

Memoria Prácticas

- Añadida Práctica 7.
- Añadidos ejemplos con LEX en la Práctica 4.
- He usado JFLAP para la realización y comprobación de algunos ejercicios de esta práctica
- El resto de prácticas fueron revisadas en clase

Francisco Javier Navarro Morales

26 de enero de 2016

Índice

1. Practica 1: $G = (V, T, P, S)$, donde $V = S, A, B$, $T = a, b$, el símbolo de partida es S y las reglas son: 3
2. Practica 2: Determinar si la gramática $G = (\{S, A, B\}, \{a, b, c, d\}, P, S)$ genera un lenguaje de tipo 3 donde P es el conjunto de reglas de producción: 5
3. Practica 3: Diseñar un autómata finito determinístico que acepte cadenas que contienen las subcadenas '0110' y '1000'. Las subcadenas pueden aparecer juntas o separadas, también pueden contener más símbolos $(0+1)$ delante o detrás de ellas. 7
4. Practica 4: Realizar programas usando la herramienta de reconocimiento de expresiones regulares LEX 9
5. Practica 5: Dado $L = \{ 0 \cup 1 \mid U \in \{0,1\}^* \}$ obtener: 13
6. Practica 6: Elegir una gramática libre del contexto que contenga producciones nulas e unitarias. Aplicar transformaciones para dejarlas en Forma Normal de Chomsky. Opcionalmente pasar a Forma Normal de Greibach. 16
7. Practica 7: Obtener la gramática a partir de un autómata con pila 19

1. Practica 1: $G = (V, T, P, S)$, donde $V = S, A, B$, $T = a, b$, el símbolo de partida es S y las reglas son:

$$\begin{array}{llll} S \rightarrow aB, & S \rightarrow bA, & A \rightarrow a, & A \rightarrow aS, \\ A \rightarrow bAA, & B \rightarrow b, & B \rightarrow bS, & B \rightarrow aBB \end{array}$$

Esta gramática genera el lenguaje: $L(G) = \{u | u \in \{a, b\}^+ \text{ y } Na(u) = Nb(u)\}$, es decir, la gramática genera palabras con el mismo número de 'a' que de 'b'.

Podemos extraer las siguientes interpretaciones de las reglas de producción:

- Interpretación de '**A**' \rightarrow Genera palabras con un Símbolo Terminal 'a' de más.
- Interpretación de '**B**' \rightarrow Genera palabras con un Símbolo Terminal 'b' de más.
- Interpretación de '**S**' \rightarrow Genera una cadena con el mismo numero de 'a' que 'b'.

Hay que demostrar dos cosas:

- Todas las palabras generadas por la gramática tienen el mismo número de a que de b.
- Cualquier palabra con el mismo número de a que de b es generada.

Vamos a ir desarrollando cada posibilidad de forma que para cada paso se apliquen todas las reglas de producción posibles, inicialmente tenemos dos posibilidades:

$$^1S \Rightarrow aB, \quad ^2S \Rightarrow bA$$

Vamos a iniciar el desarrollo para la primera:

$$\begin{array}{l} S \Rightarrow aB \Rightarrow ab, \text{ generamos la palabra } \mathbf{ab}. \\ S \Rightarrow aB \Rightarrow abS \Rightarrow abaB \Rightarrow abab, \text{ generamos la palabra } \mathbf{abab}. \\ S \Rightarrow aB \Rightarrow aabb \Rightarrow aabB \Rightarrow aabb, \text{ generamos la palabra } \mathbf{aabb}. \end{array}$$

Continuamos el desarrollo para la segunda posibilidad:

$$\begin{array}{l} S \Rightarrow bA \Rightarrow ba, \text{ generamos la palabra } \mathbf{ba}. \\ S \Rightarrow bA \Rightarrow baS \Rightarrow babA \Rightarrow baba, \text{ generamos la palabra } \mathbf{baba}. \\ S \Rightarrow bA \Rightarrow bbaA \Rightarrow bbaA \Rightarrow bbaa, \text{ generamos la palabra } \mathbf{bbaa}. \end{array}$$

Hemos comprobado que podemos conseguir generar cadenas básicas con $N_a(u) = N_b(u)$, dónde primero hay símbolos terminales 'a' y luego 'b' ó hay símbolos terminales '(ab)⁺', y lo mismo cambiando el orden de 'a' y 'b'. Si queremos generar cualquier cadena perteneciente al lenguaje lo podemos conseguir combinando las anteriores palabras generadas. Por ejemplo generemos una palabra usando todas las reglas de producción:

$$\begin{array}{l} S \xrightarrow{1} aB \xrightarrow{8} aaBB \xrightarrow{7} aabSB \xrightarrow{1} aabaBB \xrightarrow{8} aabaaBBB \xrightarrow{7} aabaabSBB \xrightarrow{2} aabaabbABB \\ \xrightarrow{5} aabaabbbAABB \xrightarrow{3} aabaabbbaABB \xrightarrow{4} aabaabbbbaaSBB \xrightarrow{2} aabaabbbaabABB \xrightarrow{3} \\ aabaabbbaabaBB \xrightarrow{6} aabaabbbaababB \xrightarrow{6} aabaabbbaababb \end{array}$$

Podemos apreciar que encima de cada fecha indicamos el número de la regla utilizada (las identificamos contando de izquierda a derecha y de arriba hacia abajo) en este ejemplo hemos usado las 8 reglas de producción que nos brinda la gramática, comprobando que la palabra que genera **'aabaabbbbaababb'** contiene el mismo número de símbolos terminales 'a' que 'b', en concreto $N = 7$. Con este ejemplo demostramos también que combinando las reglas de producción podemos generar cualquier palabra con la peculiaridad de que el número de 'a' coincide con el de 'b'.

2. Practica 2: Determinar si la gramática $G = (\{S, A, B\}, \{a, b, c, d\}, P, S)$ genera un lenguaje de tipo 3 donde P es el conjunto de reglas de producción:

$$S \rightarrow AB, \quad A \rightarrow Ab, \quad A \rightarrow a, \quad B \rightarrow cB, \quad B \rightarrow d$$

Esta gramática genera el lenguaje $L(G) = \{ab^i c^j d : i, j \in \mathbb{N}\}$

Partiendo de los datos del problema vamos a comprobar que esa gramática genera todas las palabras del lenguaje.

$$S \rightarrow \underline{AB} \rightarrow \underline{Ab}B \rightarrow \underline{Abb}B \rightarrow ab^i \underline{B} \rightarrow ab^i c \underline{B} \rightarrow ab^i cc \underline{B} \rightarrow ab^i c^j \underline{B} \rightarrow ab^i c^j d$$

Vemos que todas las palabras que genera dicha gramática pertenecen al lenguaje y no genera ninguna que no pertenezca al lenguaje. Podemos sacar la siguiente interpretación:

- A -> genera subcadenas que empiezan por 'a' seguido de un numero i de 'b'.
- B -> genera subcadenas que empiezan por un numero j de 'c' y acaban en 'd'.
- Estamos ante una gramática libre del contexto (de tipo 2) puesto que sus producciones son de la forma "terminal-Variable y Variable-terminal" ó "Variable-Variable". Por lo tanto el lenguaje que genera es también de tipo 2.

Hasta ahora no hemos hecho mas que comprobar que los datos del problema son correctos, el siguiente paso es ver si conseguimos modificar la gramática, en concreto sus producciones, para conseguir que esta genere un lenguaje regular. Podemos intentar esto porque no hay relación numérica entre los símbolos terminales del lenguaje.

Según la **Jerarquía de Chomsky** un lenguaje es regular si las reglas de producción son de tipo 3, para esto las producciones de la gramática deben ser del tipo: $A \rightarrow uB$ || $A \rightarrow u$.

Vamos a usar las siguientes reglas de producción:

$$S \rightarrow aB, \quad B \rightarrow bB, \quad B \rightarrow C, \quad C \rightarrow cC, \quad C \rightarrow d$$

- $S \rightarrow aB$: S es el símbolo inicial. Esta prod. genera la primera "a" y da paso a la variable "B".
- $B \rightarrow bB$: Generar un numero i de "b".
- $B \rightarrow C$: Pasa a la variable "C".

- $C \rightarrow cC$: Genera un número j de "c".
- $C \rightarrow d$: Para finalizar esta producción nos permite colocar el último símbolo terminal "d" necesario para formar palabras correctas.

En conclusión hemos encontrado unas reglas de producción de tipo 3, que hacen la gramática regular, capaces de generar el mismo lenguaje que generaban las reglas de producción de partida. **Demostramos que la nueva gramática es de tipo 3 por lo tanto el lenguaje también lo es.**

3. Practica 3: Diseñar un autómata finito determinístico que acepte cadenas que contienen las subcadenas '0110' y '1000'. Las subcadenas pueden aparecer juntas o separadas, también pueden contener más símbolos (0+1) delante o detrás de ellas.

Para comenzar esta práctica lo más sencillo es **diseñar el autómata no determinístico** que acepta las cadenas que nos pide el ejercicio. Nos ayudamos de los ejemplos de los autómatas que son capaces de reconocer palabras que contienen las subcadenas '0110' ó '1000, de esta forma conseguimos diseñar el autómata que necesitamos de forma muy fácil, para ello **seguimos los siguientes pasos**:

- Q4 deja de ser un nodo final y P0 deja de ser un nodo inicial.
- Unimos los dos autómatas añadiendo una transición nula entre Q4 y P0.

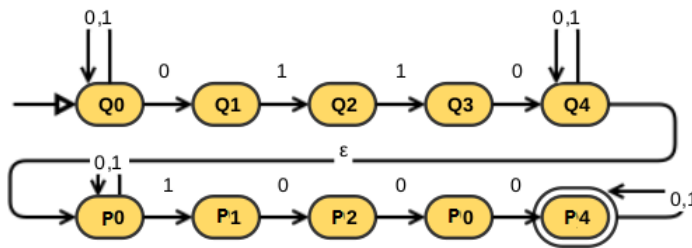


Figura 3.1: Autómata Finito No Determinístico(AFND)

Los autómatas no determinísticos son muy ineficientes porque exploran todas las posibilidades. Para resolver el problema que tenemos entre las manos de una forma más eficiente vamos a transformarlo en un autómata que para cada símbolo leído sepa que acción debe realizar. Es un proceso exponencial a medida en que crece el número de nodos en el 'AFND', para no equivocarnos vamos a realizar una tabla donde indiquemos para cada nodo cual es la acción que realiza al leer cierto símbolo de la cinta de entrada.

Símbolo	Q0	Q1	Q2	Q3	Q4
0	{Q0,Q1}	{ \emptyset }	{ \emptyset }	{Q4,P0}	{Q4,P0}
1	{Q0}	{Q2}	{Q3}	{ \emptyset }	{Q4,P0,P1}

Tabla 3.1: Tabla nodos Q.

4. Practica 4: Realizar programas usando la herramienta de reconocimiento de expresiones regulares LEX

Este programa de ejemplo es capaz de reconocer números (tanto reales como enteros) e identificadores formados por al menos un carácter seguido de más caracteres o dígitos. En el caso de los enteros también lleva la suma de estos que al final muestra junto al número de reales, enteros e identificadores reconocidos.

Escribe el siguiente código en un fichero llamado `ejemplo.c`:

```
1  /*Definir Variables Globales*/
2  %{
3  int ent=0, real=0, ident=0, sumaent=0, i=0;
4  %}
5  /*Definir Expresiones Regulares*/
6  car      [a-zA-Z]
7  digito   [0-9]
8  signo    (\-|\+ )
9  suc      ({digito}+)
10 enter    ({signo}?{suc})
11 real1    ({enter}\.{digito}*)
12 real2    ({signo}?\.{suc})
13 /*Definir Acciones*/
14 %%
15 {enter}      {ent++; sscanf(yytext,"%d",&i); sumaent +=
               i; printf("Numero entero: %s\n\n",yytext);}
16
17 ({real1}|{real2}) {real++; printf("Num. real: %s\n\n",yytext
               );}
18
19 {car}({car}|{digito})* {ident++; printf("Var. ident: %s\n\n",
               yytext);}
20
21 .|\n          {;}
22 %%
23 yywrap()
24 {printf("Numero de Enteros: %d, reales: %d, ident: %d,
25 Suma de Enteros: %d",ent,real,ident,sumaent); return 1;}
```

Ahora compila usando `lex` y `gcc`:

```
$ lex ejemplo
$ gcc lex.yy.c -o prog -ll
```

Por último ejecutamos de la siguiente forma:

```
$ ./prog < 'fichero_entrada' > 'fichero_salida'
```

Implementación del autómata de la "Practica 3: Diseñar un autómata finito determinístico que acepte cadenas que contienen las subcadenas '0110' y '1000'":

```

1  /*Definicion de variables, includes, etc*/
2  %{
3  int i=0,t=0;
4  char id[]=" ";
5  %}
6
7  /*Definicion de expresiones regulares*/
8  car      [a-zA-Z]
9  digito   [0-9]
10 patron  (0|1)*0110(0|1)*1000(0|1)*
11
12 /*Cuando encuentre una expresion regular hace la funcion
    correspondiente*/
13 %%
14
15 {car}({car}|{digito})* {t++; sscanf(yytext,"%s",id);}
16 {patron}               {i++; printf("\nCadena con id %s -> %s es
    una cadena valida.",id, yytext);}
17 .|\n {;}
18
19 %%
20
21 /*Cuando acabe de leer el fichero se ejecuta la seccion yylex (
    main) ejecuta la funcion yywrap que muestra la informacion*/
22 yywrap(){
23     i?printf("***Hay %d/%d aceptadas***\n",i,t):printf("\nNo
    hay cadena(s) valida(s)\n");
24     return 1;
25 }

```

```

navarro@navarro:~/Escritorio/++Tercer año Grado/MC/practica/Practica4/Ejercicio 3$ lex practica3.c && gcc lex.yy.c -o practica3 -ll
navarro@navarro:~/Escritorio/++Tercer año Grado/MC/practica/Practica4/Ejercicio 3$ ./practica3 < cadenaP3.txt > salida_practica3.txt
navarro@navarro:~/Escritorio/++Tercer año Grado/MC/practica/Practica4/Ejercicio 3$ cat cadenaP3.txt salida_practica3.txt
CAD1: 00011010110100110010101001101000
CAD2: 00000001111000111100001111000000
CAD3: 01110110110111111111110000001000
CAD4: 00000000011000000000100000000000
CAD5: 00000000000000000000000000000000
CAD6: 11111111111111111111111111111111
CAD7: 0010000000110
CAD8: 0110100000000
Cadena con id CAD1 -> 00011010110100110010101001101000 es una cadena valida.
Cadena con id CAD3 -> 01110110110111111111110000001000 es una cadena valida.
Cadena con id CAD4 -> 00000000011000000000100000000000 es una cadena valida.
Cadena con id CAD8 -> 0110100000000 es una cadena valida.***Hay 4/8 aceptadas***

```

Figura 4.1: Captura de ejecución del programa

Implementación de un reconocedor para saber si un email tiene un formato correcto (nombre @ dominio) :

```

1  /*Definicion de variables, includes, etc*/
2  %{
3  int ent=0;
4  %}
5
6  /*Definicion de expresiones regulares*/
7  car      [a-zA-Z]
8  digito   [0-9]
9  signo    (\-)
10 iden     ({car}|{digito})(([_\.\-]?({car}|{digito}))+*)
11 dominio  @([A-Za-z0-9]+)(([_\.\-]?[a-zA-Z0-9]+)*)\.( [A-Za-z]{2,})
12 email    ^{iden}{dominio}$
13
14 /*Cuando encuentre una expresion regular hace la funcion
    correspondiente*/
15 %%
16
17 {email}      {ent++; printf("\nEl email %s tiene un formato
    valido(texto@dominio) \n", yytext);}
18
19 .|\n {;}
20
21 %%
22
23 /*Cuando acabe de leer el fichero se ejecuta la seccion yylex (
    main) ejecuta la funcion yywrap que muestra la informacion*/
24 yywrap(){
25     printf("\nNumero de email validos: %d\n %s",ent,yytext);
26     return 1;
27 }

```

```

navarro@navarro:~/Escritorio/++Tercer año Grado/MC/practica/Practica4/Comprobar mails$ lex comprobar_email.c && gcc lex.yy.c -o emails -ll
navarro@navarro:~/Escritorio/++Tercer año Grado/MC/practica/Practica4/Comprobar mails$ ./emails < emails.txt > salida_emails.txt
navarro@navarro:~/Escritorio/++Tercer año Grado/MC/practica/Practica4/Comprobar mails$ cat emails.txt salida_emails.txt
hola@gmail.com
prueba@ugr.es
miguel.45.com
test.45@test.com.es
perico@hotmail
El email hola@gmail.com tiene un formato valido(texto@dominio)

El email prueba@ugr.es tiene un formato valido(texto@dominio)

El email test.45@test.com.es tiene un formato valido(texto@dominio)

Numero de email validos: 3

```

Figura 4.2: Captura de ejecución del programa

Implementación de sumador de dos números romanos :

```
1 WS  [ \t]+
2
3 %%
4 int total=0;
5
6 I    total += 1;
7 IV   total += 4;
8 V    total += 5;
9 IX   total += 9;
10 X    total += 10;
11 XL   total += 40;
12 L    total += 50;
13 XC   total += 90;
14 C    total += 100;
15 CD   total += 400;
16 D    total += 500;
17 CM   total += 900;
18 M    total += 1000;
19
20 {WS}    |
21 \n    return total;
22 %%
23 int main (void) {
24     int first, second;
25
26     first = yylex ();
27     second = yylex ();
28
29     printf ("%d + %d = %d\n", first, second, first+second);
30     return 0;
31 }
```

```
navarro@navarro:~/Escritorio/++Tercer año Grado/MC/practica/Practica4/Numero Romano$ lex -f contador_numeros_romanos.c && gcc lex.yy.c -o sumador-romanos -ll
navarro@navarro:~/Escritorio/++Tercer año Grado/MC/practica/Practica4/Numero Romano$ ./sumador-romanos
XX
IV
20 + 4 = 24
```

Figura 4.3: Captura de ejecución del programa

5. Practica 5: Dado $L = \{ 0 \cup 1 \mid U \in \{0,1\}^* \}$ obtener:

1. Expresión Regular
2. Autómata Finito Determinístico
3. Gramática lineal por la derecha y por la izquierda

1) La expresión regular correspondiente al lenguaje dado es ' $0(0+1)^*1$ '.

2) Para obtener el Autómata Finito Determinístico (AFD) nos ayudamos de la expresión regular que tenemos formada, de ella podemos extraer que necesitamos reconocer cadenas que comiencen por un 0 seguida de n veces 0 ò 1 (con $n \geq 0$) y acaben en 1.

Vamos a dar un paso antes de obtener el AFD, se trata de obtener un autómata que nos resuelva el problema aunque este no sea determinístico.

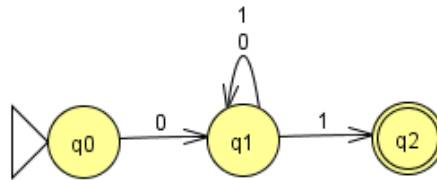


Figura 5.1: Autómata Finito NO Determinístico (AFND)

Y ahora como aprendimos en la práctica3 pasaremos este 'AFND' a un 'AFD'.

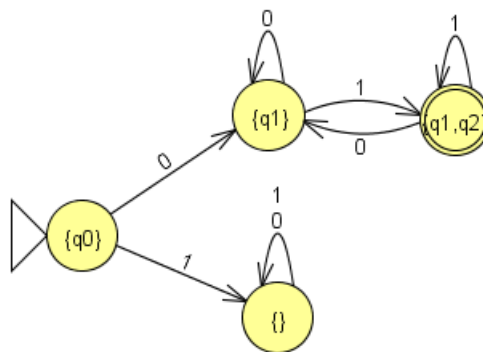


Figura 5.2: Autómata Finito Determinístico (AFD) *{}:Estado Error*

3) Siguiendo los pasos dados por el problema podemos generar una gramática lineal por la derecha a partir del autómata determinístico:

- Cada estado determina una variable de la gramática.

- Para cada transición entre estados añadimos una producción. En la parte izquierda ponemos el estado añadiendo en la parte derecha el símbolo leído + el estado al que se dirige.
- Añadir una producción $V \rightarrow \lambda$, para cada V siendo este estado final.

La gramática asociada a este autómata es (q_0 -inicial):

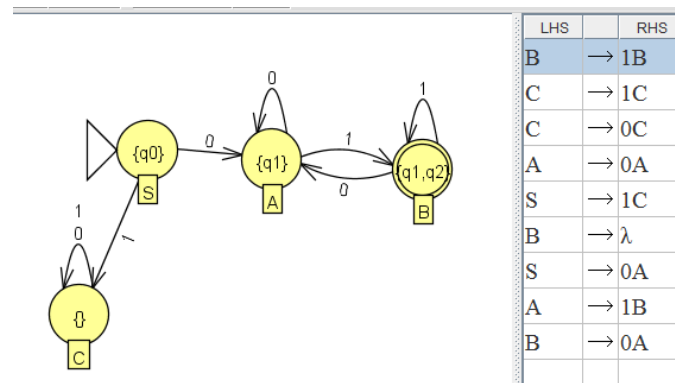


Figura 5.3: AFD - Gramática Lineal por la derecha

Etiquetamos los nodos con una letra para generar una gramática de forma más clara

Llegados hasta este paso solo nos falta obtener la gramática lineal por la izquierda, para ello invertimos el autómata (cambiando la dirección de las transiciones entre estados y cambiamos los estados finales por los iniciales), construimos la gramática lineal por la derecha asociada e invertimos la parte derecha de las producciones.

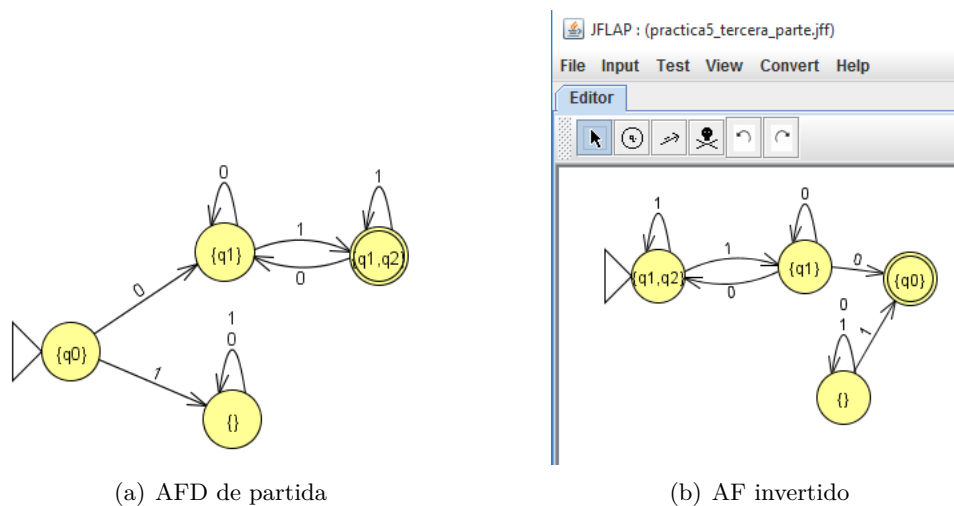
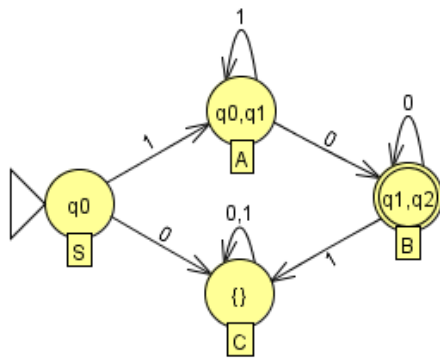


Figura 5.4: Pasando GLD a GLI

Al invertir el diagrama de la figura (a) obtenemos el de la figura (b) resultante que es un diagrama no determinístico, por lo cual necesitamos pasarlo a determinístico. Renombramos los estados, pasamos a determinístico y obtenemos la gramática lineal por la derecha asociada al AFD, por último giramos la parte derecha de las producciones, resultado:



(a) AFD de partida

JFLAP : <untitled2>

File Input Test Convert Help		
Editor		
Table Text Size		
LHS		
S	→	C0
S	→	A1
A	→	B0
A	→	A1
B	→	B0
B	→	C1
C	→	C0
C	→	C1
B	→	λ

(b) AFD de partida

Figura 5.5: AF – Gramática Lineal Izquierda

6. Practica 6: Elegir una gramática libre del contexto que contenga producciones nulas e unitarias. Aplicar transformaciones para dejarlas en Forma Normal de Chomsky. Opcionalmente pasar a Forma Normal de Greibach.

Sea la gramática:

$$\begin{aligned} S \rightarrow SS, \quad S \rightarrow CA, \quad A \rightarrow bAA, \quad A \rightarrow aC, \quad A \rightarrow B, \\ B \rightarrow aSS, \quad B \rightarrow BC, \quad C \rightarrow CC, \quad C \rightarrow \varepsilon \end{aligned}$$

El primer paso para normalizar la gramática es eliminar las producciones nulas siguiendo los siguientes pasos:

1. Marcar (para eliminar) todas las producciones nulas del tipo $A \rightarrow \varepsilon$.
2. Para cada producción que en su parte derecha tuviese como mínimo una de las producciones que hemos marcado, tenemos que añadir todas las producciones que podríamos conseguir usando la producción eliminada.
3. Eliminar todas las producciones que marcamos en el primer paso.

Aplicando el primer paso del algoritmo..

$$\begin{aligned} S \rightarrow SS, \quad S \rightarrow CA, \quad A \rightarrow bAA, \quad A \rightarrow aC, \quad A \rightarrow B, \\ B \rightarrow aSS, \quad B \rightarrow BC, \quad C \rightarrow CC, \quad C \rightarrow \varepsilon \end{aligned}$$

Aplicando el segundo paso del algoritmo..

$$\begin{aligned} S \rightarrow SS, \quad S \rightarrow CA, \quad A \rightarrow bAA, \quad A \rightarrow aC, \quad A \rightarrow B, \\ B \rightarrow aSS, \quad B \rightarrow BC, \quad C \rightarrow CC, \quad C \rightarrow \varepsilon, \\ S \rightarrow A, \quad A \rightarrow a, \quad B \rightarrow B, \quad C \rightarrow C \end{aligned}$$

Por último aplicando el tercer paso del algoritmo obtenemos una gramática equivalente sin producciones nulas:

Gramática sin producciones nulas

$$\begin{aligned} S \rightarrow SS, \quad S \rightarrow CA, \quad A \rightarrow bAA, \quad A \rightarrow aC, \\ A \rightarrow B, \quad B \rightarrow aSS, \quad B \rightarrow BC, \quad C \rightarrow CC, \\ S \rightarrow A, \quad A \rightarrow a, \quad B \rightarrow B, \quad C \rightarrow C \end{aligned}$$

Seguimos procesando la misma gramática para llegar a nuestro objetivo final. Ahora queremos eliminar las transiciones unitarias, siguiendo los siguientes pasos:

1. Nos ayudamos de un conjunto H donde introduciremos para de valores siguiendo los siguientes criterios:
 - a) Para cada producción del tipo $A \rightarrow B$, añadimos en H el par (A,B).
 - b) Si en H encontramos pares del tipo (S,A)-(A,B), añadimos el par (S,B) en H.
2. Eliminamos las producciones unitarias.
3. Para cada par en H realizamos el siguiente proceso, suponiendo que tenemos el par (A,B) tenemos que añadir una producción por cada producción del tipo $B \rightarrow$, tal que $A \rightarrow$ <parte derecha de las producciones que sean del tipo $B \rightarrow$. Consiguiendo generar por A las producciones que se generaban desde B.

Aplicando el algoritmo..

$$\begin{array}{l} S \rightarrow SS, \quad S \rightarrow CA, \quad A \rightarrow bAA, \\ A \rightarrow aC, \quad A \rightarrow B, \quad B \rightarrow aSS, \\ B \rightarrow BC, \quad C \rightarrow CC, \quad S \rightarrow A, \\ A \rightarrow a, \quad B \rightarrow B, \quad C \rightarrow C \end{array}$$

$$H = (A,B), (S,A), (B,B), (C,C), (S,B)$$

Quitar $A \rightarrow B$ y añadir:

- $A \rightarrow aSS, \quad A \rightarrow BC$

Quitar $S \rightarrow B$ y añadir:

- $S \rightarrow aSS, \quad S \rightarrow BC$

Quitar $S \rightarrow A$ y añadir:

- $S \rightarrow bAA, \quad S \rightarrow aC, \quad S \rightarrow a$

Para $B \rightarrow B$ y $C \rightarrow C$ no hay producciones que añadir.

Gramática sin producciones nulas, ni unitarias

$$\begin{array}{l} S \rightarrow SS, \quad S \rightarrow CA, \quad S \rightarrow bAA, \quad S \rightarrow aC, \quad S \rightarrow a, \\ S \rightarrow aSS, \quad S \rightarrow BC, \quad A \rightarrow bAA, \quad A \rightarrow aC, \quad A \rightarrow a, \\ A \rightarrow aSS, \quad A \rightarrow BC, \quad B \rightarrow aSS, \quad B \rightarrow BC, \quad C \rightarrow CC \end{array}$$

Pasar a Chomsky:

Sea la gramática:

$$\begin{array}{l} \underline{S \rightarrow SS}, \quad S \rightarrow bAA, \quad \underline{S \rightarrow a}, \quad S \rightarrow aSS, \\ A \rightarrow bAA, \quad \underline{A \rightarrow a}, \quad A \rightarrow aSS, \end{array}$$

Las que aparecen subrayadas están ya en Forma Normal de Chomsky

El resto de producciones las tenemos que transformar hasta que sean de la forma:

- $A \rightarrow a\alpha$, donde a es un símbolo terminal y α pertenece a V^*
- $A \rightarrow \alpha$, donde α pertenece a V^* siempre que $|\alpha| \geq 2$

Sustituimos $S \rightarrow bAA$ por: Sustituimos $A \rightarrow bAA$ por: Añadir:

- $S \rightarrow C_b AA$
- $A \rightarrow C_b AA$
- $\underline{C_b \rightarrow b}$

Sustituimos $S \rightarrow aSS$ por: Sustituimos $A \rightarrow aSS$ por: Añadir:

- $S \rightarrow C_a SS$
- $A \rightarrow C_a SS$
- $\underline{C_a \rightarrow a}$

Continuamos con el refinamiento:

Sustituimos $S \rightarrow C_b AA$ por: Sustituimos $A \rightarrow C_b AA$ por: Añadir:

- $S \rightarrow C_b D_1$
- $A \rightarrow C_b D_1$
- $D_1 \rightarrow AA$

Sustituimos $S \rightarrow C_a SS$ por: Sustituimos $A \rightarrow C_a SS$ por: Añadir:

- $S \rightarrow C_a D_2$
- $A \rightarrow C_a D_2$
- $D_2 \rightarrow SS$

Gramática en Forma Normal de Chomsky

$$\begin{array}{l} S \rightarrow SS, \quad S \rightarrow a, \quad A \rightarrow a, \quad S \rightarrow C_b D_1, \quad S \rightarrow C_a D_2, \\ A \rightarrow C_b D_1, \quad A \rightarrow C_a D_2, \quad C_a \rightarrow a, \quad C_b \rightarrow b, \\ D_1 \rightarrow AA, \quad D_2 \rightarrow SS, \end{array}$$

7. Practica 7: Obtener la gramática a partir de un autómata con pila

Por pila vacía. $M = (q1, q2, 0, 1, R, X, \delta, q1, R, \theta)$

$$\delta(q1, 0, R) = (q1, XR) \quad \delta(q1, 0, X) = (q1, XX)$$

$$\delta(q1, \epsilon, R) = (q1, \epsilon) \quad \delta(q1, 1, X) = (q2, \epsilon)$$

$$\delta(q2, 1, X) = (q2, \epsilon) \quad \delta(q2, \epsilon, R) = (q2, \epsilon)$$

La interpretación de los estados es la siguiente:

- En el estado $q1$ leemos 0 de la cinta de entrada escribiendo una X en la pila.
- En el estado $q2$ leemos 1 haciendo matching con las X de la pila.

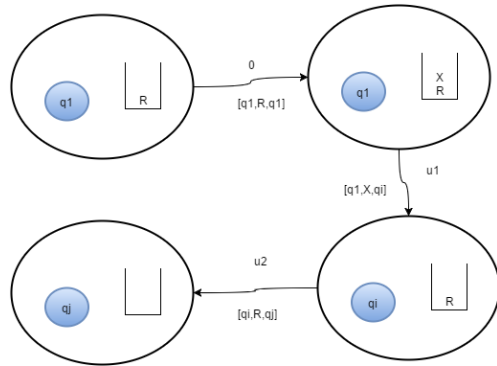
Ya estamos listos para comenzar a extraer las producciones de la gramática.

En primer paso añadimos una transición del tipo $S \rightarrow [q0, z0, q]$ siendo $q0$ el estado inicial, $z0$ el símbolo inicial de la pila y $q \in Q$.

$$S \rightarrow [q1, R, q1] \quad S \rightarrow [q1, R, q2]$$

En segundo paso para cada transición que lea y escriba en la pila tenemos que añadir una producción por cada estado hasta dejar la pila vacía. Este paso queda más claro con el siguiente ejemplo:

$$\underline{\delta(q1, 0, R) = (q1, XR)}$$



$$\begin{aligned} [q1, R, q1] &\rightarrow 0[q1, X, q1][q1, R, q1] \\ [q1, R, q1] &\rightarrow 0[q1, X, q2][q2, R, q1] \\ [q1, R, q2] &\rightarrow 0[q1, X, q1][q1, R, q2] \\ [q1, R, q2] &\rightarrow 0[q1, X, q2][q1, R, q2] \end{aligned}$$

$$\underline{\delta(q1, 0, X) = (q1, XX)}$$

$$\begin{aligned} [q1, X, q1] &\rightarrow 0[q1, X, q1][q1, X, q1] \\ [q1, R, q1] &\rightarrow 0[q1, X, q2][q2, X, q1] \\ [q1, R, q2] &\rightarrow 0[q1, X, q1][q1, X, q2] \\ [q1, R, q2] &\rightarrow 0[q1, X, q2][q1, X, q2] \end{aligned}$$

$$\underline{\delta(q1, \epsilon, R) = (q1, \epsilon)}$$

$$[q1, R, q1] \rightarrow \epsilon$$

$$\underline{\delta(q2, \epsilon, R) = (q2, \epsilon)}$$

$$[q2, R, q2] \rightarrow \epsilon$$

$$\underline{\delta(q2,1,X) = (q2, \epsilon)}$$

$$[q2,X,q2] \rightarrow 1$$

$$\underline{\delta(q1,1,X) = (q2, \epsilon)}$$

$$[q1,X,q2] \rightarrow 1$$

Y de esta forma obtenemos la gramática asociada al lenguaje $L = 0^i 1^i : i \geq 0$