Universidad Nacional de Cajamarca
"Norte de la Universidad Peruana"

$(0+1)(0+1)^*$

UNIDAD II: AUTOMATAS

EXPRESIONES REGULARES Y AUTOMATAS

ING. SANDRA RODRIGUEZ AVILA

INTRODUCCION

RECORDANDO LAS GRAMÁTICAS DE TIPO 3 O REGULARES

- Las reglas P de una gramática del tipo 3, tienen la forma:

  $A \rightarrow aB \quad A \rightarrow a$ ó

  $A \rightarrow Ba \quad A \rightarrow a$ donde $A, B \in N$, $a \in T$

- Los lenguajes regulares están definidos por las gramáticas del tipo 3.

OPERACIONES CON LOS LENGUAJES REGULARES

a) Unión o alternativa :

$$L1 \cup L2 = L1 + L2=\{x / x \in L1 \vee x \in L2\}$$

b) Concatenación :

$$L1.L2=\{x1x2 / x1 \in L1 \wedge x2 \in L2\}$$

OPERACIONES CON LOS LENGUAJES REGULARES

c) Potencia de un lenguaje : L^{i}

d) Cierre reflexivo u operación estrella : L^*

$$L^* = \{\lambda\} \cup \{L\} \cup \{LL\} \cup \{LLL\} \ldots$$

e) Cierre transitivo o positivo : L^+

$$L^+ = \{L\} \cup \{LL\} \cup \{LLL\} \ldots$$

CONTENIDO

- EXPRESIONES REGULARES
  - OPERACIONES CON EXPRESIONES REGULARES
  - PRECEDENCIA DE LAS OPERACIONES
  - TEOREMA Y PROPIEDADES
  - EJEMPLOS DE EXPRESIONES REGULARES

- AUTOMATAS
  - TIPOS
  - AUTOMATAS FINITOS: Definición formal
  - AUTOMATAS A PILA: Definición formal

EXPRESIONES REGULARES

$(0+1)(0+1)^*$

- Notación normalizada para describir los lenguajes regulares.
- Son metalenguajes.
- En una expresión regular (e.r.) se pueden utilizar todos los símbolos del alfabeto V y λ. Los operadores que también se pueden utilizar son: $+$, $.$, *, $^+$, $(\ )$

OPERACIONES CON LAS EXPRESIONES REGULARES

$(0+1)(0+1)^*$

a) **Unión o alternativa** : Si α y β son expresiones regulares, $\alpha + \beta$ es una expresión regular tal que,

$$(\alpha + \beta) = (\alpha \mid \beta) = \{\alpha\} \cup \{\beta\}$$

b) **Concatenación** : Si α y β son expresiones regulares, $\alpha\beta$ es una expresión regular tal que,

$$\alpha\beta = (\alpha\beta) = \{\alpha\}\{\beta\}$$

OPERACIONES CON LAS EXPRESIONES REGULARES

$(0+1)(0+1)^*$

c) **Cierre reflexivo u operación estrella** : Si α es una expresión regular, entonces α^* es una expresión regular que denota $(\alpha)^*$.

d) **Cierre positivo** : Si α es una expresión regular, entonces α^+ es una expresión regular que denota $(\alpha)^+$.

PRECEDENCIA DE LAS OPERACIONES

- Uso de paréntesis para indicar la precedencia.
- Cuando no se utilizan paréntesis para evaluar una expresión regular, hay que tener en cuenta:

1ª.- Uso de paréntesis

2ª.- Operación cierre reflexivo y cierre positivo

3ª.- Operación concatenación

4ª.- Alternativa

$(0+1)(0+1)^*$

TEOREMA Y PROPIEDADES

$(0+1)(0+1)^*$

Teorema: Dos expresiones regulares son iguales, si designan al mismo conjunto regular.

Propiedades: A partir del teorema anterior se pueden enunciar las siguientes propiedades :

a) Asociatividad de la operación concatenación

$$\alpha(\beta\gamma)= (\alpha\beta)\gamma$$

b) Distributividad de la operación alternativa respecto de la concatenación

$$\alpha\beta + \alpha\gamma = \alpha(\beta+\gamma)$$

TEOREMA Y PROPIEDADES

$(0+1)(0+1)^*$

c) λ es el elemento neutro de la concatenación, es decir

$$\alpha \lambda = \lambda\alpha = \alpha$$

d) Propiedades de la operación cierre:

1. $(\alpha + \beta)^* = (\alpha^* + \beta^*)^* = (\alpha^*\beta^*)^*$
2. $(\alpha + \lambda)^* = (\alpha^* + \lambda) = \alpha^*$
3. $\alpha\alpha^* + \lambda = \alpha^*$
4. $\lambda^* = \lambda$
5. $\alpha^+ = \alpha\alpha^*$

EJEMPLOS DE EXPRESIONES REGULARES

(0+1)(0+1)*

Ejemplo 1

Sea V = {a,*b*} y la expresión regular aa*bb*. Indicar el lenguaje que denota, y algunas cadenas de dicho lenguaje.

Solución : Algunas cadenas son:

a*b* aa*b* a*bbbb* aaa*bb*

El lenguaje se describe como:

L={cadenas que comienzan por una a y continúan con varias o ninguna a, y siguen con una *b* y continúan con varias o ninguna *b*}

EJEMPLOS DE EXPRESIONES REGULARES

(0+1)(0+1)*

Ejemplo 2

Sea V = {*0,1*}. La expresión regular 1(01)* denota el conjunto de cadenas que empiezan por *1* y van seguidas por *(01)* cualquier numero de veces o ninguna.

Ejemplo 3

Sea V={*0,1*}, la expresión regular $(0 + 1)^+$ denota el conjunto de números en base 2.

EJEMPLOS DE EXPRESIONES REGULARES

$(0+1)(0+1)^*$

Ejemplo 3

Sea el vocabulario {*0,1*}, la expresión regular $(0+1)^+$ denota el conjunto de números en base 2.

Ejemplo 4

Sea el vocabulario {*0, 1, 2*}, la expresión regular $(0 +1 +2)^+$ denota el conjunto de números en base 3.

EJEMPLOS DE EXPRESIONES REGULARES

(0+1)(0+1)*

Ejemplo 5

Dar una **expresión regular** para identificador:

<letra> (<letra> + <dígito>)*

También se puede definir identificador como:

$(a + b + c + \cdots + z)\ (a + b + c + \cdots + z \mid 0 \mid 1 \mid \cdots \mid 9)^*$

L = {*a*, b, c,…z, a1, numero1, 1x, λ, xyz , c1c, c1000,

EJEMPLOS DE EXPRESIONES REGULARES

$(0+1)(0+1)^*$

Ejemplo 6

Dar una expresión regular para los números reales sin exponente del lenguaje Pascal estándar.

$$(\lambda \mid + \mid -)(< \textit{dígito} >< \textit{dígito} >^{*} \cdot < \textit{dígito} >^{*}< \textit{dígito} >)$$

AUTÓMATAS

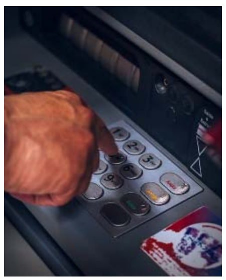
- La palabra *autómata* evoca algo que pretende imitar las funciones propias de los seres vivos, especialmente relacionadas con el movimiento, por ejemplo el típico robot antropomorfo.
- En el campo de los Traductores, Procesadores, Compiladores e Intérpretes, lo fundamental no es la simulación del movimiento, sino la simulación de procesos para tratar información.

CLASIFICACIÓN DE LOS AUTÓMATAS - GRAMÁTICAS Y LENGUAJES

Gramáticas	Lenguajes	Máquinas
Tipo 0	Sin restricciones	Máquinas de Turing
Tipo 1	Sensible al contexto	Autómatas linealmente acotados
Tipo 2	Libre de contexto	Autómatas a Pila
Tipo 3	Regular	Autómatas Finitos

AUTÓMATAS FINITOS o AF

- Reconocen los lenguajes regulares o de tipo 3 y se pueden representar intuitivamente por una cinta y una cabeza de lectura:



AUTÓMATAS FINITOS o AF

- La *cinta de entrada*, sólo contiene símbolos de un determinado alfabeto, y se mueve en una sola dirección.
- El *control de estados*, determina el funcionamiento del autómata.
- Una sentencia de un lenguaje determinado, colocada en la cinta, y leída por el autómata finito, es reconocida por éste, si el control de estados llega a un estado final.

DEFINICIÓN FORMAL DE UN AUTÓMATA FINITO

$$AF = (Q, \mathrm{Te}, \delta, q1, F)$$

donde :

- Q = {conjunto finito de estados}
- Te = {conjunto finito de símbolos de entrada}
- *δ:Q x Te* → Q* es la *función de transición*
- *q1* ∈ *Q*, es el *estado inicial*
- *F C Q:* es el *conjunto de estados finales*

DEFINICIÓN FORMAL DE UN AUTÓMATA FINITO

- *Configuración:* (q, w)

q es el estado actual y w es la cadena por leer.

- *Configuración inicial:* (q1, t)

q es el estado inicial y t es la tira o cadena de entrada.

- *Configuración final:* (qi, λ)

qi Є F y λ indica que no queda nada por reconocer.

- M*ovimiento:* (q,aw) → (q’,w) y se debe de cumplir que δ(q, a)=q’.

LENGUAJE RECONOCIDO POR UN AF

- El lenguaje reconocido por un autómata finito, es:

$$L(AF) = \{t/t \in Te^*, (q1,t) \rightarrow (qi, \lambda), qi \in F\}$$

Teorema 1

- Para toda gramática regular, G3, existe un autómata finito, AF, tal que: L(AF) = L(G3)

- **Teorema 2**

Para todo autómata finito, AF, existe una gramática regular, G3, tal que: L(G3) = L(AF)

AUTOMATA A PILA o AP

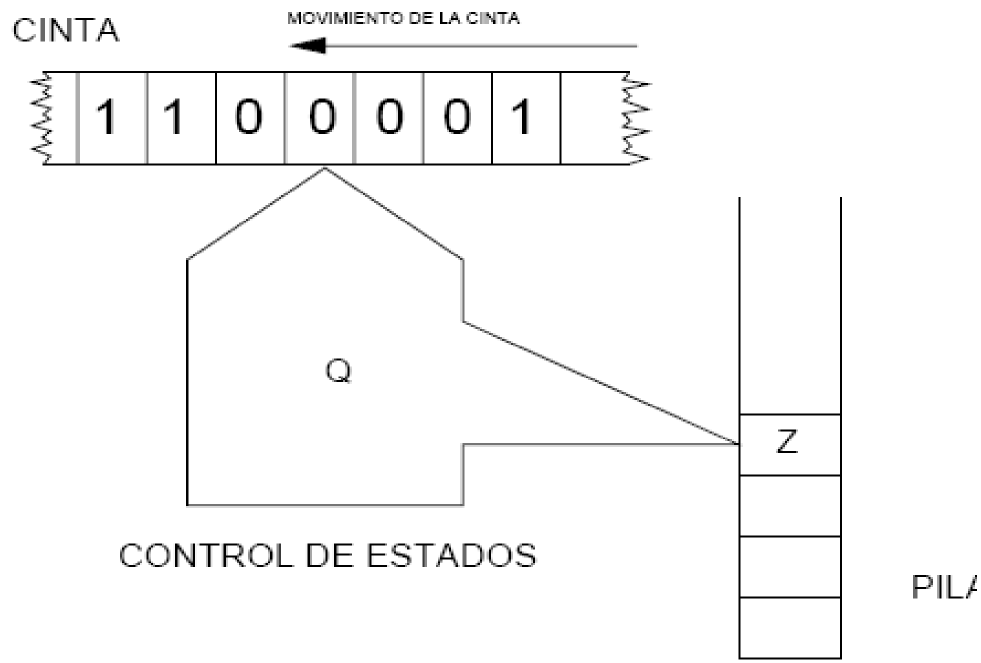
- Reconocen lenguajes libres de contexto, o de tipo 2.
- Se pueden representar como una máquina de Turing, que sólo puede leer de una cinta, y que puede guardar resultados intermedios en una pila.
  - La cinta se desplaza en un sólo sentido, y su cabeza sólo puede leer.
  - La pila, está limitada en un extremo por definición, cuando se lee un elemento de la pila, este desaparece o se saca, y cuando se escribe en la pila, se introduce un elemento.

AUTOMATA A PILA o AP

- Las operaciones elementales son de dos tipos :
  - *Dependientes de la entrada* : se lee un símbolo , y se desplaza la cinta, y en función del estado en que se encuentra la cinta, y Z (el valor de la pila), el control pasa a otro estado q’, y en la pila se introduce Z’, o se extrae Z, o no se hace nada.
  - *Independientes de la entrada* : puede ocurrir lo mismo que en el caso anterior, sólo que no se lee ningún símbolo, la cinta no se mueve, lo que permite manejar la pila sin las informaciones de entrada.

AUTOMATA A PILA o AP

- En cualquier caso, si se vacía la pila (es decir se extrae todas las Z) el autómata se detiene.



AUTOMATA A PILA

- Se puede definir formalmente por 7 elementos:

$$AP = (Q, Te, Tp, \delta, q0, Zo, F)$$

donde :

- Q es el *conjunto finito de estados*.
- Te es el *alfabeto de entrada*, es finito.
- Tp es el *alfabeto de pila*.

AUTOMATA A PILA o AP

- δ es la *función de transición*, y es una aplicación de la forma :

$$\delta : Q \times \{Te \cup \{\lambda\}\} \times Tp \rightarrow Q \times Tp^*$$

- q0 es el *estado inicial*, y cumple q0Є Q.
- Zo es el *símbolo inicial que contiene la pila* antes de comenzar, evidentemente Zo Є Tp
- F es el *conjunto de estados finales*, evidentemente F C Q

AUTOMATA A PILA O AP

- **Configuración**:

$$(q,w,\alpha)$$

donde :

- q representa el *estado actual del autómata*
- w es la *cadena de entrada que resta por analizar*, siendo $w \in Te^*$; si $w=\lambda$ se asume que toda la cadena de entrada ya ha sido leída.
- α es el *contenido de la pila*, en el instante considerado, y $\alpha=\lambda$ Indica que la pila está vacia. Por supuesto $\alpha \in Tp^*$

AUTOMATA A PILA O AP

- **Movimiento**: transición entre configuraciones.

$$(q, aW, z) \rightarrow (q', W, \gamma)$$

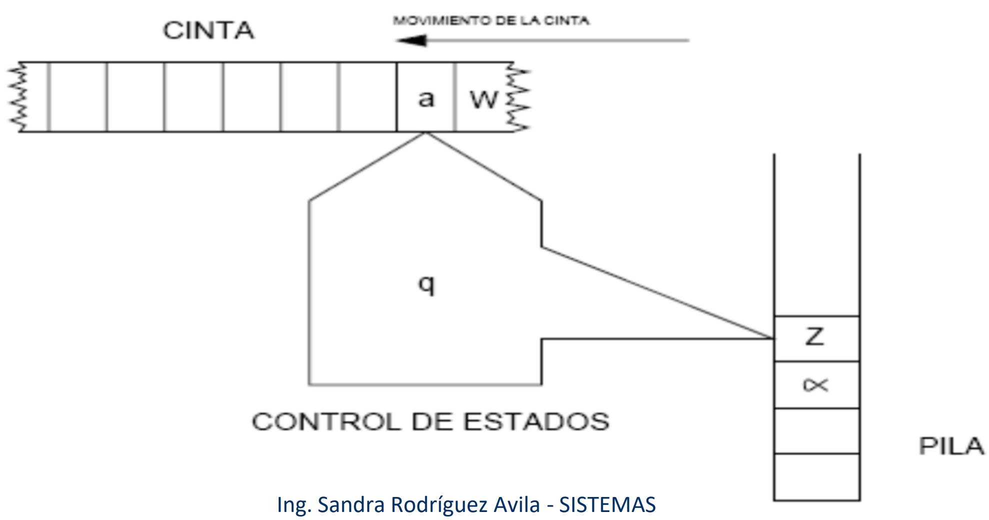
Donde la función de transición δ, es:

$$\delta(q, aW, Z) \rightarrow (q', \gamma)$$

siendo $q \in Q$, $a \in Te \cup \{\lambda\}$, $W \in Te^*$, $Z \in Tp$, $\gamma \in Tp^*$

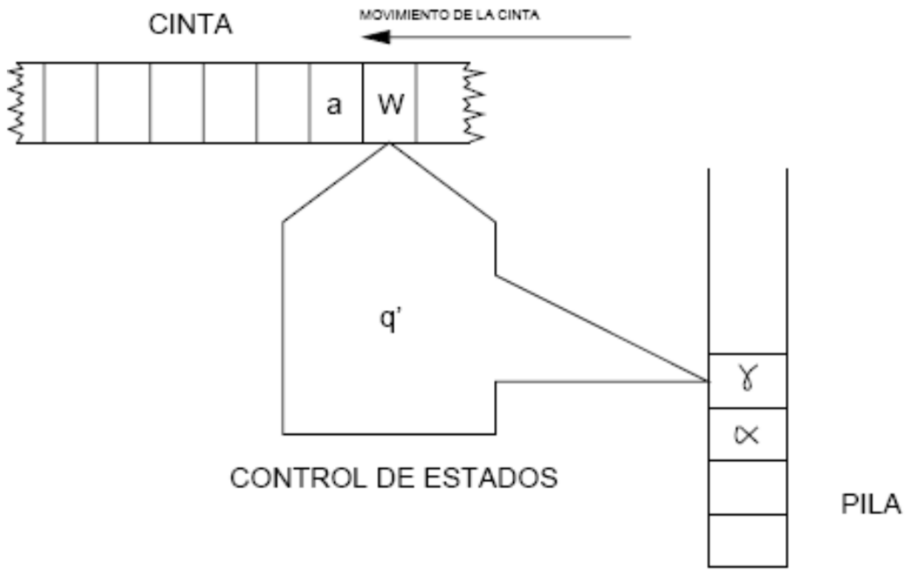
AUTOMATA A PILA o AP

Situación del AP antes del movimiento. La pila contiene determinados símbolos representados por la concatenación de Z y α , siendo el situado más a la izquierda Z el que se encuentra en cabeza de la pila.



AUTOMATA A PILA O AP

- El AP pasa a un estado q’, avanzando la cabeza de lectura al siguiente símbolo y procediendo a realizar determinadas sustituciones en la cabeza de la pila.



LENGUAJE RECONOCIDO POR UN AP

Se puede definir de dos formas :

a) Cuando el AP alcanza una configuración final :

$$L(AP) = \{w/w \in \mathrm{Te}^* \ y \ (q_0,w,Z_0) \rightarrow (qf,\lambda,\alpha)\}$$

b) El AP reconoce la cadena cuando la pila queda vacía, independientemente del estado:

$$L(AP) = \{w/\mathrm{w} \in \Sigma^* \ y \ (q_0,\mathrm{w},Z_0) \rightarrow (q',\lambda,\lambda)\}$$

LENGUAJE RECONOCIDO POR UN AP

Teorema

- Para toda gramática libre de contexto G2, existe un reconocedor constituido por un autómata de pila RAP, tal que el lenguaje generado por la gramática L(G2) es reconocido por el autómata RAP.

$$L(G2)=L(RAP)$$

CONCLUSIONES

- Las expresiones regulares son una notación normalizada de los lenguajes regulares (Tipo 3).
- Los autómatas son dispositivos formales que reconocen lenguajes definidos por las gramáticas de los 4 tipos según la clasificación de Noam Chomsky:
  - Los AF reconocen lenguajes regulares (Tipo 3)
  - Los AP reconocen lenguajes definidos por gramáticas de contexto libre (Tipo 2)

BIBLIOGRAFIA

- ALFONSECA Enrique, ALFONSECA Manuel y MORIYON Roberto**.** ***Teoría de Autómatas y Lenguajes Formales***. 2007. Madrid. Editorial Mc Graw Hill.
- http://dehesa.unex.es/bitstream/10662/2367/1/978-84-691-6345-0.pdf
- http://di002.edv.uniovi.es/~cueva/publicaciones/libros/36_LGA.pdf

RECURSOS GRAFICOS

- Pixabay
- Pexels
- Icon-Icons