Tema 2

Lenguajes y Operaciones con lenguajes

Definición de lenguaje

- Un lenguaje sobre Σ es un subconjunto de Σ^*
- L es un lenguaje sobre Σ si $L \subseteq \Sigma^*$
- ¿Cuántos lenguajes hay para un Σ dado?

¿Cuántos lenguajes hay por alfabeto?

- Sea m_0 … m_n la lista de cadenas finitas formadas con símbolos de Σ
- Sea S_0 … S_n la lista de todos los subconjuntos (languajes) en Σ^*

	m_0	m_1	m_2	…	m_n
S_0	0	0	0	…	0
S_1	0	0	1	…	0
S_2	1	0	1	…	0
…	…	…	…	…	…
S_n	1	1	1	…	1

El conjunto diagonal: $D(i) = S_i(i)$

	m_0	m_1	m_2	…	m_n
S_0	0	0	0	…	0
S_1	0	0	1	…	0
S_2	1	0	1	…	0
…	…	…	…	…	…
S_n	1	1	1	…	1

$$D = \{m_2, \ldots, m_n\}$$

El conjunto *anti*-diagonal

	m_0	m_1	m_2	…	m_n
S_0	1	0	0	…	0
S_1	0	1	1	…	0
S_2	1	0	0	…	0
…	…	…	…	…	…
S_n	1	1	1	…	0

$$\overline{D} = \{m_0, m_1, \ldots\}$$

¿Es el conjunto *anti-diagonal* un lenguaje?

- $\overline{D}$ no está en la lista S_0 … S_n
  - Difiere de S_0 en el primer elemento
  - Difiere de S_1 en el segundo elemento
  - …
  - Difiere de S_n en el elemento $n + 1$
- El conjunto potencia de Σ^* no es contable!
- 2^{Σ^*} no es contable!
- Hay más lenguajes de los que nos podemos contar! (¿imaginar?)

Operaciones de conjuntos sobre lenguajes

- Si $L_1, L_2 \subseteq \Sigma^*$
  - Unión: $L_1 \cup L_2$ es un lenguaje
  - Intersección: $L_1 \cap L_2$ es un lenguaje
  - Diferencia: $L_1 - L_2$ es un lenguaje
- Si $L_1 \subseteq \Sigma_1^*$ y $L_2 \subseteq \Sigma_2^*$ entonces L_1 y L_2 son subconjuntos de $(\Sigma_1 \cup \Sigma_2)^*$:
  - $\overline{L_1} = \Sigma_1^* - L_1$
  - $\overline{L_1} = (\Sigma_1 \cup \Sigma_2)^* - L_1$

- Concatenación de lenguajes
  - Si $L_1, L_2 \subseteq \Sigma^*$ la concatenación de L_1 y L_2, L_1L_2 es

    $L_1L_2 = \{xy \mid x \in L_1 \ \& \ y \in L_2\}$
- Ejemplo:
  - $L_1 = \{\text{homo, dis}\}$
  - $L_2 = \{\text{forme, funcional}\}$
  - $L_1L_2 = \{\text{homoforme, homofuncional, disforme, disfuncional}\}$
- Concatenación con $\{\Lambda\}$:
  - $\{\Lambda\}L = L\{\Lambda\} = L$
  - $\{\Lambda\}$ es la identidad de concatenación de lenguajes

Concatenación de Lenguajes

funcional	homofuncional	disfuncional
forme	homoforme	disforme
L_2 \ L_1	homo	dis

$$L_1L_2 = \{xy \mid x \in L_1 \ \& \ y \in L_2\}$$

"Potencias" de un lenguaje

- La concatenación de un lenguaje consigo mismo:
  - $L^0 = \{\Lambda\}$
  - $L^1 = L$
  - $L^2 = L^1L^1$
  - $L^3 = L^2L^1$

    …
  - $L^n = L^{n-1}L^1$
- En particular:
  - $L^1 = L^0L^1$

Identidad de concatenación para lenguajes

Λ	homoΛ	disΛ
L_2 \ L_1	homo	dis

⇒

Λ	homo	dis
L_2 \ L_1	homo	dis

Cerraduras de los lenguajes

- Cerradura de Kleen (Kleen-star) L^*:

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

- Cerradura de Kleen (Kleen-plus) L^+:

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

Cerraduras de los lenguajes

- Operación Kleen-star L^*:
  - $L = \{0, 1\}$ ($\Sigma = \{0, 1\}$ but $\Sigma \neq L$)
  - $L^* = L^0 \cup L^1 \cup L^2 \cup \ldots$
  - $L^* = \{0, 1\}^* = \{\Lambda, 0, 1, 00, 01, 10, 11, \ldots\}$
- Operación Kleen-plus L^+:
  - $L = \{0, 1\}$
  - $L^+ = L^1 \cup L^2 \cup \ldots$
  - $L^+ = \{0, 1\}^+ = \{0, 1, 00, 01, 10, 11, \ldots\}$

Cerraduras de los lenguajes

- El lenguaje vacío: Φ
- El lenguaje con la cadena vacía: $\{\Lambda\}$
- Lenguajes finitos sin Λ
- Lenguajes finitos con Λ
- Lenguajes infinitos (denumerables) sin Λ
- Lenguajes infinitos (denumerable) con Λ

Notación para las potencias

- Si $a \in \Sigma, x \in \Sigma^*$ & $L \subseteq \Sigma^*$
  - $a^k = aa\ldots a$ k veces
  - $x^k = xx\ldots x$ k veces
  - $\Sigma^k = \Sigma\Sigma\ldots\Sigma$ k veces
  - $L^k = LL\ldots L$ k veces
- Para el caso de $k = 0$
  - $\left.\begin{array}{l} a^0 = \Lambda \\ x^0 = \Lambda \end{array}\right\}$ Ident. de concatenación de cadenas
  - $\left.\begin{array}{l} \Sigma^0 = \{\Lambda\} \\ L^0 = \{\Lambda\} \end{array}\right\}$ Ident. de concatenación de lenguajes

Cerraduras de Φ

- Expansión of las potencias:
  - $\Phi^0 = \{\Lambda\}$
  - $\Phi^1 = \Phi$ (no tiene elementos)
  - $\Phi^2 = \Phi\,\Phi = \Phi$
  - $\Phi^3 = \Phi\,\Phi^2 = \Phi\,\Phi = \Phi$
  - ...
- Cerraduras:
  - $\Phi^* = \Phi^0 = \{\Lambda\}$
  - $\Phi^+ = \Phi^1 = \Phi$ (no tiene elementos)

Cerraduras de $\{\Lambda\}$

- Expansión of las potencias:
  - $L^0 = \{\Lambda\}^0 = \{\Lambda\}$
  - $L^1 = L = \{\Lambda\}$
  - $L^2 = L^1L^1 = \{\Lambda\}\{\Lambda\} = \{\Lambda\Lambda\} = \{\Lambda\}$
  - $L^3 = L^1L^2 = \{\Lambda\}\{\Lambda\} = \{\Lambda\Lambda\} = \{\Lambda\}$
  - ...
- Cerraduras:
  - $\{\Lambda\}^* = \{\Lambda\}$
  - $\{\Lambda\}^+ = \{\Lambda\}$
  - Si $L = \{\Lambda\}$ entonces $L = L^* = L^+$

Cerraduras de lenguajes finitos sin Λ

- Sea $L = \{0, 1\}$ ($\Sigma = \{0, 1\}$ pero $\Sigma \neq L$)
  - $L^0 = \{\Lambda\}$
  - $L^1 = L = \{0, 1\}$
  - $L^2 = L^1L^1 = \{0,1\}\{0,1\} = \{00,01,10,11\}$
  - $L^3 = L^1L^2 = \{0,1\}\ \{00,01,10,11\} = \{000, 001,010,011,100,101,110,111\}$
- Cerraduras:
  - $L^* = \{\Lambda,0,1,00,01,10,11,000,001,010,011,\ldots\}$
  - $L^+ = \{0,1,00,01,10,11,000,001,010,011,\ldots\}$

Cerraduras de lenguajes finitos sin Λ

- Sea $L = \{0, 11\}$ (i.e. dos palabras) $(\Sigma = \{0, 1\})$
  - $L^0 = \{\Lambda\}$
  - $L^1 = L = \{0, 11\}$
  - $L^2 = L^1L^1 = \{0, 11\}\{0, 11\} = \{00, 011, 110, 1111\}$
  - $L^3 = L^1L^2 = \{0, 11\}\{00, 011, 110, 1111\}$
    $= \{000, 0011, 0110, 01111, 1100, 11011, 11110, 111111\}$
- L^n es el conjunto de cadenas que resulta de la concatenación de *n cadenas* o *palabras* de L (NO de los símbolos de Σ)
- Cerraduras:
  - $L^* = \{\Lambda, 0, 11, 00, 011, 110, 1111, 000, 0011, 0110, 01111 \ldots\}$
  - $L^+ = \{0, 11, 00, 011, 110, 1111, 000, 0011, 0110, 01111 \ldots\}$

Computación gráfica de L^{n+1}

L^2

1111	01111	111111	
110	0110	11110	
011	0011	11011	
00	000	1100	
	0	11	L

$L^3 = L\,L^2$

¡La concatenación de un lenguaje con sus potencias es conmutativa!

L

11	0011	01111	11011	111111	
0	000	0110	1100	11110	
	00	011	110	1111	L^2

$L^2\,L = L\,L^2$

Todo lenguaje conmuta con su cerradura

- $L^+ = L\,L^* = L^*\,L$
  - $L^* = L^0 + L^1 + L^2 + \ldots$
  - $L^+ = L^1 + L^2 + \ldots$
- $L\,L^* = L(L^0 + L^1 + L^2 + \ldots)$
  $= L(\{\Lambda\} + L^1 + L^2 + \ldots)$
  $= L\{\Lambda\} + LL^1 + LL^2 + \ldots$
  $= L + L^2 + L^3 + \ldots$
  $= L^+$
- $L^*\,L = (L^0 + L^1 + L^2 + \ldots)\,L$
  $= (\{\Lambda\} + L^1 + L^2 + \ldots)\,L$
  $= \{\Lambda\}L + L^1L + L^2L + \ldots$
  $= L + L^2 + L^3 + \ldots$
  $= L^+$
- $L\{\Lambda\} + LL^1 + LL^2 + \ldots = \{\Lambda\}L + L^1L + L^2L + \ldots$

Cerraduras de lenguajes finitos con Λ

- Sea $L = \{\Lambda, 0, 11\}$ $(\Sigma = \{0, 1\})$
  - $L^0 = \{\Lambda\}$
  - $L^1 = L = \{\Lambda, 0, 11\}$ (i.e. L^1 contiene a L^0)
  - $L^2 = L^1L^1 = \{\Lambda, 0, 11\}\{\Lambda, 0, 11\}$
    $= \{\Lambda\Lambda, \Lambda 0, \Lambda 11, 0\Lambda, 00, 011, 11\Lambda, 110, 1111\}$
    $= \{\Lambda, 0, 11, 0, 00, 011, 11, 110, 1111\}$
    $= \{\Lambda, 0, 11, 00, 011, 110, 1111\}$ (i.e. L^2 contiene a L)
  - $L^2 = L \cup \{0, 11\}\{0, 11\} = L \cup (L^1 - L^0)(L^1 - L^0)$
  - $L^3 = L^1L^2 = L^2 \cup (L^1 - L^0)\,(L^2 - L^1)$
- Cerraduras:
  - $L^* = \{\Lambda, 0, 11, 00, 011, 110, 1111, \ldots\}$
  - $L^+ = \{\Lambda, 0, 11, 00, 011, 110, 1111, \ldots\}$
  - $L^* = L^+$

Cálculo de L^3

L^2

1111	Λ1111	01111	111111	
110	Λ110	0110	11110	
011	Λ011	0011	11011	
00	Λ00	000	1100	
11	Λ11	011	1111	
0	Λ0	00	110	
Λ	ΛΛ	0Λ	11Λ	
	Λ	0	11	L

Cálculo de L^3: aplicando la id.

¡Λ es un reproductor de lenguajes!

L^2	L^2			
1111	1111	01111	111111	
110	110	0110	11110	
011	011	0011	11011	
00	00	000	1100	
11	11	011	1111	
0	0	00	110	
Λ	Λ	0	11	L^1
	Λ	0	11	L

The contribution of L^3: $(L^1 - L^0)(L^2 - L^0)$

L^2	L^2			
1111	1111	01111	111111	
110	110	0110	11110	
011	011	0011	11011	
00	00	000	1100	
11	11	011	1111	
0	0	00	110	
Λ	Λ	0	11	L^1
	Λ	0	11	L

$L^3 = L^0 \cup L^1 \cup L^2 \cup (L^1 - L^0)(L^2 - L^0)$

La contribución de L^3

L^2	L^2			
1111	1111	01111	111111	
110	110	0110	11110	
011	011	0011	11011	
00	00	000	1100	
11	11	011	1111	$L^2 - L^1$
0	0	00	110	
Λ	Λ	0	11	L^1
	Λ	0	11	L

$L^3 = L^0 \cup L^1 \cup L^2 \cup (L^1 - L^0)(L^2 - L^1)$

Si $\Lambda \in L$

$$L^n = \bigcup_{i=0}^{n-1} L^i \bigcup (L^1 - L^0)(L^{n-1} - L^{n-2})$$

¿Es verdad?

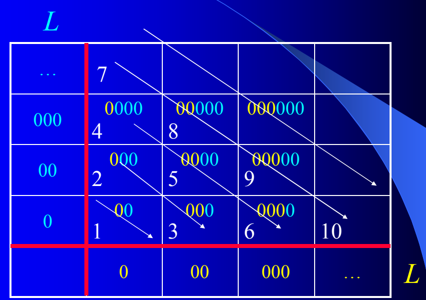
Cerraduras de lenguajes infinitos sin Λ

- Sea $L = \{0, 00, 000, \ldots\}$ $(\Sigma = \{0\})$
  - $L^0 = \{\Lambda\}$
  - $L^1 = L = \{0, 00, 000, \ldots\}$
  - $L^2 = L^1L^1 = \{0, 00, 000, \ldots\}\{0, 00, 000, \ldots\}$
    $= \{00, 000, 0000, \ldots, 000, 0000, 00000, \ldots\}$
    $= \{00, 000, 0000, 00000, \ldots\} = L^1 - \{0\}$
- Cerraduras:
  - $L^* = \{\Lambda, 0, 00, 000, \ldots\}$
  - $L^+ = \{0, 00, 000, \ldots\}$
  - $L^+ = L$

Enumeración de la concatenación de lenguajes infinitos

L					
…					
000	0000	00000	000000		
00	000	0000	00000		
0	00	000	0000		
	0	00	000	…	L

$L^2 = \{00,000,000,0000,0000,0000,\ldots\}$



Cerraduras de lenguajes infinitos con Λ

- Sea $L = \{\Lambda, 0, 00, 000, \ldots\}$ $(\Sigma = \{0\})$
  - $L^0 = \{\Lambda\}$
  - $L^1 = L = \{\Lambda, 0, 00, 000,\ldots\}$
  - $L^2 = L^1L^1 = \{\Lambda, 0, 00, 000,\ldots\}\{\Lambda, 0, 00, 000,\ldots\}$
    $= \{\Lambda\Lambda, \Lambda 0, \Lambda 00, \Lambda 000 \ldots, 0\Lambda, 00, 000, 0000, \ldots, 000, 0000, 00000,\ldots\}$
    $= \{\Lambda, 0, 00, 000, 0000, 00000\ldots\}$
  - $L^2 = L^1$
- Cerraduras:
  - $L^* = \{\Lambda, 0, 00, 000, \ldots\}$
  - $L^+ = \{\Lambda, 0, 00, 000, \ldots\}$
  - $L^* = L^+ = L$ y, además, $L^{i+1} = L^i$ para todo $i > 0$

La cerradura otra vez!

- Todas las cadenas formadas por 0 o más elementos de *L,* permitiendo repeticiones, está en L^*
  - El primer elemento de la cadena está en *L*
  - Los primeros *dos* elementos de la cadena están en L^2, incluso si son el mismo!
  - Los primeros *tres* elementos de la cadena están en L^3, ya que se componen de la concatenación de L^2 con *L,* incluso si son el mismo
  - …

Cerraduras

- La cerradura de $L = \{0, 1\}$ es el conjunto de todas las cadenas formadas por 0, 1 & Λ
- 100111 pertenece a la cerradura:
  - 100111 pertenece a L
  - 100111 pertenece a L^2
  - 100111 pertenece a L^3
  - 100111 pertenece a L^4
  - 100111 pertenece a L^5
  - 100111 pertenece a L^6

Cerraduras

- La cerradura de $L = \{10, 111\}$ es el conjunto de todas las cadenas formadas por 10 y 111, incluyendo repeticiones y Λ:
- 10111111 pertenece a la cerradura:
  - 10111111 pertenece a L
  - 10111111 pertenece a L^2
  - 10111111 pertenece a L^3
- Pero 101111 no está incluida!
  - 101111 pertenece a L^1
  - 101111 pertenece a L^2
  - 101111 no pertenece a ninguna potencia de *L*!

Cerraduras

- La cerradura de $L = \{11, 111, 11111, 1111111\ldots\}$ es el conjunto de cadenas formadas por una secuencia de números primos en notación monádica, y Λ!
- 1111111111 pertenece a la cerradura:
  - 1111111111 pertenece a L
  - 1111111111 pertenece a L^2
  - 1111111111 pertenece a L^3
- Hay algún número $n > 1$ que no esté incluido en la cerradura de *L*?

Cerradura

- Probablemente es difícil conceptualizar la cerradura de un lenguaje
- Pero es mucho más simple verificar si una cadena está en la cerradura de un lenguaje:
  - Leer la cadena de izquierda de a derecha y verificar que cada subcadena pertenece al lenguaje
  - Si la cadena se lee completamente y es posible extraer palabras del lenguaje en cada paso, la cadena está en la cerradura del lenguaje
- Es muy simple generar las cadenas en la cerradura: concatenar todos los símbolos de L, permitiendo repetiones, en un orden dado

Dr. Luis Pindea, IIMAS-UNAM, 2010

Ejemplo 1

- Dar un ejemplo de lenguajes INFINITOS L_1, L_2 sobre $\Sigma = \{a, b\}$ donde $L_1 \not\subset L_1L_2$ & $L_2 \not\subset L_1L_2$. Justificar la respuesta dando cadenas en L_1 y L_2 que no estén en L_1L_2

Dr. Luis Pindea, IIMAS-UNAM, 2010

Ejemplo 1

- Dar un ejemplo de lenguajes INFINITO L_1, L_2 sobre $\Sigma = \{a, b\}$ donde $L_1 \not\subset L_1L_2$ & $L_2 \not\subset L_1L_2$. Justificar la respuesta dando cadenas en L_1 y L_2 que no estén en L_1L_2
- Asumir:
  - $L_1 = \{a, aa, aaa, \ldots\}$
  - $L_2 = \{b, bb, bbb, \ldots\}$
  - $L_1L_2 = \{ab, abb, aab, abbb, aabb, aaabb \ldots\}$

Dr. Luis Pindea, IIMAS-UNAM, 2010

Ejemplo 1

- Dar un ejemplo de lenguajes INFINITOS L_1, L_2 sobre $\Sigma = \{a, b\}$ donde $L_1 \not\subset L_1L_2$ & $L_2 \not\subset L_1L_2$. Justificar la respuesta dando cadenas en L_1 y L_2 que no estén en L_1L_2
- Asumir:
  - $L_1 = \{a, aa, aaa, \ldots\}$
  - $L_2 = \{b, bb, bbb, \ldots\}$
  - $L_1L_2 = \{ab, abb, aab, abbb, aabb, aaabb \ldots\}$
- Consecuentemente:
  - $L_1 \not\subset L_1L_2$ & $L_2 \not\subset L_1L_2$
  - En particular, $a^k \in L_1$ pero $a^k \notin L_1L_2$
  - De manera similar, $b^k \in L_2$ pero $b^k \notin L_1L_2$

Dr. Luis Pindea, IIMAS-UNAM, 2010

Ejemplo 2

- Dar un ejemplo de lenguajes L_1, L_2 y L_3 tales que
  $(L_1 \cap L_2)L_3 \neq (L_1L_3) \cap (L_2L_3)$

Dr. Luis Pindea, IIMAS-UNAM, 2010

Ejemplo 2

- Dar un ejemplo de lenguajes L_1, L_2 y L_3 tales que
  $(L_1 \cap L_2)L_3 \neq (L_1L_3) \cap (L_2L_3)$
- Sea $L_1 = \{\Lambda\}$, $L_2 = \{a, b\}^+ = \Sigma^+$ y $L_3 = \{a, b\}^* = \Sigma^*$

Dr. Luis Pindea, IIMAS-UNAM, 2010

Ejemplo 2

- Dar un ejemplo de lenguajes L_1, L_2 y L_3 tales que
  $(L_1 \cap L_2)L_3 \neq (L_1L_3) \cap (L_2L_3)$
- Sea $L_1 = \{\Lambda\}$, $L_2 = \{a, b\}^+ = \Sigma^+$ y $L_3 = \{a, b\}^* = \Sigma^*$
- Lado izquierdo:
  $$\begin{aligned}(L_1 \cap L_2)L_3 &= (\{\Lambda\} \cap \Sigma^+)\Sigma^* \\ &= \Phi\Sigma^* \\ &= \Phi\end{aligned}$$

Ejemplo 2

- Dar un ejemplo de lenguajes L_1, L_2 y L_3 tales que
  $(L_1 \cap L_2)L_3 \neq (L_1L_3) \cap (L_2L_3)$
- Sea $L_1 = \{\Lambda\}$, $L_2 = \{a, b\}^+ = \Sigma^+$ y $L_3 = \{a, b\}^* = \Sigma^*$
- Lado izquierdo:
  $$\begin{aligned}(L_1 \cap L_2)L_3 &= (\{\Lambda\} \cap \Sigma^+)\Sigma^* \\ &= \Phi\Sigma^* \\ &= \Phi\end{aligned}$$
- Lado derecho:
  $$\begin{aligned}(L_1L_3) \cap (L_2L_3) &= (\{\Lambda\}\Sigma^*) \cap (\Sigma^+\Sigma^*) \\ &= \Sigma^* \cap \Sigma^+ \\ &= \Sigma^+\end{aligned}$$

Ejemplo 2

- Dar un ejemplo de lenguajes L_1, L_2 y L_3 tales que
  $(L_1 \cap L_2)L_3 \neq (L_1L_3) \cap (L_2L_3)$
- Sea $L_1 = \{\Lambda\}$, $L_2 = \{a, b\}^+ = \Sigma^+$ y $L_3 = \{a, b\}^* = \Sigma^*$
- Lado izquierdo:
  $$\begin{aligned}(L_1 \cap L_2)L_3 &= (\{\Lambda\} \cap \Sigma^+)\Sigma^* \\ &= \Phi\Sigma^* \\ &= \Phi\end{aligned}$$
- Lado derecho:
  $$\begin{aligned}(L_1L_3) \cap (L_2L_3) &= (\{\Lambda\}\Sigma^*) \cap (\Sigma^+\Sigma^*) \\ &= \Sigma^* \cap \Sigma^+ \\ &= \Sigma^+\end{aligned}$$
- Consecuentemente, $\Phi \neq \Sigma^+$ y $(L_1 \cap L_2)L_3 \neq (L_1L_3) \cap (L_2L_3)$

Resumen

- Dado un alfabeto Σ
- Hay un universo de cadenas Σ^*
- Un lenguaje L es un subconjunto de Σ^*
- Hay mucho, muchos, subconjuntos de Σ^*: 2^{Σ^*}
- Algunos de estos son finitos; otros infinitos pero contables
- Dado un conjunto de lenguajes iniciales, es posible generar nuevos lenguajes a través de las operaciones de conjuntos, la concatenación y las operaciones de cerradura.

Problemas y lenguajes

- Un problema: dado un lenguaje y una cadena decidir si la cadena pertenece al lenguaje :
  - Especificación declarativa: $s \in L$
  - Especificación procedural:

(L, s) ⟶ **Procedimiento de decisión: El algoritmo** ⟶ Si o no!

- Dos lados de la moneda:
  - Generar todas las cadenas del lenguaje
  - Verificar si una cadena dada pertenece al lenguaje

Universo de los lenguajes formales