TEORÍA DE AUTÓMATAS Y LENGUAJES FORMALES

Francisco Vico

departamento
Lenguajes y
Ciencias de la Computación

área de conocimiento
Ciencias de la Computación e
Inteligencia Artificial

ETSI Informática
Universidad de Málaga

fjvico@uma.es
geb.uma.es/fjv

10 de diciembre de 2015

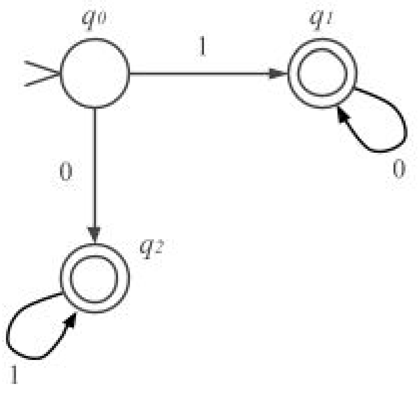
Notation

ℕ	the set of nonnegative integers (or natural numbers), i.e., {0, 1, 2, … } (U+2115)
ℙ	the set of positive integers (U+2119)
ℝ	the set of real numbers (U+211D)
ℤ	the set of integers (U+2124)
∅	the empty set (U+2205)
⊆	the (infix) subset relation between sets (U+2286)
⊂	the (infix) proper subset relation between sets (U+2282)
∪	the infix union operation on sets (U+222A)
∩	the infix intersection operation on sets (U+2229)
~	the prefix complementation operation on sets (U+007E)
−	the infix set difference operation on sets (U+2212)
×	the infix cartesian product of sets (U+00D7)
A^n	the postfix n-fold cartesian product of A, i.e. A× … ×A (n times)
2^A	the powerset of A

1. En términos de conjuntos, proponer ejemplos de los distintos conceptos definidos en este tema. Dados $A = \{1 \dots 5\}$ y $B = \mathbb{N}$:
   a. Encontrar una partición de A.
   b. Encontrar una partición de B.
   c. Encontrar una relación entre A y B.
   d. Encontrar una función entre A y B.
   e. Encontrar una relación binaria sobre A.

2. Representar estos tres lenguajes en forma gramatical:
   a. $ab^n,\ n \in \mathbb{N}$
   b. $\{a, aa, aaa\}$
   c. $a^n b^{n-1},\ n \in \mathbb{N} = \{a, aab, aaabb, \ldots\}$

3. Razonar si es regular el lenguaje $\{ x \in \Sigma^* \mid |x|_0 = |x|_1 \}$.
4. Definir un AFD que represente el lenguaje $1(0+1)^*0$.
5. ¿ $\Sigma^+ = \Sigma^*$?
6. Sea $L = \{0\}\cdot\{1\}^*$
   a. ¿ qué operaciones definen este lenguaje ?
   b. ¿ es cierto que $\{ n \in \mathbb{N} \mid \exists\ w \in L, |w|_1 = n \} = \mathbb{N}$?
   c. ¿ es cierta la misma expresión para $L = \{0\}\cdot\{1\}^m$, con $m \in \mathbb{N}$?
7. ¿ $\| L((\{A\}, \Sigma, \{A \rightarrow 1 \mid 0A\}, A)) - 00^*1 \| \in \mathbb{N}$?
8. $M = (K, \Sigma, R, s, F)$ es un AF dado por el diagrama:



   a. ¿ M es un AFD o un AFND ?
   b. Define una expresión regular que represente a $L(M)$
   c. Determina $\| L(M) \cap 1^* \|$
   d. Escribe una computación completa para la cadena 100
   e. Si $M' = (K \cup \{q_3\}, \Sigma, R \cup \{(q_1, 1, q_3), (q_2, 0, q_3)\}, s, F)$, ¿ $M' \equiv M$?
   f. ¿ Qué longitud tiene la computación que acepta 01^n ?

   g. ¿ En cuántos pasos se completa la transición $(q_0, 100^n) \vdash^* (q_1, 0)$, con $n \in \mathbb{N}$?

9. Dada la gramática $G = (\{ A, B \}, \Sigma, \{ A \to 10, A \to B0, B \to 1A \}, A)$
   a. ¿ $11110000 \in L(G)$?
   b. ¿ De qué tipo es G ?
   c. ¿ De qué tipo es $L(G)$?
   d. Define matemáticamente $L(G)$.
   e. Deduce qué longitud tiene la derivación que produce 1^n0^n y demuestra
      $$A \Rightarrow^{l(n)} 1^n0^n \ , \ \forall n \in \mathbb{N}$$
   f. ¿ Cúal de estas expresiones representa a $L(G)$: 1^*0^*, 11^*00^*, $1(10)^*0$?
   g. ¿ Es un lenguaje $\Sigma^* - L(G)$?
10. Responde brevemente a estas preguntas.
   a. ¿ Cuántas expresiones regulares pueden formularse ?
   b. ¿ Todo lenguaje de contexto libre es representable con expresiones regulares ?
   c. ¿ Una expresión regular es una cadena ?
   d. ¿ ER es un lenguaje ?
   e. ¿ En qué operaciones sobre lenguajes se fundamentan las expresiones regulares ?
   f. ¿ Por qué razón el AFD se denomina "determinista" ?
   g. ¿ Qué cardinal tiene L(AFD) ?
   h. ¿ En qué se relaciona L.3 con L(AFD) ?
   i. ¿ Todo AFD es AFND ?
11. Representa en el modelo de AFND estos lenguajes
   a. 01*
   b. (0+1)*
   c. (0+1)
   d. (001+110)*+010
12. Encuentra una cadena distinguible con 0011 respecto a $\{ x \in \Sigma^* \mid |\, x\,|_0 = |\, x\,|_1 \}$ y otra indistinguible. Justifica los ejemplos.
13. Dado un lenguaje L y dos cadenas x e y sobre el mismo alfabeto de L, ¿cuáles de estas afirmaciones son ciertas?
   a. $x^2 \in L \land y^2 \notin L \land \mathrm{x} \neq \mathrm{y} \Rightarrow x$ e y son indistinguibles respecto a L
   b. $x^2 \in L \land yx \notin L \Rightarrow x$ e y son distinguibles respecto a L
   c. $xz \in L \land yz \notin L \land z \in \Gamma^* \land \Sigma \neq \Gamma \Rightarrow x$ e y son distinguibles respecto a L
14. ¿Es cierta esta expresión: $(v, w), (x, y) \in I_L \Rightarrow (x, w) \in I_L$? Razona la respuesta.
15. Demostrar que $L = \{ x \in \Sigma^* \mid |\, x\,|_0 = |\, x\,|_1 \}$ no es un lenguaje regular por el método del bombeo.
16. Sea $\Sigma = \{a, b, c, d\}$. Demostrar, usando alternativamente el teorema de Myhill-Nerode y el lema de bombeo, que el lenguaje $L = \{ a^nb^mc^nd^m \mid n, m \geq 0 \}$ no es regular.

Soluciones propuestas

1. En términos de conjuntos, proponer ejemplos de los distintos conceptos definidos en este tema. Dados $A = \{1 \dots 5\}$ y $B = \mathbb{N}$:
   a. Encontrar una partición de A : $\{ \{1, 2\}, \{3, 4\}, \{5\} \}$
   b. Encontrar una partición de B : $\{ \{ X \in B \mid X \leq 10 \}, \{ X \in B \mid X > 10 \} \}$
   c. Encontrar una relación entre A y B : $\{ (1, 10), (2, 20), (2, 2), (10, 20) \}$
   d. Encontrar una función entre A y B : $\{ (X, Y) \in A \times B \mid Y = 10\, X \}$
   e. Encontrar una relación binaria sobre A : $\{ (X, Y) \in A^2 \mid Y = 2\, X \}$

2. Representar estos tres lenguajes en forma gramatical:
   a. $ab^n,\ n \in \mathbb{N}$

   $N = \{ A, B \}$
   $T = \{ a, b \}$
   $P = \{ \quad A \rightarrow aB$
   $\qquad\quad B \rightarrow bB \mid b \ \}$
   $S = A$

   b. $\{a, aa, aaa\}$

   $N = \{ A \}$
   $T = \{ a, b \}$
   $P = \{ A \rightarrow a \mid aa \mid aaa \}$
   $S = A$

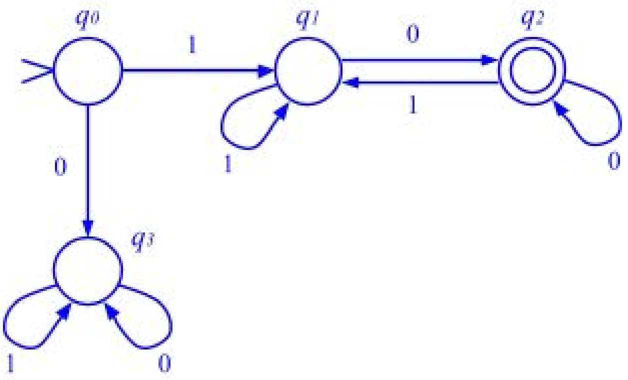
   c. $a^n b^{n-1},\ n \in \mathbb{N} = \{a, aab, aaabb, \dots\}$

   $N = \{ A, B \}$
   $T = \{ a, b \}$
   $P = \{ \quad A \rightarrow aB$
   $\qquad\quad B \rightarrow aBb \mid \varepsilon \}$
   $S = A$

3. Razonar si es regular el lenguaje $\{ x \in \Sigma^* \mid |x|_0 = |x|_1 \}$.

No, porque los símbolos en las cadenas de un lenguaje regular se generan independientemente, no podríamos hacer que coincidieran “siempre” el número de 0s y de 1s.

4. Definir un AFD que represente el lenguaje $1(0+1)^*0$.



5. ¿ $\Sigma^+ = \Sigma^*$?

En este caso, no son iguales los cierres estricto y amplio, difieren en { ε }.

6. Sea $L = \{ 0 \} \cdot \{ 1 \}^*$.
   a. ¿ Qué operaciones definen este lenguaje ?

      Cierre de lenguajes y concatenación de lenguajes.

   b. ¿ Es cierto que $\{ n \in \mathbb{N} \mid \exists\ w \in L, |w|_1 = n \} = \mathbb{N}$?

      Sí, hay cadenas con todos los números de unos posibles.

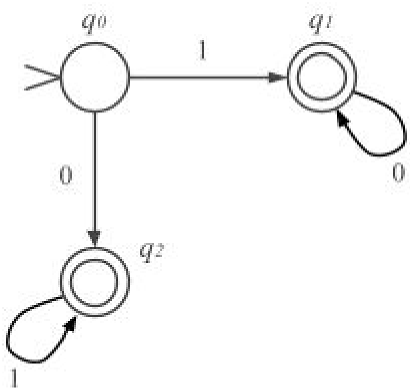
   c. ¿ Es cierta la misma expresión para $L = \{ 0 \} \cdot \{ 1 \}^m$, con $m \in \mathbb{N}$?

      No, en este caso hay un número finito de cadenas en *L*.

7. ¿ $\| L ((\{ A \}, \Sigma, \{ A \to 1 \mid 0A \}, A)) - 00^*1 \| \in \mathbb{N}$?

Es el lenguaje { 1 } y su cardinalidad es 1.

8. $M = (K, \Sigma, R, s, F)$ es un AF dado por el diagrama:



a. ¿ M es un AFD o un AFND ?

Es un AFND, puede realizar computaciones bloqueadas.

b. Define una expresión regular que represente a $L(M)$

$10^{*}+01^{*}$

c. Determina $\| L(M) \cap 1^{*} \|$

El lenguaje resultante es $\{1\}$, por tanto, su cardinalidad es 1.

d. Escribe una computación completa para la cadena 100

$((q_0, 100), (q_1, 00), (q_1, 0), (q_1, \varepsilon))$

e. Si $M' = (K \cup \{q_3\}, \Sigma, R \cup \{ (q_1, 1, q_3), (q_2, 0, q_3)\}, s, F)$, ¿ $M' \equiv M$?

Sí, el nuevo estado es un estado pozo.

f. ¿ Qué longitud tiene la computación que acepta 01^n ?

Tiene longitud $n + 1$; en general, una cadena $w \in \Sigma^{*}$ se acepta en $| w |$ pasos.

g. ¿ En cuántos pasos se completa la transición $(q_0, 100^n) \vdash^{*} (q_1, 0)$, con $n \in \mathbb{N}$?

Se completa en $n + 1$ pasos.

9. Dada la gramática $G = (\{ A, B \}, \Sigma, \{ A \rightarrow 10, A \rightarrow B0, B \rightarrow 1A \}, A)$
   a. ¿ $11110000 \in L(G)$?

   Sí, la genera la derivación

   $A \Rightarrow B0 \Rightarrow 1A0 \Rightarrow 1B00 \Rightarrow 11A00 \Rightarrow 11B000 \Rightarrow 111A000 \Rightarrow 11110000$.

b. ¿De qué tipo es G ?

Es de tipo 2 y no es de tipo 3.

c. ¿De qué tipo es $L(G)$?

Es de tipo 2 y no es de tipo 3.

d. Define matemáticamente $L(G)$

$\{ w \in \Sigma^* \mid w = 1^n0^n , n \in \mathbb{N}\}$

e. Deduce qué longitud tiene la derivación que produce 1^n0^n y demuestra

$$A \Rightarrow^{l(n)} 1^n0^n , \forall n \in \mathbb{N}$$

La longitud es $2n - 1$.
Demostramos $A \Rightarrow^{2n\text{-}1} 1^n0^n$, $\forall n \in \mathbb{N}$ por inducción sobre n
C.B.: $A \Rightarrow 10$, por aplicación de la regla $A \rightarrow 10$
H.I.: $A \Rightarrow^{2n} 1^n0^n$
P.I.: $A \Rightarrow^{2n+1} 1^{n+1}0^{n+1}$
para que se verifique la H.I., $A \Rightarrow^{2n\text{-}1} 1^n0^n$, las dos últimas producciones deben ser $A \Rightarrow^{2n\text{-}1} 1^{n\text{-}1}A0^{n\text{-}1} \Rightarrow 1^n0^n$ (esta última, por aplicación de la regla $A \rightarrow 10$), por tanto, tras $n - 1$ producciones se pueden aplicar, en este orden, las reglas $A \rightarrow B0$, $B \rightarrow 1A$ y $A \rightarrow 10$, obteniendo la derivación
$A \Rightarrow^{2n\text{-}1} 1^{n\text{-}1}A0^{n\text{-}1} \Rightarrow 1^{n\text{-}1}B0^n \Rightarrow 1^nA0^n \Rightarrow 1^{n+1}0^{n+1}$ en $2(n - 1) + 3$ pasos, es decir, la derivación tienen una longitud $2n$+1.

f. ¿Cuál de estas expresiones representa a $L(G)$: 1^*0^*, 11^*00^*, $1(10)^*0$?

Ninguna, son expresiones regulares y el lenguaje no es regular.

g. ¿Es un lenguaje $\Sigma^* - L(G)$?

Sí, es el complementario de $L(G)$.

10. Responde brevemente a estas preguntas.

a. ¿Cuántas expresiones regulares pueden formularse?

Infinitas numerables ($\aleph_0$).

b. ¿Todo lenguaje de contexto libre es representable con expresiones regulares?

No, sólo los que también son regulares.

c. ¿Una expresión regular es una cadena?

Sí, es una cadena sobre lo que llamamos un meta-alfabeto.

d. ¿ER es un lenguaje?

Sí, es un meta-lenguaje.

e. ¿En qué operaciones sobre lenguajes se fundamentan las expresiones regulares?

Cierre, concatenación y unión de lenguajes.

f. ¿Por qué razón el AFD se denomina “determinista” ?

Porque su comportamiento está determinado siempre.

g. ¿Qué cardinal tiene *L*(AFD)?

Infinito numerable ($\aleph_0$).

h. ¿En qué se relaciona *L*.3 con *L*(AFD)?

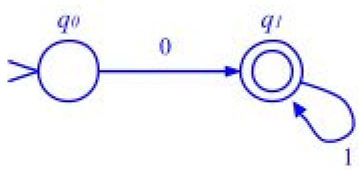
L.3 = *L*(AFD)

i. ¿Todo AFD es AFND?

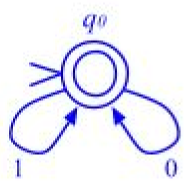
Sí, pues toda función es relación.

11. Representa en el modelo de AFND estos lenguajes

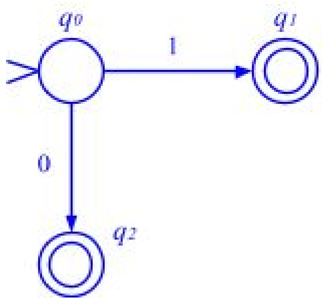
a. 01*



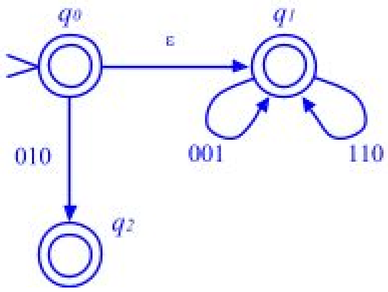
b. (0+1)*



c. (0+1)



d. (001+110)*+010



12. Encuentra una cadena distinguible con 0011 respecto a $L = \{ x \in \Sigma^* \mid |x|_0 = |x|_1 \}$ y otra indistinguible. Justifica los ejemplos.

0011 y 0 son distinguibles respecto a L porque concatenando la cadena 1 como sufijo hace que $00111 \notin L$ y $01 \in L$.

0011 y 01 son indistinguibles respecto a L porque cualquier cadena w que se concatene como sufijo hace que ambas pertenezcan (si $w \in L$) o no pertenezcan a L (si $w \notin L$).

13. Dado un lenguaje L y dos cadenas x e y sobre Σ, ¿cuáles de estas afirmaciones son ciertas?

a. $x^2 \in L \wedge y^2 \notin L \wedge x \neq y \Rightarrow$ x e y son distinguibles respecto a L

Es falsa, ya que no estamos concatenando la misma cadena.

b. $x^2 \in L \wedge yx \notin L \Rightarrow$ x e y son distinguibles respecto a L

Es cierta, pues existe una cadena (x) que, concatenada como sufijo, hace que una pertenezca a L y la otra no pertenezca.

c. $xz \in L \wedge yz \notin L \wedge z \in \Gamma^* \wedge \Sigma \neq \Gamma \Rightarrow$ x e y son distinguibles respecto a L

Es cierta, aunque z sea una cadena sobre un alfabeto diferente, es obvio que no contiene símbolos que no pertenezcan a Σ, pues xz es cadena de L.

14. ¿Es cierta esta expresión: $(v, w), (x, y) \in I_L \Rightarrow (x, w) \in I_L$? Razona la respuesta.

No es cierta, la indistinguibilidad de cadenas es transitiva, pero no se aplica aquí.

15. Demostrar que $L = \{ x \in \Sigma^* \mid |x|_0 = |x|_1 \}$ no es un lenguaje regular por el método del bombeo.

Tomemos como n el número de ceros ($|w|_0 = n$) de la cadena, por lo que $|w| = 2n$. Entonces se cumple:

1) $x = uvw$
2) $|uv| \leq n$
3) $|v| > 0$
4) $\forall m \geq 0\ uv^m w \in L$

Podemos elegir v de forma que cumpla $0 < |v| \leq n \wedge |v|_0 \neq |v|_1$, entonces, para cualquier valor de $m \neq 1$, no se cumple que $uv^m w \in L$. Por ejemplo: $x = 0110$ y $n = 2$, podemos asignar $u = 0$, $v = 1$, $w = 10$ y vemos que $uv^0w =$ 010 ó $uv^2w =$ 01110 no cumplen la condición 4).

16. Sea $\Sigma = \{a, b, c, d\}$. Demostrar, usando alternativamente el teorema de Myhill-Nerode y el lema de bombeo, que el lenguaje $L = \{ a^n b^m c^n d^m \mid n, m \geq 0 \}$ no es regular.

Por Myhill-Nerode: Sea el conjunto de cadenas $S = a^*$. Todas las cadenas de S son distinguibles dos a dos: sean $x = a^p$, $y = a^q \in S$, $p \neq q$, entonces la cadena $z = bc^p d$ es tal que $xz \in L$ pero $yz \notin L$. Por tanto x e y son distinguibles $\forall x,y \in S$. Esto significa que cada una de estas cadenas está en una clase de equivalencia distinta de la relación de indistinguibilidad I_L. Luego existen infinitas clases de equivalencia. Por el teorema de Myhill-Nerode, esto quiere decir que L no es regular.

Por bombeo: Sea $n \in \mathbf{N}$, y sea $x = a^n bc^n d \in L$, $|x| = 2n+2 \geq n$. Supongamos que $\exists\ u, v, w \in \Sigma^*$ que cumplen conjuntamente todas las condiciones del lema del bombeo para lenguajes regulares. Entonces:

1. $x = uvw$
2. $|uv| \leq n \Rightarrow$ (por 1) $uv = a^k$ con $0 \leq k \leq n \Rightarrow$ (por 1) $w = a^{n-k} bc^n d$
3. $|v| > 0 \Rightarrow$ (por 2) $v = a^j$, $u = a^{k-j}$ con $0 < j < k$,
4. $\forall m \geq 0$, $uv^m w \in L \Rightarrow$ Si $m=0$, $uw \in$ L. Pero $uw = a^{k-j} a^{n-k} bc^n d = a^{k-j+n-k} bc^n d = a^{n-j} bc^n d \notin L$ porque j>0. Por lo tanto hemos llegado a un absurdo, de donde L no es regular.