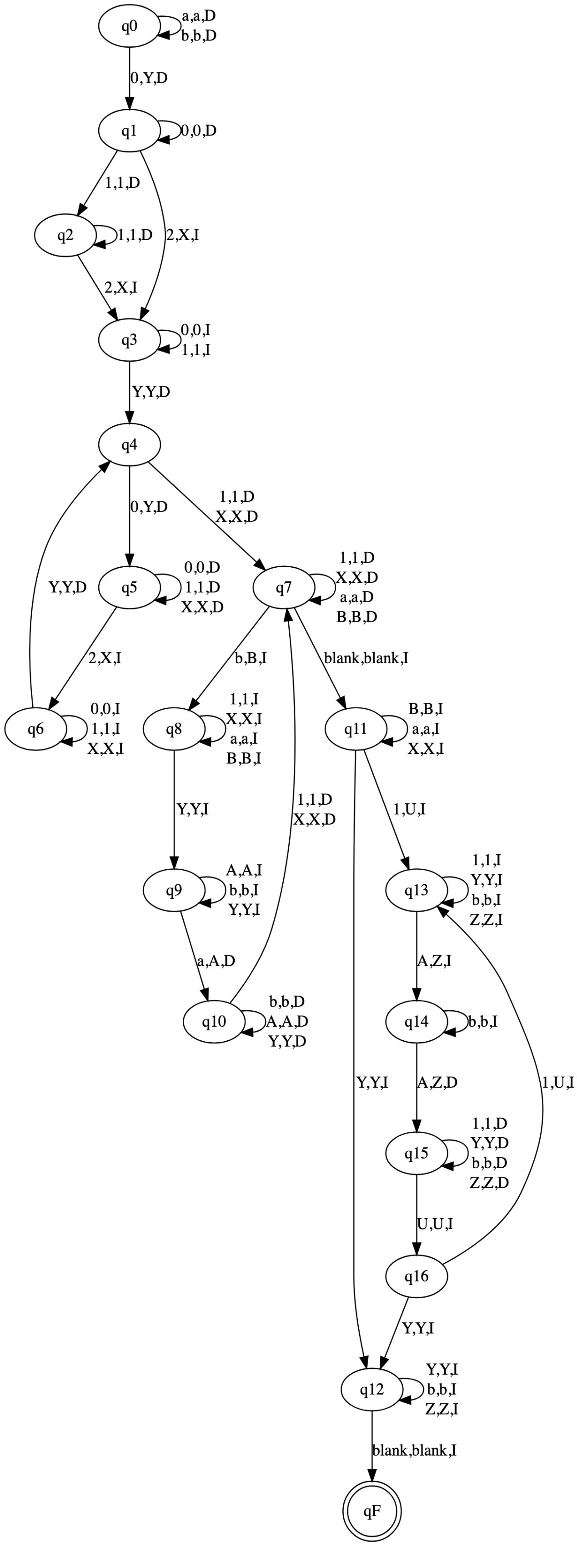
S' → aS'Xb
S' → 0C2
C → 0C2 | ε

Xb → bX
2X → X2
0X → 011
1X → 111

ab → ba
ba → ab

A → aA | ε
B → bB | ε

c) Construya un autómata M_2 / $L(M_2) = L_2$



Ejercicio 3

Dados los siguientes lenguajes:

L_3 = { x / x ∈ {a,b,#}* es de la forma $a^{2k+j}\#b^{k+j+p}\#a^{p+1}$; k,j>0, p≥0 }

L_4 = { x / x ∈ {a,b,#}* es de la forma $a^{2k+j}\#b^{p}\#a^{r+1}$; k,j,r>0, p≥0 }

a) Construya gramáticas simplificadas G_3 / $L(G_3) = L_3$ y G_4 / $L(G_4) = L_4$. Justifique porque ambas están simplificadas.
b) Construya un autómata M_3 / $L(M_3) = L_3$. ¿Es determinista? Justifique.
c) ¿Puede afirmar que L_3 y L_4 son lenguajes libre de contexto? ¿Y recursivamente enumerables? Justifique.

Solución

a) Construya gramáticas simplificadas G_3 / $L(G_3) = L_3$ y G_4 / $L(G_4) = L_4$. Justifique porque ambas están simplificadas.

L_3 es un lenguaje libre de contexto no regular.
Se dará una gramática independiente de contexto simplificada G_3 que genera las tiras de L_3.

Observar que las tiras de L_3 pueden expresarse como $x = a^{2k}a^{j}\#b^{j}b^{k}\ b^{p}\#a^{p}a$, con k,j >0 y p≥0.

Esto sugiere la gramática $G_3 = (\{S,S_1,S_2,S_3\}, \{a,b,\#\},P,S)$, con P formado por las siguientes producciones:

$S \rightarrow S_1S_2a$
$S_1 \rightarrow aaS_1b \mid aaS_3b$
$S_3 \rightarrow aS_3b \mid a\#b$
$S_2 \rightarrow bS_2a \mid \#$

G_3 está simplificada porque:
1) No tiene producciones-ε
2) No tiene producciones unitarias
3) Todas las variables son alcanzables desde S (S_1 y S_2 a través de $S \rightarrow S_1 S_2a$, y S_3 a partir de $S_1 \rightarrow aaS_3b$ con S_1 alcanzable)
4) Todas las variables son positivas. S_2 y S_3 directamente a través de las producciones
$S_2 \rightarrow \#$ y $S_3 \rightarrow a\#b$.
$S_1 \Rightarrow aaS_3b$ con S_3 positiva
$S \Rightarrow S_1S_2a$ con S_1 y S_3 positivas

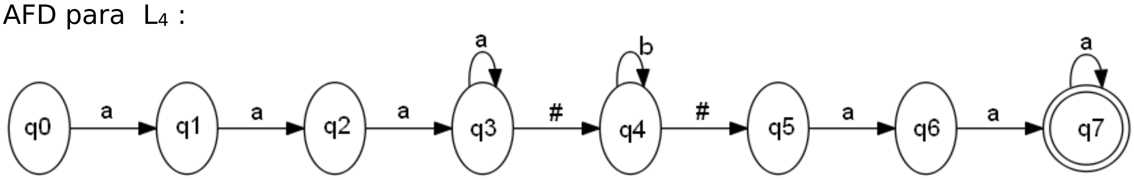
Otra posibilidad es escribir las tiras de L_3 como $x = a^{j}a^{2k}\#b^{k}b^{j}\ b^{p}\#a^{p}a$, con k,j >0 y p≥0.
Las producciones de P en este caso podrían ser:

$S \rightarrow S_1S_2a$
$S_1 \rightarrow aS_1b \mid aS_3b$
$S_3 \rightarrow aaS_3b \mid aa\#b$
$S_2 \rightarrow bS_2a \mid \#$

Nota: Las gramáticas y los autómatas deben corresponderse con el tipo del lenguaje considerado en cada caso, según la Jerarquía de Chomsky. Se valora positivamente la simplicidad de las soluciones propuestas así como una breve explicación de éstas. Todas las respuestas deben estar debidamente justificadas.

L_4 es un lenguaje regular, siendo $L_4 = L(r)$ con $r = aaaa^*\#b^*\#aaa^*$.

AFD para L_4 :



A partir de este AFD y del algoritmo correspondiente que identifica estados con variables y transiciones con producciones se construye la gramática lineal derecha G_4 / $L_4 = L(G_4)$.

$G_4 = (\{S,A,B,C,D\}, \{a,b,\#\},P,S\}$ con P formado por las siguientes producciones:

S → aaaA
A → aA | #B
B → bB | #C
C → aD
D → aD | a

G_4 está simplificada porque:

1) No tiene producciones-ε
2) No tiene producciones unitarias
3) Todas las variables son alcanzables desde S (por ser construida a partir de un AFD en el que todos los estados son alcanzables desde el estado inicial)
4) Todas las variables son positivas (porque desde cada estado se alcanza el estado de aceptación del AFD)

Una GLD equivalente con menos producciones y variables y también simplificada puede ser:

S → aS | aaa#A
A → bA | #aB
B → aB | a

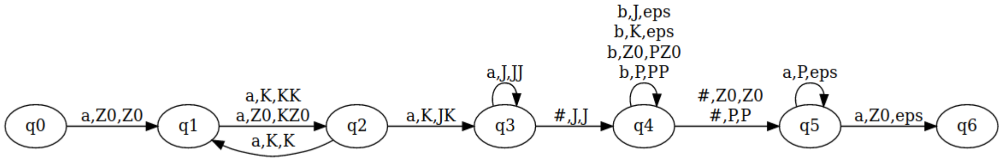
Una gramática lineal izquierda alternativa puede ser la siguiente:

S → Sa | A#aa
A → Ab | B#
B → Ba | aaa

b) Construya un autómata M_3 / $L(M_3) = L_3$. ¿Es determinista? Justifique.

Las tiras del lenguaje las podemos ver también con la forma

$a^{2k}a^j\#b^{k+j}b^p\#a^pa$; $k,j>0$, $p\geq 0$



El autómata no es determinista porque, por ejemplo, desde el estado q_2 tenemos dos transiciones posibles asociadas al símbolo ‘a’ con tope de stack ‘K’.

c) ¿Puede afirmar que L_3 y L_4 son lenguajes libre de contexto? ¿Y recursivamente enumerables? Justifique.

L_3 es Libre de Contexto porque se escribió una Gramática Libre de Contexto que lo genera. Por Jerarquía de Chomsky, es Recursivamente Enumerable.

Por su parte L_4, como es Regular (al escribirse una Gramática Lineal) es – por el Teorema de la Jerarquía de Chomky – Libre de Contexto y nuevamente por la Jerarquía de Chomky también Recursivamente Enumerable.