Equivalencia AP y GLC

Teorema

El conjunto de lenguajes reconocido por gramáticas libres de contexto **es exactamente igual** al conjunto de lenguajes reconocidos por autómatas de pila **NO** deterministas

- Ida (⇒): Si L = L(G) para alguna GLC G entonces existe un APND M tal que L = L(M)
- Vuelta (<=): Si L = L(M) para algún APND M entonces existe una GLC G tal que L = L(G)

Ida

El conjunto de lenguajes reconocido por gramáticas libres de contexto **es exactamente igual** al conjunto de lenguajes reconocidos por autómatas de pila **NO** deterministas

Ida (⇛): Si L = L(G) para alguna GLC G entonces existe un APND M tal que L = L(M)

Vamos a convertir una GLC G arbitraria en un **APND** M tal que M simula derivaciones **más a la izquierda** de G

Recordemos que una derivación más a la izquierda suele verse así

$$G: \qquad S \Rightarrow \cdots\cdots \Rightarrow \underbrace{\sigma_1\sigma_2\cdots\sigma_k}_{\text{Substring de terminales generado}} V \underbrace{\alpha_1\alpha_2\cdots\alpha_m}_{\text{Forma Sentencial para derivar el resto del string}} \Rightarrow \cdots$$

$\sigma_i \in \Sigma$

$\alpha_i \in (V \cup \Sigma)$

Donde V es la variable izquierda de la derivación

Entonces, sea esta nuestra derivación

$$G: \quad S \Rightarrow \cdots\cdots \Rightarrow \underbrace{\sigma_1\sigma_2\cdots\sigma_k}_{\substack{\text{Substring de}\\ \text{terminales}\\ \text{generado}}}\underbrace{V\alpha_1\alpha_2\cdots\alpha_m}_{\substack{\text{Forma Sentencial}\\ \text{para derivar el}\\ \text{resto del string}}} \Rightarrow \cdots$$

$$\sigma_i \in \Sigma$$

$$\alpha_i \in (V \cup \Sigma)$$

Creamos una simulación usando un APND

$M:$

Input

σ_1	σ_2	$\cdots$	σ_k	σ_{k+1}	$\cdots$

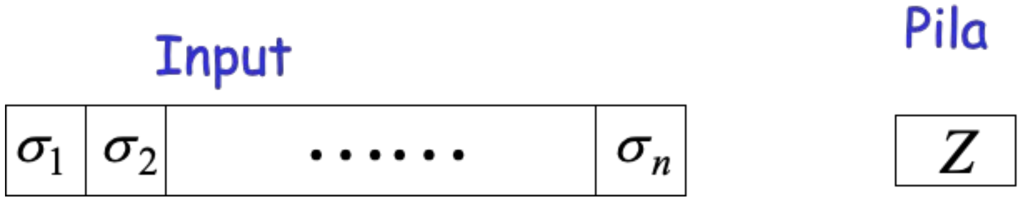
Pila

V
α_1
$\vdots$
α_m
Z

Y cuando la derivación llega a la cadena de terminales

$$S \Rightarrow \cdots\cdots\cdots \Rightarrow \sigma_1 \sigma_2 \cdots \sigma_n$$

La pila quedaría en el siguiente estado (se alcanza el fin del input)

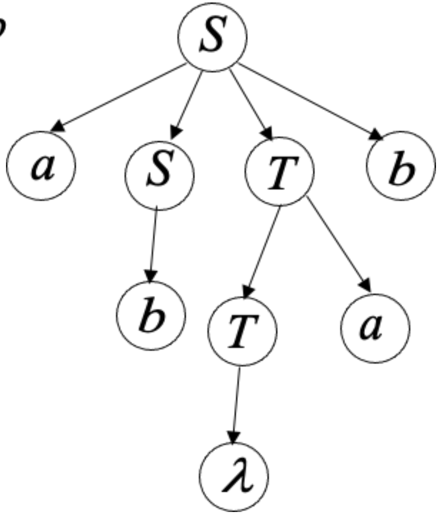


Intuición

Teniendo la siguiente gramática y su derivación más a la izquierda

$$S \to aSTb$$
$$S \to b$$
$$T \to Ta$$
$$T \to \lambda$$

$$S \Rightarrow aSTb \Rightarrow abTb \Rightarrow abTab \Rightarrow abab$$



Podemos recorrer el árbol (t) utilizando la pila (s)

```
def simulate(t, s):
  if t.is_empty(): print([])
  else:
    s.push(t.root())
    while not s:
      top = s.top()
      if top.leaf(): print(top); s.pop()
      else: s.push(top.children())
```

Ejemplito

Gramática:

$S \to aSTb$

$S \to b$

$T \to Ta$

$T \to \lambda$

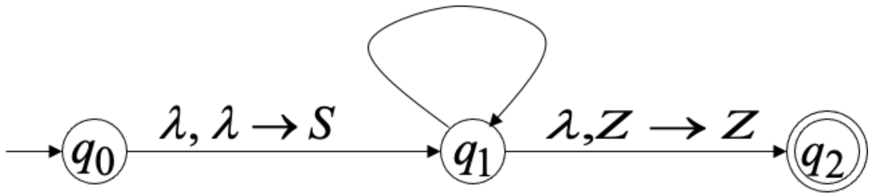
APND:

$\lambda, S \to aSTb$

$\lambda, S \to b$

$\lambda, T \to Ta \quad a, a \to \lambda$

$\lambda, T \to \lambda \quad b, b \to \lambda$



Ejemplito

Vemos una derivación más a la izquierda

$$\boxed{\begin{array}{l} S \rightarrow aSTb \\ S \rightarrow b \\ T \rightarrow Ta \\ T \rightarrow \lambda \end{array}} \qquad S \Rightarrow aSTb \Rightarrow abTb \Rightarrow abTab \Rightarrow abab$$

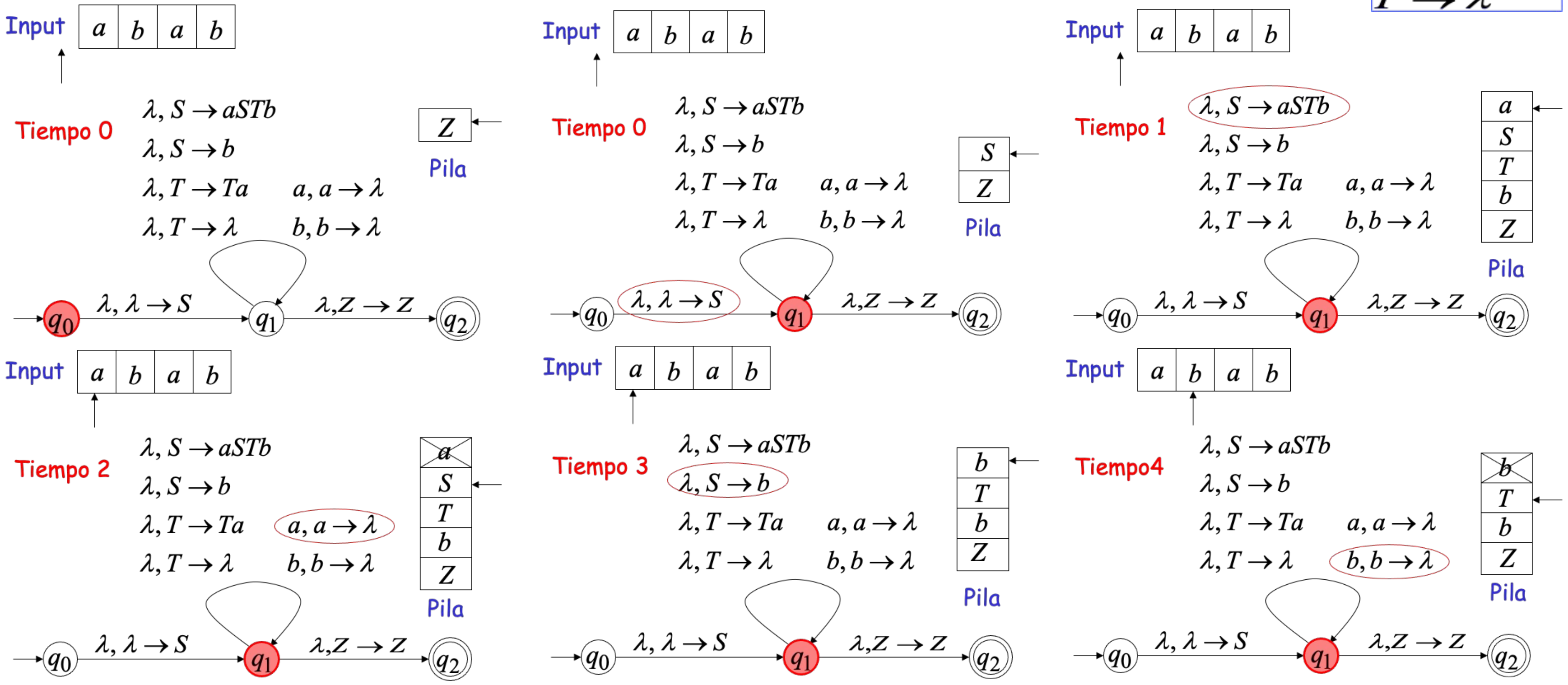
Ejemplito

$S \rightarrow aSTb$

$S \rightarrow b$

$T \rightarrow Ta$

$T \rightarrow \lambda$



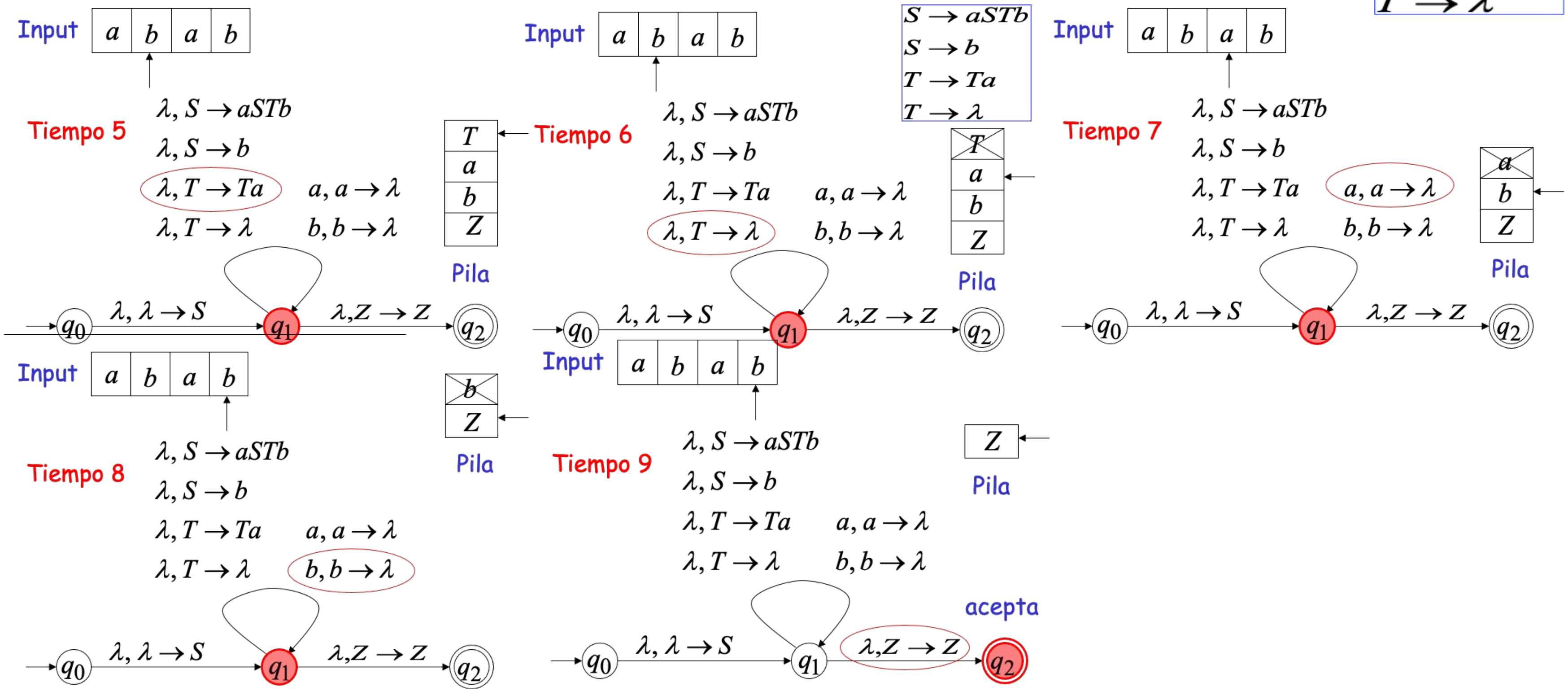
Ejemplito

$S \rightarrow aSTb$

$S \rightarrow b$

$T \rightarrow Ta$

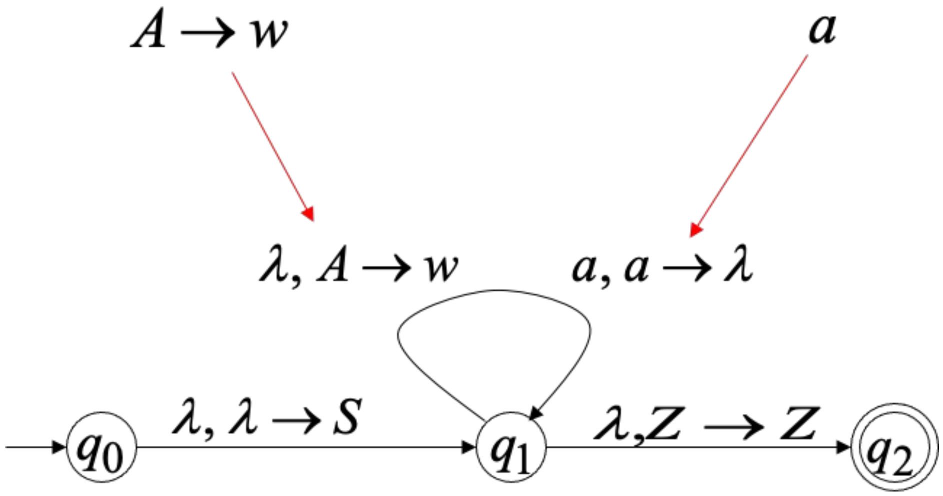
$T \rightarrow \lambda$



Vemos que en general, podemos construir un APND M a partir de una GLC G (con **L(G) = L(M)**)

para cualquier producción

para cualquier terminal



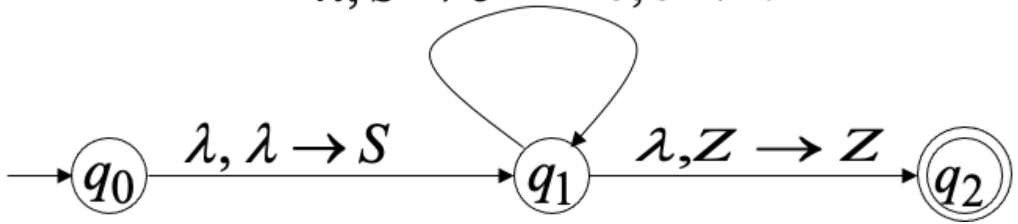
Otro ejemplito

$$S \rightarrow aSa$$
$$S \rightarrow bSb$$
$$S \rightarrow c$$

$$L(M) = \{vcv^R : v \in \{a,b\}^*\}$$

APND : M

$$\lambda, S \rightarrow aSa \qquad a, a \rightarrow \lambda$$
$$\lambda, S \rightarrow bSb \qquad b, b \rightarrow \lambda$$
$$\lambda, S \rightarrow c \qquad c, c \rightarrow \lambda$$



Vuelta

El conjunto de lenguajes reconocido por gramáticas libres de contexto **es exactamente igual** al conjunto de lenguajes reconocidos por autómatas de pila **NO** deterministas

Vuelta (<=): Si L = L(M) para algún APND M entonces existe una GLC G tal que L = L(G)

Vamos a convertir un APND M arbitrario en una **GLC** C donde C simula el comportamiento de M

Una derivación de la gramática G se lee

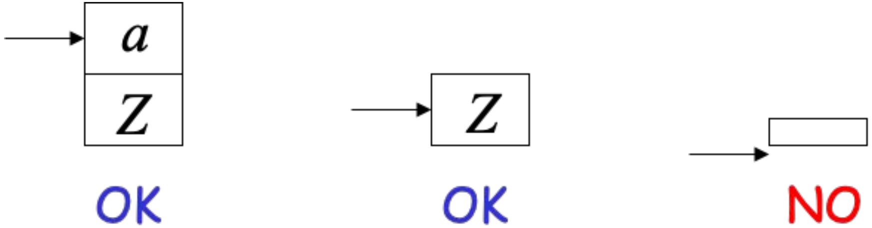
$$S \Rightarrow \cdots \Rightarrow \underset{\text{Input procesado}}{\overset{\text{terminales}}{abc\ldots}}\,\underset{\text{Contenidos de la pila}}{\overset{\text{variables}}{ABC\ldots}} \Rightarrow \ldots \Rightarrow abc\ldots$$

Intuición

Previo a la construcción de la gramática, debemos modificar el APND de modo que

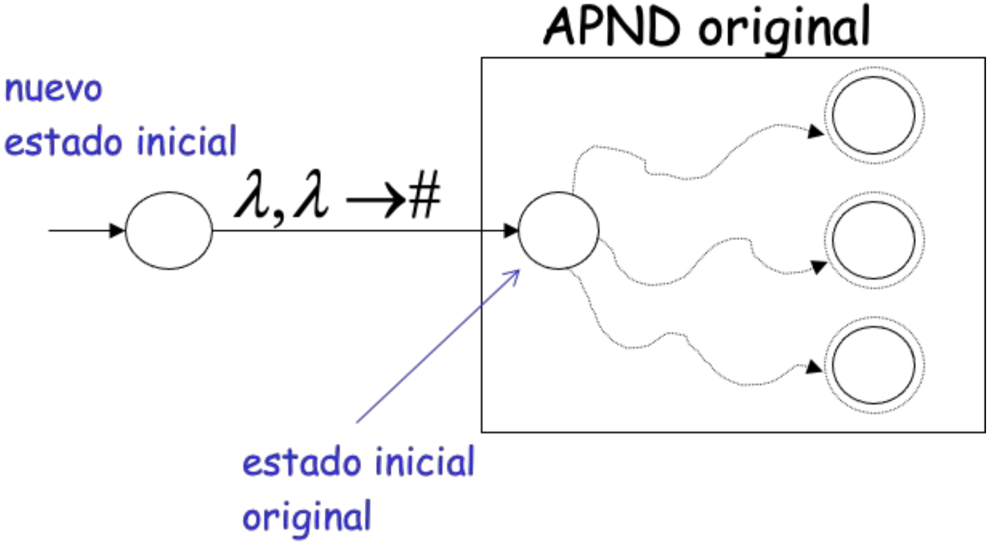
1. La pila nunca está vacía (sólo cuando acepta el string)
2. Tiene un único estado final y vacía la pila cuando acepta el string
3. Tiene transiciones de una forma especial

La pila nunca está vacía



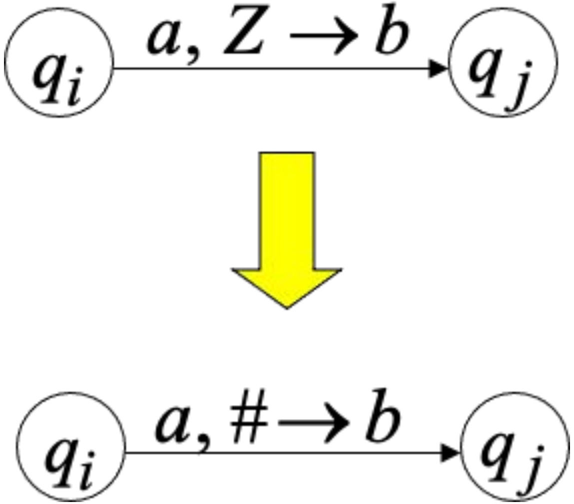
La pila nunca está vacía

Colocando el símbolo # en la pila



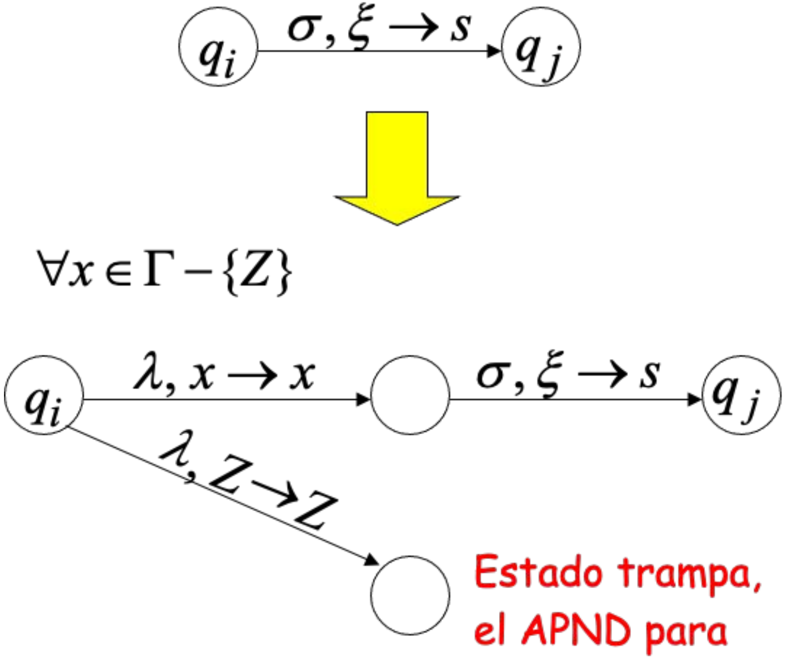
La pila nunca está vacía

En las transiciones, reemplazamos toda instancia de **Z** con **#**



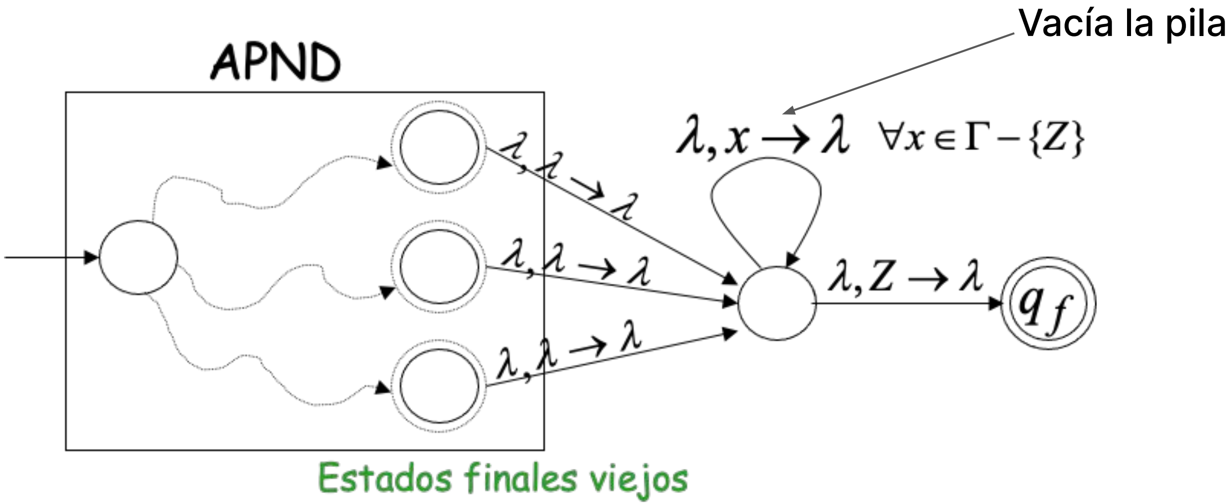
La pila nunca está vacía

Transformamos todas las transiciones de modo que si el autómata intenta sacar o reemplazar el símbolo Z se detendrá



Tiene un único estado final y vacía la pila cuando acepta el string

Modificamos el APND de modo que vacíe la pila y tenga un único estado final



Transiciones especiales

Las transiciones deben tener alguna de las siguientes formas

$$q_i \xrightarrow{\sigma,\, B \to \lambda} q_j$$

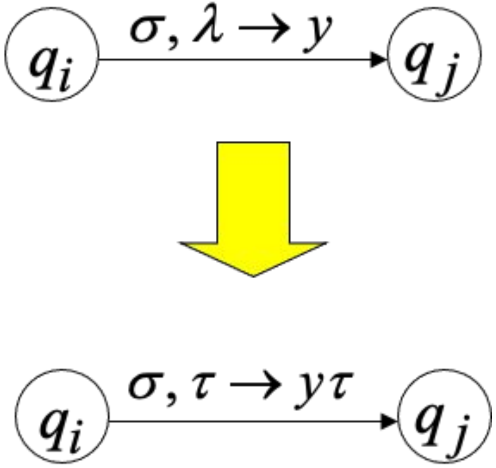
o bien

$$q_i \xrightarrow{\sigma,\, B \to CD} q_j$$

$B, C, D:$ símbolos de la pila

Transiciones especiales

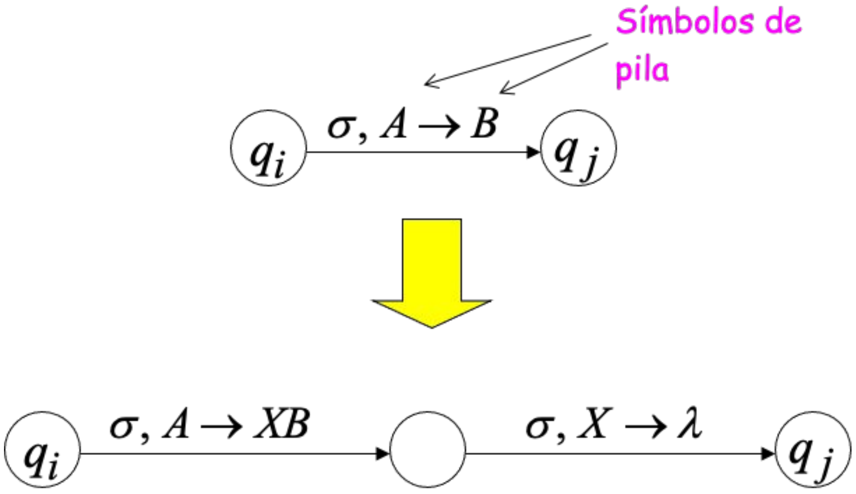
Realizamos las siguientes conversiones



$$\forall \tau \in \Gamma - \{Z\}$$

Transiciones especiales

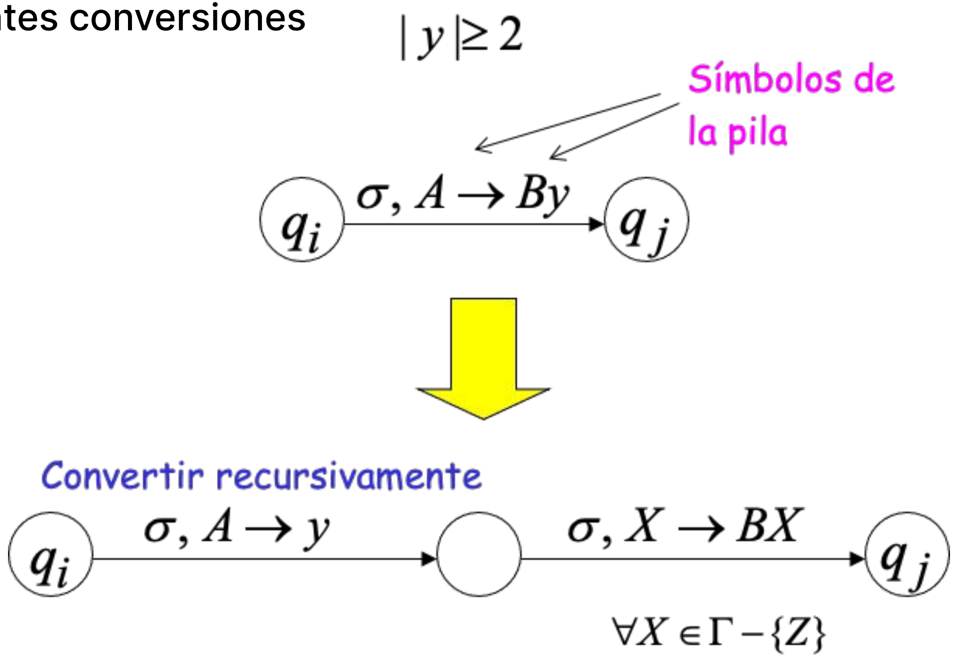
Realizamos las siguientes conversiones



$$X \in \Gamma - \{Z\}$$

Transiciones especiales

Realizamos las siguientes conversiones



Ejemplo de APND en forma correcta

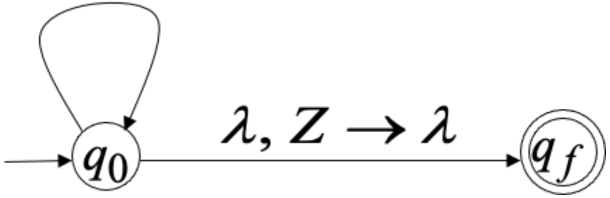
$$L(M) = \{w : \quad n_a = n_b\}$$

Z : símbolo inicial de la pila

$a, Z \rightarrow 0Z \qquad b, Z \rightarrow 1Z$

$a, 0 \rightarrow 00 \qquad b, 1 \rightarrow 11$

$a, 1 \rightarrow \lambda \qquad b, 0 \rightarrow \lambda$



Construcción

Construimos la gramática G

Los **no terminales** de la gramática va a definirse a partir de $(q_i B q_j)$ Símbolo de la pia

estados del AP

Los **terminales** serán todos los símbolos input del AP

Construcción

Para cada transición $q_i \xrightarrow{a,\ B \to \lambda} q_j$ agregamos una producción a G, $(q_i B q_j) \to a$

Para cada transición $q_i \xrightarrow{a,\ B \to CD} q_j$ agregamos producciones

$(q_i B q_k) \to a(q_j C q_l)(q_l D q_k)$ para todos los posibles estados $\mathbf{q_k}$, $\mathbf{q_l}$

Construcción

El no terminal inicial se define como

Símbolo inicial de la pila

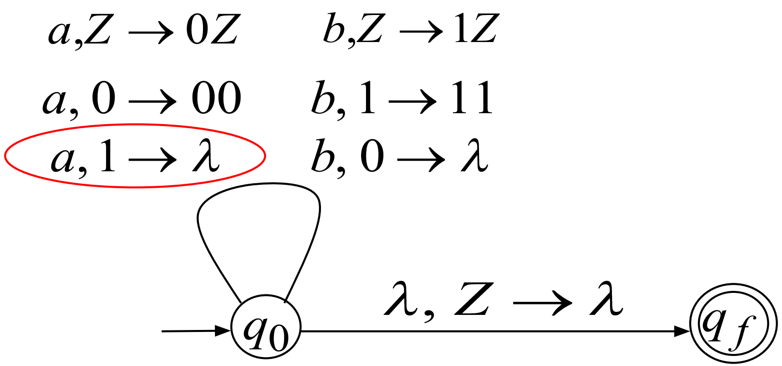
$$(q_o Z q_f)$$

Estado inicial

Estado final

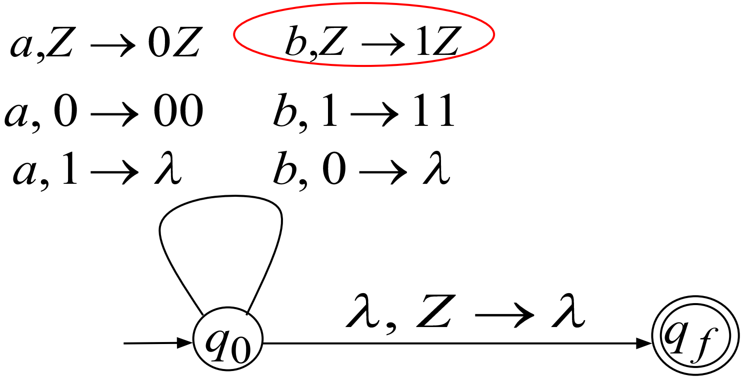
Ejemplo

Sea el siguiente AP



Creamos la producción a la grammatica G $(q_0 1 q_0) \to a$

Ejemplo

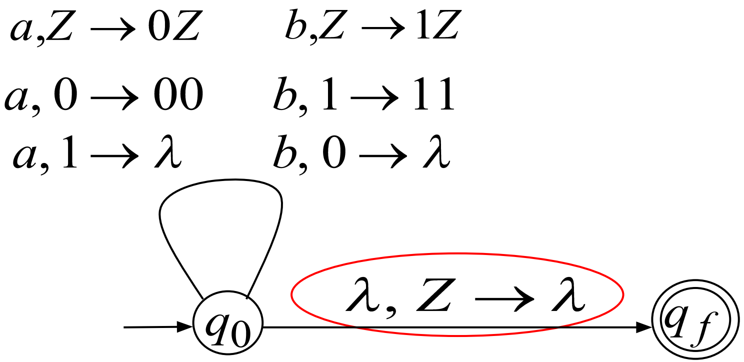


Creamos las producciones a la grammatica G

$$(q_0 Z q_0) \to b(q_0 1 q_0)(q_0 Z q_0) \mid b(q_0 1 q_f)(q_f Z q_0)$$

$$(q_0 Z q_f) \to b(q_0 1 q_0)(q_0 Z q_f) \mid b(q_0 1 q_f)(q_f Z q_f)$$

Ejemplo



Creamos la producción a la grammatica G $(q_0 Z q_f) \rightarrow \lambda$

Ejemplo

Variable inicial de la gramática $(q_0 Z q_f)$: variable inicial

Producciones

$$(q_0 Z q_0) \rightarrow b(q_0 1 q_0)(q_0 Z q_0) \mid b(q_0 1 q_f)(q_f Z q_0)$$
$$(q_0 Z q_f) \rightarrow b(q_0 1 q_0)(q_0 Z q_f) \mid b(q_0 1 q_f)(q_f Z q_f)$$

$$(q_0 1 q_0) \rightarrow b(q_0 1 q_0)(q_0 1 q_0) \mid b(q_0 1 q_f)(q_f 1 q_0)$$
$$(q_0 1 q_f) \rightarrow b(q_0 1 q_0)(q_0 1 q_f) \mid b(q_0 1 q_f)(q_f 1 q_f)$$

$$(q_0 Z q_0) \rightarrow a(q_0 0 q_0)(q_0 Z q_0) \mid a(q_0 0 q_f)(q_f Z q_0)$$
$$(q_0 Z q_f) \rightarrow a(q_0 0 q_0)(q_0 Z q_f) \mid a(q_0 0 q_f)(q_f Z q_f)$$

$$(q_0 0 q_0) \rightarrow a(q_0 0 q_0)(q_0 0 q_0) \mid a(q_0 0 q_f)(q_f 0 q_0)$$
$$(q_0 0 q_f) \rightarrow a(q_0 0 q_0)(q_0 0 q_f) \mid a(q_0 0 q_f)(q_f 0 q_f)$$

$$(q_0 1 q_0) \rightarrow a$$
$$(q_0 0 q_0) \rightarrow b$$

$$(q_0 Z q_f) \rightarrow \lambda$$

Una derivación de la nueva gramática

Vemos una derivación de la cadena *abba*

$$(q_0Zq_f) \to a(q_0 0 q_0)(q_0Zq_f) \qquad (q_0 0 q_0) \to b$$

$$(q_0Zq_f) \Rightarrow a(q_0 0 q_0)(q_0Zq_f) \Rightarrow$$

$$(q_0Zq_f) \to b(q_0 1 q_0)(q_0Zq_f)$$

$$ab(q_0Zq_f) \Rightarrow$$

$$(q_0 1 q_0) \to a$$

$$abb(q_0 1 q_0)(q_0Zq_f) \Rightarrow$$

$$(q_0Zq_f) \to \lambda$$

$$abba(q_0Zq_f) \Rightarrow abba$$

En general podemos ver que $(q_i A q_j) \stackrel{*}{\Rightarrow} w$ si y sólo si el APND va desde q_i a q_j leyendo la cadena *w* y la pila no cambia debajo de A. Entonces A es eliminada de la pila

Y en consecuencia, $(q_0 Z q_f) \stackrel{*}{\Rightarrow} w$ **si y sólo si w es aceptado por el APND**