ESQUEMAS

TEORÍA DE AUTÓMATAS Y LENGUAJES FORMALES

Por José Román Hernández Martín

Σ

http://www.emezeta.com/

TEORIA DE AUTOMATAS Y LENGUAJES FORMALES (TALF)

* no responder con ejemplos las demostraciones !!!

- CONCEPTOS PREVIOS
  - PROPOSICIÓN ($\equiv$) — CONJUNCIÓN ($\overset{y}{\wedge}$) — CONDICIONAL ($\overset{entonces}{\rightarrow}$)
  - NEGACIÓN ($\neg$) — DISYUNCION ($\overset{ó}{\vee}$) — BICONDICIONAL ($\leftrightarrow$)
  - RECIPROCA ($\frac{A \rightarrow B}{B \rightarrow A}$) — CONTRAPUESTA ($\frac{\neg P \rightarrow \neg Q}{\neg Q \rightarrow \neg P}$) — CONTRADICCIÓN (siempre falso)
  - TAUTOLOGÍA (siempre verdadero)
  - TEORIA DE CONJUNTOS
    - LEY DE DEMORGAN $\overline{A \cup B} = \bar{A} \cap \bar{B}$, $\overline{A \cap B} = \bar{A} \cup \bar{B}$
    - PRODUCTO CARTESIANO $A \times B = \{(x,y) \mid x \in A, y \in B\}$
    - RELACIÓN: Relación de A con B, R; $R \subseteq A \times B$
    - PARTICIÓN $P = \{x \mid x \subseteq C\}$
    - FUNCIONES
      - TOTAL $f: A \rightarrow B \equiv f \subseteq A \times B$ $Dom(f) = A$
      - PARCIAL $f: A \rightarrow B = f \subseteq A \times B$ $Dom(f) \subset A$
    - EQUIPOTENCIA $\cong$ $|A| = |B|$
  - TECNICAS DE DEMOSTRACIÓN
    - DIRECTA
    - INDIRECTA (contraposición)
    - INDUCCIÓN
    - REDUCCIÓN
- CONCEPTOS BÁSICOS
  - ALFABETO (Σ) $\Sigma \neq \emptyset$ $\Sigma = \{\alpha, \epsilon\}$
  - CADENA VACÍA (ϵ) EPSILON
  - LONGITUD ($|w|$)
  - LENGUAJE CJTO DE CADENAS DE UN Σ.
  - LENGUAJE UNIVERSAL (Σ^*) CJTO DE CADENAS (TODAS) POSIBLES DE Σ.
  - CIERRE DE KLEENE $L^* = \{\epsilon, a, b, aa, ab, ba, bb..\}$
  - REPETICIONES DE UNA CADENA
    - $w^n = abab$
    - $w = ab, n = 2$
  - PREFIJOS / SUFIJOS $\underline{z}abc$ $abc\underline{z}$
    - * toda cadena es pref/suf de si misma: $w = w\epsilon$, $w = \epsilon w$
  - PREFIJOS PROPIOS
  - CONCATENACIÓN
  - POTENCIACIÓN

TEORIA DE AUTOMATAS Y LENGUAJES FORMALES (TALF)

LENGUAJES REGULARES

ϕ, $\{\varepsilon\}$, $(\forall x \in \Sigma)$ $\{x\}$, L_1 y L_2 son L.R. $\cup, \cdot, {}^*$ son L.R.

EXPRESIONES REGULARES

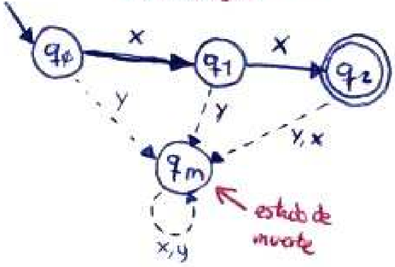
$\phi, \varepsilon, (\forall y \in \Sigma)$ y, r, s son L.R. $\Rightarrow rs, r|s, r*$ son L.R.

- PROPIEDADES UNIÓN ($\cup$): $r|(s|t) = (r|s)|t$, $r|s = s|r$, $\phi|r = r|\phi = r$, $r|r = r$
- PROPIEDADES CONCATENACIÓN ($\cdot$): $r(st) = (rs)t$, $rs \neq sr$, $\varepsilon r = r\varepsilon = r$, $\phi r = r\phi = \phi$, $r(s|t) = rs|rt$
- PROPIEDADES CLAUSURA ($*$): $\varepsilon* = \varepsilon$, $r* = \varepsilon|rr*$, $\phi* = \varepsilon$, $rr* = r*r$, $(r*|s*)* = (r|s)*$

AUTÓMATA FINITO (DFA)
DETERMINISTA

DFA $M \equiv (\Sigma, Q, q_\phi, F, \delta)$ $\quad \delta: Q \times \Sigma \rightarrow Q$

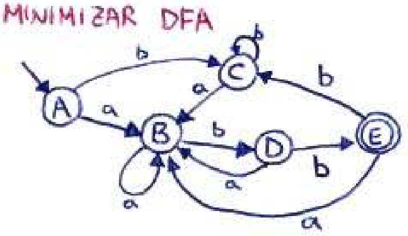
$\delta(q, a) = p$



$M = (\{x, y\}, \{q_\phi, q_1, q_2, q_m\}, q_\phi, \{q_2\}, \delta)$

δ	x	y
q_ϕ	q_1	q_m
q_1	q_2	q_m
q_2	q_m	q_m
q_m	q_m	q_m

- ESTADO INDISTINGUIBLES $q \equiv p$
- MINIMIZAR DFA



$\{E\}\ \{A\ B\ C\ D\} \Rightarrow \{E\}\ \{A\ B\ C\}\ \{D\}$

C B DC B E $\qquad$ C B DC B

$\Downarrow$

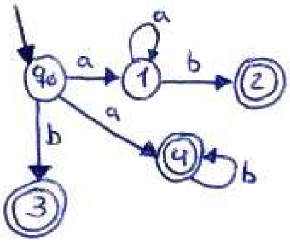
REPRES. (al azar)

$\{E\}\ \{A\ C\}\ \{B\}\ \{D\}$

C B C

ESTADOS INDISTINGUIBLES

AUTÓMATA FINITO NO DETERMINISTA (NFA)



NFA $M \equiv (Q, \Sigma, \delta, q_\phi, F)$ $\quad \delta: Q \times \Sigma \rightarrow 2^Q$

$\delta(q, a) = \{p_1, p_2 \dots\}$

* No hay est. de absorción
* No sabe qué hacer si no $\exists$ trans.

δ	a	b
q_ϕ	{1,4}	{3}
q_1	{1}	{2}
q_2	–	–
q_3	–	–
q_4	–	{4}

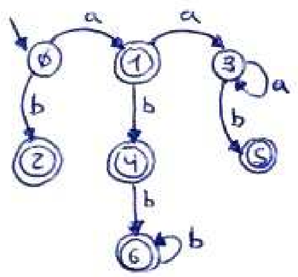
- CONVERSIÓN DE NFA A DFA

Ø $\begin{cases} \delta(\{\phi\}, a) = \{1, 4\} \\ \delta(\{\phi\}, b) = \{3\} \end{cases}$

① $\begin{cases} \delta(\{1, 4\}, a) = \{1\} \\ \delta(\{1, 4\}, b) = \{2, 4\} \end{cases}$

② $\begin{cases} \delta(\{3\}, a) = \phi \\ \delta(\{3\}, b) = \phi \end{cases}$

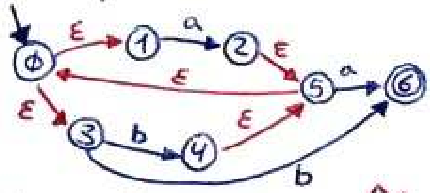
....



F = cjto que incluyen estados de aceptación en el NFA original

AUTÓMATAS FINITOS NO DETERMINISTAS (NFA) CON ε-TRANSICIONES

NFA-ε $M \equiv (Q, \Sigma, \delta, q_0, F)$ $\delta: Q \times (\Sigma \cup \{\varepsilon\}) \to 2^Q$



* ε-trans no ∃ en DFA

ε-CLAUSURA

$S = \varepsilon\text{-}cl(\emptyset) = \{\emptyset, 1, 3\}$

$\delta(S, a) = \{2\}$ $\varepsilon\text{-}cl(\{2\}) = \{2, 5, 0, 1, 3\} = B$

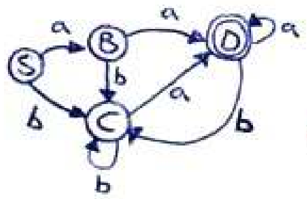
$\delta(S, b) = \{4, 6\}$ $\varepsilon\text{-}cl(\{4, 6\}) = \{4, 5, \emptyset, 1, 3\} = C$

$\delta(B, a) = \{6, 2\}$ $\varepsilon\text{-}cl(\{6, 2\}) = \{6, 2, 5, 0, 1, 3\} = D$

$\delta(B, b) = \{4, 6\}$ $\varepsilon\text{-}cl(\{4, 6\}) = C$

. . . .

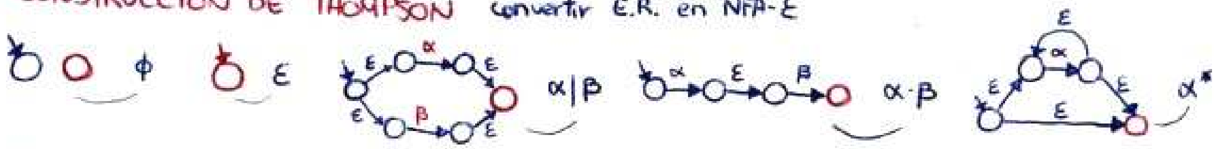
sin ε-trans ⇨



* $B \equiv C$ (indistinguibles)

CONSTRUCCIÓN DE THOMPSON

convertir E.R. en NFA-ε



$\emptyset$ ε $\alpha | \beta$ $\alpha \cdot \beta$ α^*

LEMA DE ARDEN

convertir NFA en E.R.

- $A_\emptyset = L(M)$
- $A_i = \emptyset$ (es posible)
- Si $q_i \in F$, $\varepsilon \in A_i$

$A_\emptyset = aA_1 \cup bA_2$
$A_1 = aA_3 \cup bA_2$
$A_2 = aA_3 \cup bA_2$
$A_3 = aA_3 \cup bA_2 \cup \varepsilon$

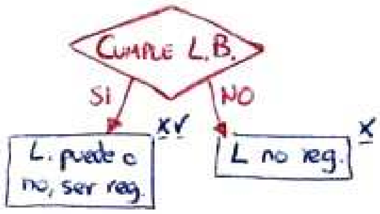
LEMA DEL BOMBEO

$\forall L \subseteq \Sigma^*,\ \exists n \in \mathbb{N} \mid$ si $w \in \Sigma^*$, $w \in L$ con $|w| \geq n$

a) $w = xyz$
b) $|xy| \leq n$
c) $|y| \geq 1$
d) $xy^iz \in L \ \ \forall i \geq \emptyset$

$L = \{a^m b^m \mid m \geq \emptyset\} \subseteq \{a, b\}^*$

$L = \{\varepsilon, ab, aabb, aaabbb, \ldots\}$



Sea $\underline{\underline{n}}$ cte del L.B. $|w| = 2n > n$ ⇒ ha de ocurrir

Sea $\underline{\underline{w}} = a^n b^n$

a) $w = a^n b^n = xyz$ ⟶ $\overbrace{aaa\ldots a}^{n}\overbrace{bbb\ldots b}^{n}$

b) $|xy| \leq n$ ⟶ xy solo contiene "a" ⇒ y solo contiene "a"

c) $|y| \geq 1$

por lo menos una "a"

d) $xy^iz \in L \forall i \geq \emptyset$ ⟶ $xy^0z \notin L$ ⇒ porque en $xy^0z = xz$ se han eliminado "a"'s (por lo menos una)

TEORIA DE AUTOMATAS Y LENGUAJES FORMALES (TALF)

GRAMATICAS Y LENGUAJES INDEPENDIENTES DEL CONTEXTO

$G \equiv (V, \Sigma, S, P)$

- Σ: cjto simb. terminales
- S: simb. arranque G $S \in V$
- V: cjto simb. no terminales $(V \neq \phi) \cap \Sigma = \phi$
- P: producciones de G $P \subseteq V \times (\Sigma \cup V)^*$

$S \to aB \mid AB$
$A \to aA \mid a$
$B \to bBa \mid b$

$G = \{\{S, A, B\}, \{a, b\}, S, P\}$

CLASIFICACIÓN CHOMSKY (GRAMATICAS)

	LENGUAJE	AUTOMATA	NORMAS PRODUCCION G.
G-0	R.Enum.	Máquina Turing	sin restricciones
G-1	dependient. contexto	-	-
G-2	indep. contex. (CFG)	PDA (a pilas)	$A \to \alpha$
G-3	regulares	autómata finito	$A \to Bu \mid v$

derivar
- $\Rightarrow$ en un paso
- $\Rightarrow^*$ en 0 o más pasos
- $\Rightarrow^+$ en 1 o más pasos
- $\Rightarrow^n$ en n pasos

DERIVACIONES

$S \Rightarrow aB \Rightarrow abBa \Rightarrow abba$

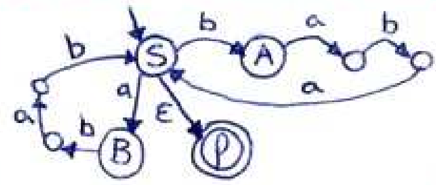
$S \Rightarrow^+ aaaAbBa \Rightarrow aaaabBa \Rightarrow aaaabba$

GRAMATICAS REGULARES

debe ser de algun tipo para ser reg.
- LINEAL POR LA DERECHA $A \to aB \mid b \mid \varepsilon$ $A, B \in V$ $a, b \in \Sigma^*$
- LINEAL POR LA IZQUIERDA $D \to Au \mid v \mid \varepsilon$ $A, D \in V$

CONVERSIÓN DE GRAMÁTICA A NFA (desde gram. lineal derecha)

$S \to aB \mid bA \mid \varepsilon$
$A \to abaS$
$B \to babS$



REGLAS

a) $(A \to aa\overset{n}{a}B) \in P \Rightarrow Q = Q \cup \{qq\overset{n-1}{q}\}$

$\delta(A, aaa) = \delta(q_1, aa) = \dots = \delta(q_{n-1}, aa) = B$

b) $(A \to aaa) \in P \Rightarrow Q = Q \cup \{qqq\}$

$\delta(A, aaa) = \delta(q, aa) = \dots = \delta(q, a) = p$

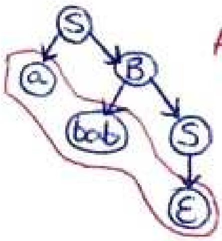
DESDE GRAM' LIN. IZQUIERDA

a) invertimos producciones $A \to Abbc \mid B \rightarrow A \to cbbA \mid B$

b) invertimos el autómata $q_0' \to p$, $p' \to q_0$, invertir transiciones (dirección)

GRAMÁTICAS INDEP. DEL CONTEXTO (CFG)

$S \Rightarrow aB \Rightarrow ababS \Rightarrow abab$



ARBOL DE ANÁLISIS SINTÁCTICO (AAS)

DERIVACIONES CANÓNICAS

$C \underset{mi}{\Rightarrow} DE \underset{mi}{\Rightarrow} dE \underset{mi}{\Rightarrow} de$ — más a la izquierda

$C \underset{mi}{\Rightarrow} DE \underset{md}{\Rightarrow} De \underset{md}{\Rightarrow} de$ ← no canónica

AMBIGÜEDAD

Gramática con más de un AAS distinto

GRAMATICAS INDEPENDIENTES DEL CONTEXTO (CFG)

SIMPLIFICACIÓN DE GRAMÁTICAS

- ELIMINACIÓN DE SIMBOLOS INUTILES
  (no participan en la derivación de las cadenas posibles)
- ELIMINACIÓN DE PRODUCCIONES VACÍAS
  (no-terminal produce epsilon: $A \to \epsilon$)
- ELIMINACIÓN DE PRODUCCIONES UNITARIAS
  (no-terminal produce un solo no-terminal: $A \to B$)

$\Sigma = \{a, b\}$

$S \to Aa \mid B \mid \cancel{X}$ b

$B \to b$ ← simplificar

$A \to aA \mid bA \mid B$ b

$\cancel{X} \to ab\cancel{X}$

orden del algoritmo:
- elim. simb. inutiles (RECOMENDADO!)
- elim. prod. vacias
- elim. prod. unitarias
- elim. simb. inutiles

RECURSIVIDAD EN GRAMATICAS $A \Rightarrow^{+} \alpha A \beta \quad A \in V$

(si derivando desde un simb. no-terminal A, obtenemos de nuevo A)

- RECURSIVA POR LA DERECHA $A \Rightarrow^{+} A\beta$
- RECURSIVA POR LA IZQUIERDA $A \Rightarrow^{+} \alpha A$

FORMAS NORMALES DE ESCRIBIR GRAMATICAS

FORMA NORMAL DE CHOMSKY (FNC)

$A \to a \quad (a \in \Sigma)$

$A \to BC \quad (A, B, C \in V)$

$S \to \epsilon \quad (\text{si } \epsilon \in L(G))$

(si no aparecen en parte dcha de ninguna prod.)

(sin producciones vacías)
(sin producciones unitarias)
(sin simbolos inutiles) } FNC

* Los AAS de G en FNC son arboles bin.
* Cualquier G puede escribirse en FNC.

FORMA NORMAL DE GREIBACH

$A \to a\alpha \quad (a \in \Sigma \quad \alpha \in V^{*})$

$S \to \epsilon \quad (\text{si } \epsilon \in L(G))$

* Si L(G) esta en FNG, $L(G) \neq \emptyset$
* Una G en FNG no es recursiva por izq.

LEMA DEL BOMBEO PARA LIC/CFL

Sea L un CFL $\mid \epsilon \notin L \quad \exists K \in \mathbb{N}$ Sea $n = |V|$ (nº de no terminales)

si $z \in L$, $|z| > K = 2^n$ $\quad S \in L \quad S \Rightarrow^{+} z$

a) $z = uvwxy$ (5 partes)

b) $|vwx| \leq K$

c) $uv^{i}wx^{i}y \in L \quad \forall i \geq 0$ (se bombean v y x)

d) $|vx| \geq 1$ (v y x no pueden ser ε, solo una si acaso)

CUMPLE L.B.
- SI → L puede o no ser CFL
- NO → L no CFL

el camino + largo raiz-hojas tiene más de n+2 nodos.

ALGORITMO DE COCKE-YOUNGER-KASAMI (CKY)

Determina si $w \in \Sigma^{*}$ puede ser generada por una CFG.

~ G en FNC

$S \to AB \mid BC$

$A \to BA \mid a$

$B \to CC \mid b$

$C \to AB \mid a$

$w = baaba \quad ¿x \in L(G)?$

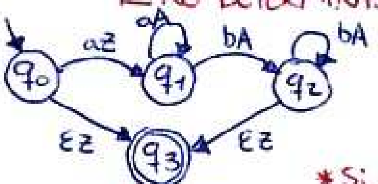
b	a	a	b	a
B	AC	AC	B	AC
AS	B			
	V_{ij}			

→ Si está el símbolo S ⇒ $S \Rightarrow^{+} w_{1n} \Rightarrow x \in L(G)$

TEORIA DE AUTOMATAS Y LENGUAJES FORMALES (TALF)

└── AUTÓMATAS CON PILAS (PDA) PDA $M \equiv (Q, \Sigma, \Gamma, \delta, q_\phi, \underset{\text{inicial pila}}{Z}, F)$

- DETERMINISTA resultado = (a,b) $\delta : Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma \to 2^{(Q \times \Gamma^*)} = P(Q \times \Gamma^*)$
- NO DETERMINISTA resultado = conjunto de (a,b) $\delta(q, a, b) = \{(p, w)\}$



$\Sigma = \{a, b\}$

$\Gamma = \{Z, A\}$

(ε,A)

δ	(a,Z)	(b,Z)	ε,Z	a,A	b,A
q_ϕ	$\{(q_1, AZ), (q_3, \varepsilon)\}$	ϕ	(q_3, ε)	ϕ	ϕ
q_1	ϕ	ϕ	ϕ	(q_1, AA)	(q_2, ε)
q_2	ϕ	ϕ	(q_3, Z)	ϕ	(q_2, ε)
q_3	ϕ	ϕ	ϕ	ϕ	ϕ

* Si no hay trans. el PDA no sabe que hacer.

* $\delta(q, \varepsilon, \underset{\text{elimina de la pila}}{A}) = \{(p, \underset{\text{se añade a la pila}}{AA})\}$

- DESCRIPCIÓN INSTANTÁNEA (D.I.) $DI \equiv (q, w, u) \in Q \times \Sigma^* \times \Gamma^*$

$(q_\phi, aabb, Z) \vdash (q_1, abb, AZ) \vdash (q_1, bb, AAZ) \vdash (q_2, b, AZ) \vdash (q_2, \varepsilon, Z)$

- CRITERIOS DE ACEPTACIÓN
  - POR ESTADO DE ACEPTACIÓN $L(M)$
  - POR PILA VACIA $N(M)$

- CONVERSION PDA $L(M)$ A PDA $N(M')$ $M \equiv (Q', \Sigma, \Gamma', \delta', q_\phi', F', Z')$

$\delta'(q_\phi', \varepsilon, Z') = \{(q_\phi, ZZ')\}$

$\delta'(q, a, X) = \delta(q, a, X)$ mismas trans. que δ

$\delta'(p, \varepsilon, A) = \{(q_s', A)\}$ $(\forall p \in F)$

$\delta'(q_s', \varepsilon, A) = \{(q_s', \varepsilon)\}$ $(\forall A \in \Gamma')$

$Q' = Q \cup \{q_0', q_s'\}$ $(q_\phi', q_s' \notin Q)$

$\Gamma' = \Gamma \cup \{Z'\}$ $(Z' \notin \Gamma)$

$F' = \phi$

q_ϕ' = new

Z' = new

- CONVERSIÓN PDA $N(M)$ A PDA $L(M')$ $M \equiv (Q', \Sigma, \Gamma', \delta', q_\phi', F', Z')$

$\delta'(q_\phi', \varepsilon, Z') = \{(q_\phi', ZZ')\}$

$\delta'(q, a, X) = \delta(q, a, X)$ mismas trans. que δ

$\delta'(p, \varepsilon, Z') = \{(q_F', Z')\}$

$Q' = Q \cup \{q_\phi', q_F'\}$ $(q_\phi, q_F' \notin Q)$

$\Gamma' = \Gamma \cup \{Z'\}$ $(Z' \notin \Gamma)$

$F' = q_F'$

q_ϕ' = new

Z' = new

- CONVERSIÓN DE G A PDA NO DETERMINISTA $\Gamma = V \cup \Sigma \cup \{Z\}$ $(Z \notin (\Sigma \cup V))$

SIEMPRE ① $\delta(q_\phi, \varepsilon, Z) = \{(q_1, SZ)\}$

SI ENCUENTRAS NO-TERM ② $\delta(q_1, \varepsilon, A) = \{(q_1, \alpha)\}$ $(A \to \alpha) \in P$ $\forall A \in V$

SI ENCUENTRAS TERMINAL ③ $\delta(q_1, a, a) = \{(q_1, \varepsilon)\}$ $\forall a \in \Sigma$

SIEMPRE ④ $\delta(q_1, \varepsilon, Z) = \{(q_2, \varepsilon)\}$

TEORIA DE AUTÓMATAS Y LENGUAJES FORMALES (TALF) 7

MAQUINAS DE TURING

MT $M \equiv (Q, \Sigma, \delta, q_0, F, \Gamma, \not b)$

$\Sigma \subseteq (\Gamma - \{\not b\})$

$\not b \in \Gamma$, $\not b \notin \Sigma$

DEFINICIÓN

$\delta : Q \times \Gamma \to Q \times \Gamma \times \{L, R\}$

$\delta(q, a) = (p, b, \underline{X})$ → Left, Right

$M \equiv (\{q_0, q_1, q_2, q_3, q_4\}, \{\emptyset, 1\}, \delta, \{q_4\}, \{\emptyset, 1, \not b\}, \not b)$

DIAGRAMA DE ESTADOS

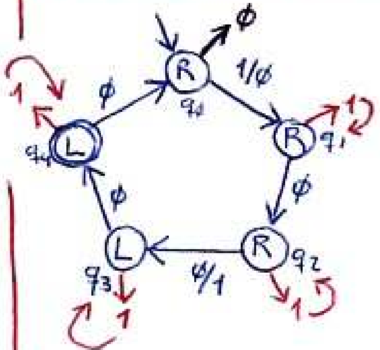
$\delta(q_0, 1) = (q_1, \emptyset, R)$

$\delta(q_1, \emptyset) = (q_2, \emptyset, R)$

$\delta(q_2, \emptyset) = (q_3, 1, L)$

$\delta(q_3, \emptyset) = (q_4, \emptyset, L)$

$\delta(q_4, \emptyset) = (q_0, \emptyset, R)$



* si no existe trans. en el diag. de estados ~~es~~ que no cambia de estado

$(q_0, \underline{1}00011) \vdash (q_1, \emptyset\underline{0}0011) \vdash (q_1, \emptyset\emptyset\underline{0}011) \vdash \ldots$

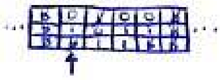
ESQUEMA DE LENGUAJES

L.R.E → L.R. → L.I.C. → L. Reg.

L.R.E: { MT para y acepta / MT no se detiene (∞)

L.R.: MT para y acepta / MT para y rechaza

TIPOS DE MAQUINAS DE TURING

- MT CON CAPACIDAD DE NO MOVERSE — $\delta : Q \times \Gamma \to Q \times \Gamma \times \{L, R, S\}$ (S → Stop)
- MT MULTIPISTA  — $\delta : Q \times \Gamma^k \to Q \times \Gamma^k \times \{L, R, S\}$
- MT CON CINTA ∞ EN UN SENTIDO
- MT MULTICINTA — $\delta : Q \times \Gamma^n \to Q \times \Gamma^n \times \{L, R, S\}^n$
- MT CON CINTA MULTIDIMENSIONAL — $\delta : Q \times \Gamma \to Q \times \Gamma \times \{L, R, S, U, D\}$ (2 dimensiones)
- MT NO DETERMINISTA — δ no es una $f(x)$ $\delta(q, a) = \{(p, a, S), (r, b, X)\}$

MAQUINAS DE TURING UNIVERSALES (MT-U/M_U)

MT-U $M \equiv (Q, \Sigma, \delta, q_0, F, \Gamma, \not b)$

estados, simbolos, movimientos

$q_i \equiv 1^i$, $a_j \equiv 1^j$, $L \equiv 1$, $R \equiv 11$, $S \equiv 111$, $a_1 \equiv \not b \equiv 1$

$|F| = 1$, $F = \{q_n\}$

$\Gamma_1 = \{\not b\}$

CONVERSIÓN DE DFA A MT

DFA M, MT M'

$Q' = Q \cup \{q'_f\}$, $q_0' = q_0$

$F' = \{q'_f\}$

$\delta'(q, a) = (\delta(q, a), a, R)$

$\delta'(q, \not b) = (q'_f, \not b, \underline{X})$ → R, L, S…

CONVERSIÓN DE PDA A MT MULTICINTA (entrada y pila)