Tecnológico de Costa Rica

Compiladores e Interpretes

Apuntes de clase

Profesor Francisco Torres

Edwin Cen 2014055617

15 marzo 2017

Apuntes del miércoles 15 de marzo 2017

Quiz de asistencia!!!!

Se hizo un breve repaso de la clase del 8 de marzo, quedamos en Σ^+.

Continuación de Análisis Léxico

Reverso de una Hilera

Sea w en Σ^*

La reversa de w, denotada como w^R o w^{-1} es la hilera w escrita al revés

Ejemplo:

u = abc, v = def

Se puede demostrar por inducción,

$(uv)^R =$ fedcba

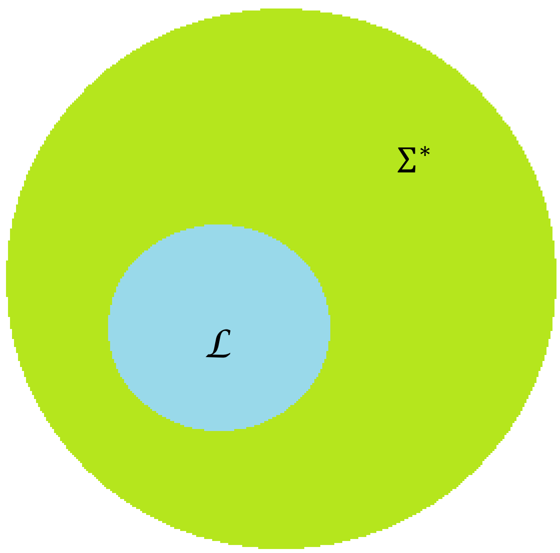
$v^R u^R \quad =$ fedcba

Lenguaje formal

Sea Σ un alfabeto, entonces un lenguaje formal $\mathcal{L}$ sobre Σ es:

- Un conjunto de hileras
- $\mathcal{L}$ es subconjunto de Σ^*

Las hileras cumplen cierta propiedad especifica con reglas precisas



Sea Σ = {0,1}

- = hileras sobre Σ que terminen en 00

Ej. $\mathcal{L} =$ 100, 1100, 00, 000, …

ε no es miembro de este lenguaje porque no termina en 00.

- $\mathcal{L}$ = hileras sobre Σ que empiecen con 1

Ej. $\mathcal{L} =$ 100, 1100, 10100, 1000, …

ε no es miembro de este lenguaje porque no empieza con 1.

- $\mathcal{L}$ = hileras sobre Σ que tengan una longitud impar

Ej. $\mathcal{L} =$ 1, 010, 0, 10001, 101, …

ε no es miembro de este lenguaje ya que la cardinalidad de la hilera es 0 y se considera par.

Sea Σ = {A, U, C, G}

- $\mathcal{L}$ = hileras sobre Σ que no incluyan la subhilera UUU

Ej. $\mathcal{L} =$ CACA, AGUA, CCG,UUG, AUU, …

- $\mathcal{L}$ = hileras sobre Σ cuya longitud sea múltiplo de 3

Ej. $\mathcal{L} =$ GUAGUAGUA, AUG, CGA, CUGAAA, …

- $\mathcal{L}$ = hileras sobre Σ que empiecen con AUG, tengan una longitud múltiplo de 3 y terminen con UAA, UAG o UGA

Ej. $\mathcal{L} =$ AUGCAGUAGUAA, AUGUAG, AUGAGGGAAUAA, …

Continuación de Análisis Léxico…

Unión de lenguajes

Operaciones sobre lenguajes formales, ya sabemos que son, ahora le vamos a hacer cosas a los lenguajes formales.

Unión de lenguajes, por el momento para la unión del lenguaje vamos a utilizar el mismo alfabeto pero si se puede unir lenguajes con diferentes alfabetos según el profe.

Sea $\mathcal{L}$ y $\mathcal{M}$ dos lenguajes sobre Σ

La unión de $\mathcal{L}$ y $\mathcal{M}$, denotado como $\mathcal{L} \cup \mathcal{M}$ es:

$$\mathcal{L} \cup \mathcal{M} = \{ x \mid x \in \mathcal{L} \textbf{ o } x \in \mathcal{M}\}$$

Las reglas de los dos lenguajes se unen.

Unión de lenguajes – Ejemplo

Sea $\mathcal{L} = \{000, 101, 100, 110\}$ y $\mathcal{M} = \{0, 1, 11\}$

$$\mathcal{L} \cup \mathcal{M} = \{000, 101, 100, 110, 0, 1, 11\}$$

#Posdata – En los conjuntos no pueden tener elementos repetidos, nunca lo hagan en quices ni en exámenes. Si no el profe les va a dar latigazos.

Sea:

- $\mathcal{L}$ = hileras sobre Σ que terminen en 00
- $\mathcal{M}$ = hileras sobre Σ que empiecen con 1

$\mathcal{L} \cup \mathcal{M}$ = hileras sobre Σ que empiecen con 1 **o** que terminen en 00

Algunas hileras que cumple las propiedades de $\mathcal{L} \cup \mathcal{M}$ son:

$\mathcal{L} \cup \mathcal{M} = \{1, 00, 100, 000, 11111, 1010101, 00100,\ldots\}$

ε no es miembro de este lenguaje.

Propiedades de la unión de lenguajes

Conmutatividad:

$\mathcal{L} \cup \mathcal{M} = \mathcal{M} \cup \mathcal{L}$

Asociatividad:

$\mathcal{L} \cup \mathcal{M} \cup \mathcal{P} = (\mathcal{L} \cup \mathcal{M}) \cup \mathcal{P} = \mathcal{L} \cup (\mathcal{M} \cup \mathcal{P})$

Idempotencia:

$\mathcal{L} \cup \mathcal{L} = \mathcal{L}$

Elemento neutro:

$\exists \mathcal{N}\ tq. \forall \mathcal{L}, \mathcal{L} \cup \mathcal{N} = \mathcal{L}$

Que es $\mathcal{N}$?

✗ Un lenguaje cuyo único elemento es ε. $\{\varepsilon\}$

✓ Conjunto vacio $\emptyset$

Son diferentes porque $\{\varepsilon\}$ tiene un elemento, mientras que $\emptyset$ no tiene. La cardinalidad de $\{\varepsilon\}$ es 1 y la cardinalidad de $\emptyset$ es 0. Por lo tanto $\mathcal{N}$ es vacio.

Cierre:

Si $\mathcal{L}, \mathcal{M} \subseteq \Sigma^*$ entonces $\mathcal{L} \cup \mathcal{M} \subseteq \Sigma^*$

Intersección de Lenguajes

Sea $\mathcal{L}$ y $\mathcal{M}$ dos lenguajes sobre Σ

La intersección de $\mathcal{L}$ y $\mathcal{M}$, denotado como $\mathcal{L} \cap \mathcal{M}$ es:

$$\mathcal{L} \cap \mathcal{M} = \{ x \mid x \in \mathcal{L} \textbf{ y } x \in \mathcal{M}\}$$

Intersección de lenguajes – Ejemplo

Sea $\mathcal{L} = \{000, 010, 100, 110\}$ y $\mathcal{M} = \{0, 1, 11\}$

$\mathcal{L} \cap \mathcal{M} = \emptyset$

Sea:

- $\mathcal{L}$ = hileras sobre Σ que terminen en 00
- $\mathcal{M}$ = hileras sobre Σ que empiecen con 1

$\mathcal{L} \cap \mathcal{M}$ = hileras sobre Σ que empiecen con 1 **y** que terminen en 00

Ej. $\mathcal{L} \cap \mathcal{M}$ ={100, 1000, 1100, 10000, 10100, 11000, 11100, …}

Las hileras 1, 00, y ε no son miembros de $\mathcal{L} \cap \mathcal{M}$ porque 1 no termina en 00, 00 no empieza en 1, y ε no cumple con ninguno.

Propiedades de la intersección de lenguajes

Conmutatividad:

$\mathcal{L} \cap \mathcal{M} = \mathcal{M} \cap \mathcal{L}$

Asociatividad:

$\mathcal{L} \cap \mathcal{M} \cap \mathcal{P} = (\mathcal{L} \cap \mathcal{M}) \cap \mathcal{P} = \mathcal{L} \cap (\mathcal{M} \cap \mathcal{P})$

Idempotencia:

$\mathcal{L} \cap \mathcal{L} = \mathcal{L}$

$\mathcal{L} \cap \emptyset = \emptyset$

Elemento neutro:

No hay elemento neutro de la intersección

Cierre:

Si $\mathcal{L}, \mathcal{M} \subseteq \Sigma^*$ entonces $\mathcal{L} \cap \mathcal{M} \subseteq \Sigma^*$

Diferencia de Lenguajes

Sea $\mathcal{L}$ y $\mathcal{M}$ dos lenguajes sobre Σ

La diferencia de $\mathcal{L}$ y $\mathcal{M}$, denotado como $\mathcal{L} - \mathcal{M}$ es:

$$\mathcal{L} - \mathcal{M} = \{ x \mid x \in \mathcal{L} \textbf{ y } x \notin \mathcal{M} \}$$

Lenguaje formado por hileras que pertenezcan a $\mathcal{L}$ pero no pertenece a $\mathcal{M}$

Diferencia de lenguajes – Ejemplo

Sean $\mathcal{L} = \{1, 11, 111, 1111, 11111\}$ y $\mathcal{M} = \{0, 1, 11\}$

$\mathcal{L} - \mathcal{M} = \{111, 1111, 11111\}$

Sean:

- $\mathcal{L}$ = hileras sobre Σ que terminen en 00
- $\mathcal{M}$ = hileras sobre Σ que empiecen con 1

$\mathcal{L} - \mathcal{M}$ = hileras sobre Σ que terminen con 00 pero no empiecen con 1

Ej. $\mathcal{L} - \mathcal{M}$ ={00, 01000, 000, 0000, 00100, …}

ε no es miembro de $\mathcal{L} - \mathcal{M}$.

Propiedades de la diferencia de lenguajes

Conmutatividad:

La diferencia de lenguajes no es conmutativa

Asociatividad:

La diferencia de lenguajes no es asociativa

Idempotencia:

La diferencia de lenguajes no es idempotente

Elemento neutro:

$\exists \mathcal{N}$ $tq. \forall \mathcal{L}, \mathcal{L} - \mathcal{N} = \mathcal{L}$

Que es $\mathcal{N}$?

- Conjunto vacio $\emptyset$

Cierre:

Si $\mathcal{L},\mathcal{M} \subseteq \Sigma^*$ entonces $\mathcal{L} - \mathcal{M} \subseteq \Sigma^*$

Complemento de Lenguajes

Sea $\mathcal{L}$ un lenguajes sobre Σ

EL complemento de $\mathcal{L}$, denotado como $\overline{\mathcal{L}}$ es:

$$\overline{\mathcal{L}} = \{ x \mid x \in \Sigma^* , x \notin \mathcal{L}\}$$

Complemento de lenguajes – Ejemplo

Sea $\mathcal{L} = \{000, 010, 100, 110\}$ un lenguaje sobre $\Sigma = \{0,1\}$

$\overline{\mathcal{L}} = \Sigma^* - \{000, 010, 100, 110\}$

Sea:

- $\mathcal{L}$ = hileras sobre Σ que terminen en 00

$\overline{\mathcal{L}}$ = cualquier hileras sobre Σ que no empiece con 00

Ej. $\overline{\mathcal{L}} = \{1, 10, 11, 010, \ldots\}$

ε es miembro de $\overline{\mathcal{L}}$.

Propiedades del complemento de lenguajes

$\overline{\Sigma^*} = \emptyset$

$\overline{\overline{\mathcal{L}}} = \mathcal{L}$

<u>Cierre:</u>

Si $\mathcal{L} \subseteq \Sigma^*$ entonces $\overline{\mathcal{L}} \subseteq \Sigma^*$

Inverso de Lenguajes

Sea $\mathcal{L}$ un lenguajes sobre Σ

EL inverso o **reflejo** de $\mathcal{L}$, denotado como L^{-1} es:

$$L^{-1} = \{ x^{-1} \mid x \in \mathcal{L} \}$$

Inverso de lenguajes – Ejemplo

Sea $\mathcal{L} = \{000, 010, 100, 110\}$ un lenguaje sobre $\Sigma = \{0,1\}$

$L^{-1} = \{000, 010, 001, 011\}$

Sea $\mathcal{L}$ = { esteban, tomas, Alejandro, amanda}

L^{-1} = {nabetse, samot, ordnajela, adnama}

Propiedades del inverso de lenguajes

- No es idempotente

<u>Cierre:</u>

Si $\mathcal{L} \subseteq \Sigma^*$ entonces $L^{-1} \subseteq \Sigma^*$

Concatenación de Lenguajes

Sea $\mathcal{L}$ y $\mathcal{M}$ dos lenguajes sobre Σ

La concatenación de $\mathcal{L}$ y $\mathcal{M}$, denotado como $\mathcal{LM}$ es:

$$\mathcal{LM} = \{ xw \mid x \in \mathcal{L} \textbf{ y } w \in \mathcal{M} \}$$

Concatenación de lenguajes – Ejemplo

Lenguaje formado por hileras creadas al concatenar una hilera de $\mathcal{L}$ con una hilera de $\mathcal{M}$.

Ej. Sea $\mathcal{L}$ = {000, 010, 100, 110} y $\mathcal{M}$ = { 0, 1, 11}

$\mathcal{LM}$ = { 0000, 0001, 00011, 0100, 0101, 01011, ...}

Sea:

- $\mathcal{L}$ = hileras sobre Σ que terminen en 00
- $\mathcal{M}$ = hileras sobre Σ que empiecen con 1

$\mathcal{LM}$ = hileras de forma ....001.....

$\mathcal{LM}$ = {001, 0001, 0011, 10010001, 101010100111, ...}

Propiedades de la concatenación de lenguajes

Conmutatividad:

La concatenación de lenguajes no es conmutativa

Asociatividad:

$\mathcal{LMP} = (\mathcal{LM})\mathcal{P} = \mathcal{L}(\mathcal{MP})$

Idempotencia:

La concatenación de lenguajes no es idempotente

Elemento neutro:

$\{\varepsilon\}$

$\exists \mathcal{N}\ tq. \forall \mathcal{L}, \mathcal{L}\mathcal{N} = \mathcal{L}$

$\mathcal{N} = \{\varepsilon\}$

Cierre:

Si $\mathcal{L},\mathcal{M} \subseteq \Sigma^*$ entonces $\mathcal{L}\mathcal{M} \subseteq \Sigma^*$

Multiplicación o potencia de Lenguajes

Sea $\mathcal{L}$ un lenguaje entonces:

L^1 = L

L^2 = LL

L^3 = LLL

L^4 = LLLL

…

Multiplicación o potencia de lenguajes – Ejemplo

Sea L = {000,010,100,110}

L^2 ={00000,000010,000100,000110,010000,010010,010100,010110,100000,100010,100100,100110,110000,110010,110100,110110}

Sea M ={000,010,100,110,ε }

M^2={00000,000010,000100,000110,010000,010010,010100,010110,100000,100010,100100,100110,110000,110010,110100,110110,000,010,100,110,ε }

ε afecta la hilera

Propiedades de la multiplicación de lenguajes

L^k = $L^{k-1}L$ = LL^{k-1}

L^0 = $\{\varepsilon\}$

<u>Cierre:</u>

Si $L \subseteq \Sigma^*$ entonces $L^k \subseteq \Sigma^*$

Lenguaje $\boldsymbol{L^*}$

(Parecido a Σ^* pero con lenguajes)

Sea $\mathcal{L}$ un lenguaje, el lenguaje L^* se define como:

$$L^* = \bigcup_{i=0}^{\infty} L^i$$

Lenguaje $\boldsymbol{L^+}$

Sea $\mathcal{L}$ un lenguaje, el lenguaje L^+ se define como:

$$L^+ = \bigcup_{i=1}^{\infty} L^i$$

Lenguajes formales y mecanismos

- Nos interesan dos mecanismos asociados a un lenguaje L^*:
    - Genera todas las hileras de L
    - No genera hileras que no pertenezcan a L
- Un reconocedor de L
    - Mecanismo binario que acepta o rechaza hileras

- Acepta únicamente hileras de L
- No rechaza hileras que pertenezcan a L

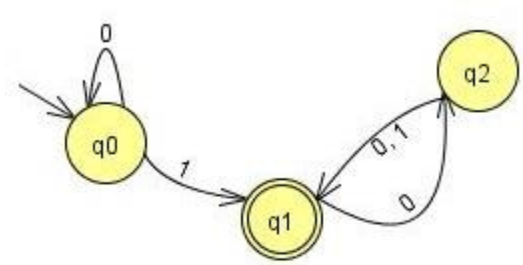
¿Qué es más difícil generar o reconocer?

Dependen, algunos son más fáciles de generar pero difícil de reconocer, otros son difíciles de generar pero fácil de reconocer.

Ej. Mozart, cuando se escucha es fácil de reconocer que es Mozart, pero crear es difícil.

Autómatas deterministicos de Estados finitos (Eargasm)

Esto es un autómata:



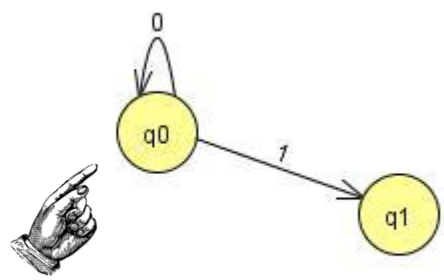
Estructura del autómata

- Es un grafo
- Los nodos tienen etiqueta(opcional)
- Hay 2 tipos de nodo
- Los arcos están etiquetados
- Los nodos son estados
- Los arcos son transiciones
- Hay un estado inicial (flecha sin origen a q0)

Mecánica del autómata

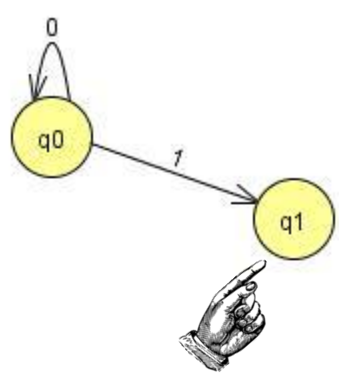
-Un autómata de este tipo procesa hileras

1	0	0	1	0	1	1	1	0	1



Se encontró 1 entonces se mueve al siguiente estado y empieza a leer el siguiente símbolo de la hilera.

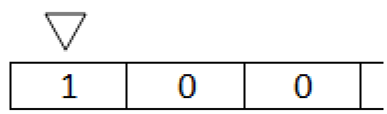
1	0	0	1	0	1	1	1	0	1



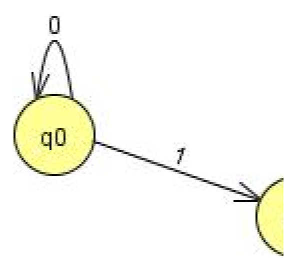
El resultado final es que la hilera se acepta o que la hilera se rechaza

- Se empieza en el estado inicial
- Siempre hay un estado actual

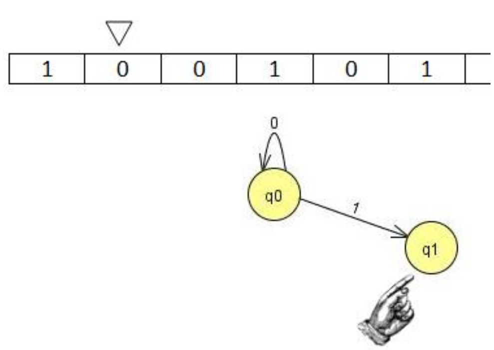
Se toma un símbolo de la hilera a la vez(símbolo actual)



Estado actual tiene transiciones a otros estados(o a sí mismo) según sea el símbolo actual



Se sigue la transición y se toma el siguiente símbolo de la hilera

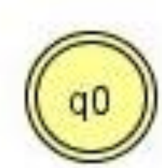


Se establece un nuevo estado actual y un nuevo símbolo actual

El proceso se repite en el nuevo estado

Eventualmente se acaba la hilera
-El estado actual podría ser de aceptación(doble circulo) o de rechazo (circulo simple)

Rechaza



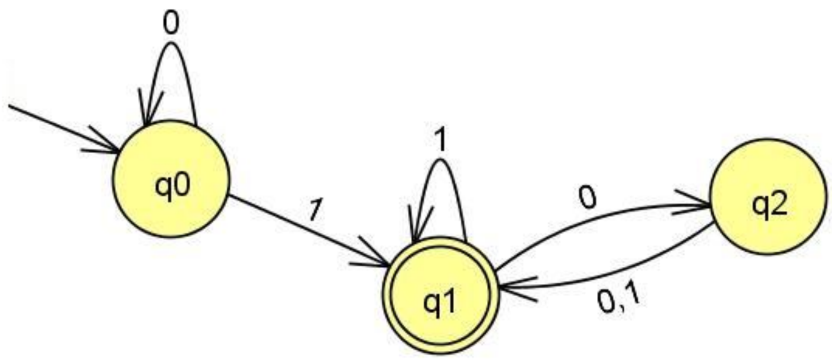
Acepta

El estado actual al acabarse la hilera determina si se acepta o se rechaza
En cualquier punto que no haya transición se rechaza la hilera,

Ahora Chuck Noris nos acompañara en algunos ejemplos,

Ejemplo 1

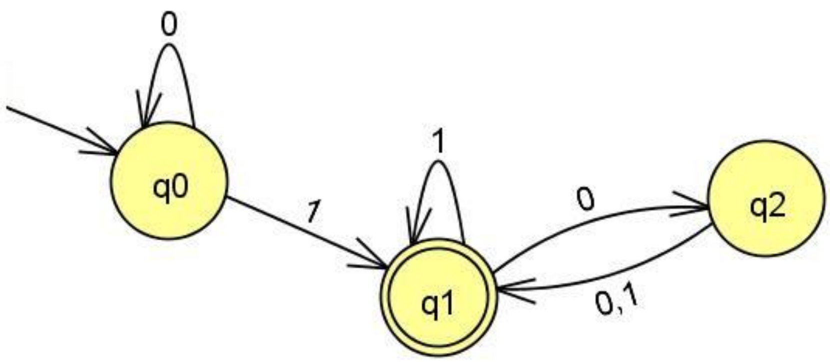
1	0	0	1	0	1	1	1	0	1	1	0	1



Se acepta la hilera

Ejemplo 2

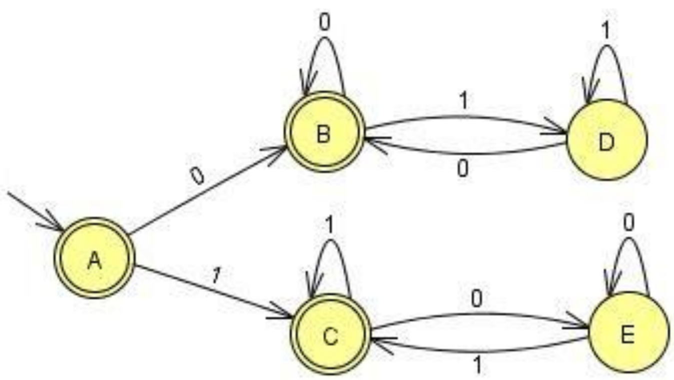
0	0	0	0



No se acepta la hilera

Ejemplo 3

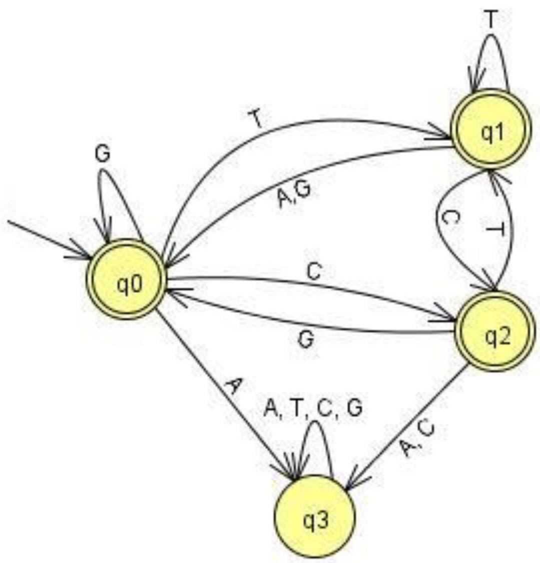
w = 111010101



Se acepta la hilera

Ejemplo 4

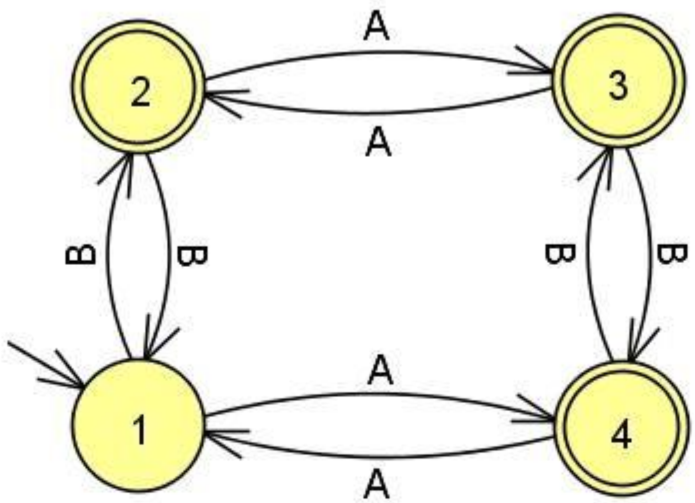
w = GTACGGTTCCG



Se acepta la hilera

Ejemplo 5

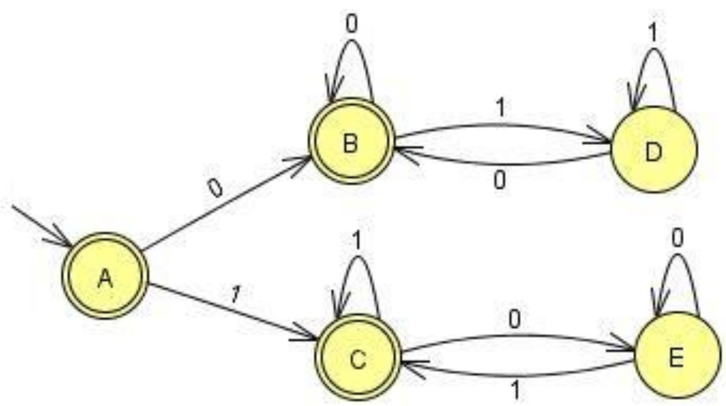
w = ABABABAAABBBAB



Se acepta la hilera

Ejemplo 6

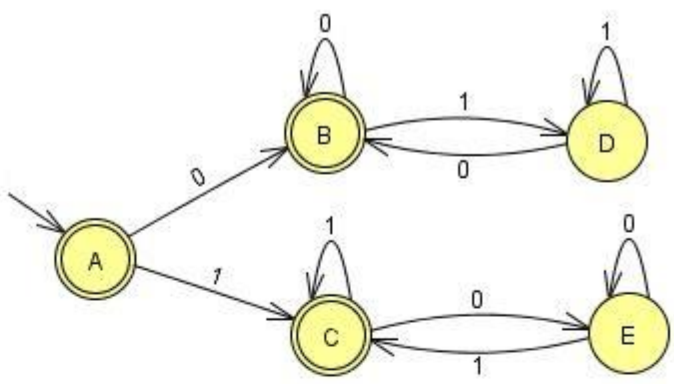
w = 00101110101010



Se acepta la hilera

Ejemplo 7

w = ε

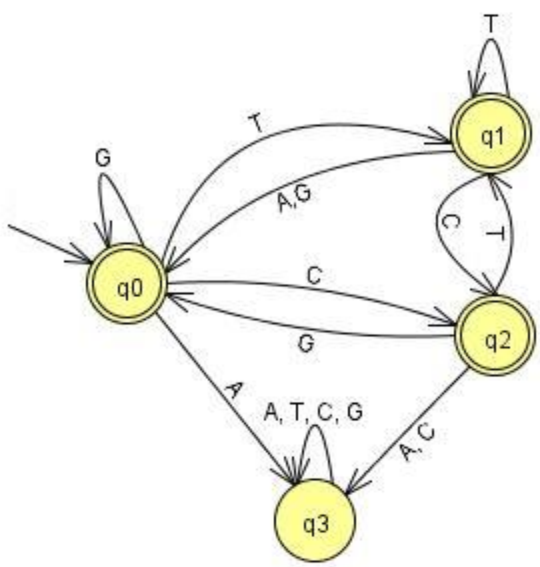


Se acepta la hilera

Porque al entrar al autómata se queda en el estado del conjunto que contiene A

Ejemplo 8

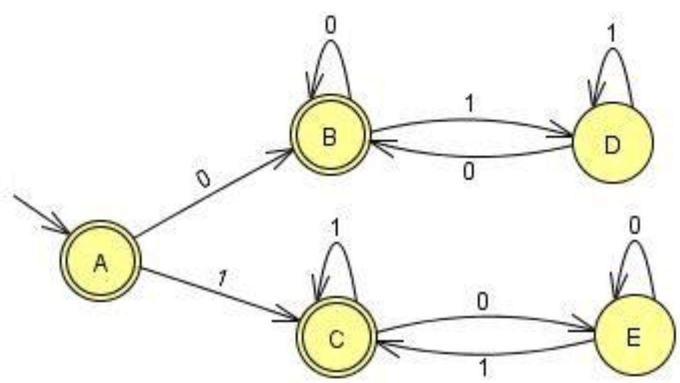
w= GTCGTTGCGACGG



No se acepta la hilera

Ejemplo 9

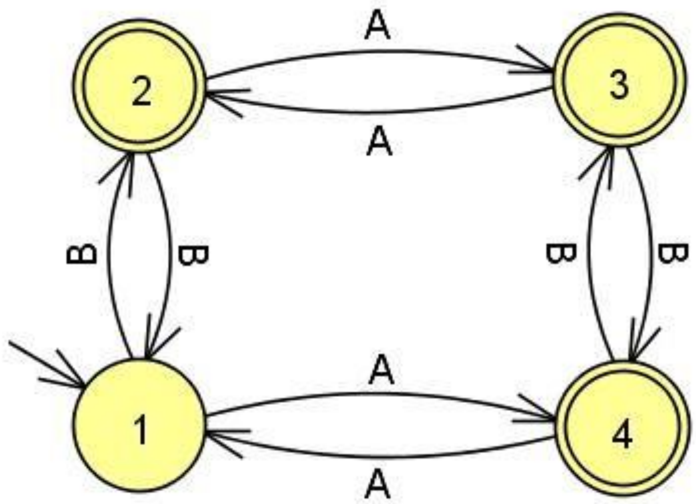
w= 111101110



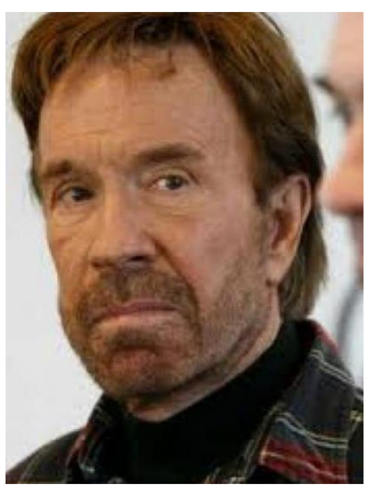
No se acepta la hilera

Ejemplo 10

w= ABAABAAAABBBAB

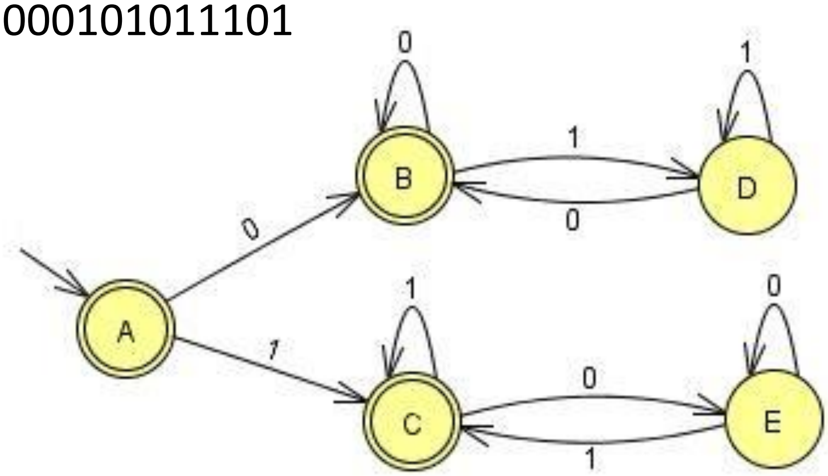


No se acepta la hilera



Ejemplo 11

w= 000101011101



No se acepta la hilera

Hasta este ejemplo llegamos en la clase. Recuerden el quiz del 22 de marzo