



Teoría de Lenguajes

DC FCEyN UBA

Clase práctica 1

Generalidades

Cadenas, lenguajes, operaciones



Repaso Conjuntos

Sean A, B conjuntos

Algunas nociones básicas a recordar:

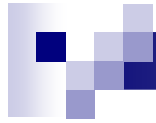
$$A \cup B = \{e \mid e \in A \vee e \in B\}$$

$$A \cap B = \{e \mid e \in A \wedge e \in B\}$$

$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

$$P(A) = \{B \mid B \subseteq A\}$$

$$A \setminus B = \{e \mid e \in A \wedge e \notin B\}$$



Lenguajes formales

Cadena y lenguaje serán
nociones centrales en la materia.



Alfabeto – “el a,b,c de nuestro tema”

Conjunto *finito* de símbolos. Los nombramos con letras griegas mayúsculas. Ejemplos:

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

$$\Gamma = \{0, 1\}$$

$$\Pi = \{\square, \triangle\}$$

Cadena, string, secuencia...

Secuencia *finita* de símbolos de un alfabeto dado.

Las nombramos con letras griegas minúsculas. Por ejemplo, las siguientes son cadenas sobre Σ :

$$\alpha = ab$$

$$\beta = babeg$$

$$\gamma = g$$

$$\left\{ \begin{array}{l} \lambda \\ a \cdot s \\ a \in \Sigma \\ s \end{array} \right.$$

Conjunto estrella... de cadenas

La cadena vacía se denota λ . (En la bibliografía se la suele encontrar como ϵ). Notar que λ *no* es un símbolo del alfabeto. Es una secuencia de cero símbolos.

Formalmente, una cadena ω de largo n es una tupla de n elementos de Σ ; es decir, $\omega \in \Sigma \times \Sigma \times \dots \times \Sigma = \Sigma^n$. Así, Σ^n es el conjunto de todas las cadenas sobre Σ de largo n . Por definición, $\Sigma^0 = \{\lambda\}$

Al conjunto de todas las cadenas posibles sobre Σ lo notamos Σ^* y se define de la siguiente manera:

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



Concatenación entre cadenas

Sean $\alpha = a_1a_2 \dots a_n, \beta = b_1b_2 \dots b_m \in \Sigma^*$. Definimos la operación concatenación como

$$\alpha.\beta = a_1a_2 \dots a_nb_1b_2 \dots b_m$$

Podemos omitir el ., entonces $\alpha.\beta$ lo podemos escribir como $\alpha\beta$.

Ejemplos:

$$\alpha = ab, \beta = babeg, \gamma = g$$

$$\alpha.\beta = abbabeg$$

$$\gamma.\gamma = \gamma\gamma = gg$$

$$\alpha\lambda = \alpha \quad (\lambda \text{ es el elemento neutro de la concatenación de cadenas})$$



Asociatividad y neutro

- La concatenación es asociativa (y no conmutativa).
- Hay un elemento neutro, que es la cadena λ .
- Esto dice que el conjunto de todas las cadenas sobre un alfabeto dado, junto con la operación de concatenación, tienen estructura de *monoide*.



Potencia

Definición (inductiva sobre n):

$$\alpha^n = \begin{cases} \lambda & , \text{ si } n = 0 \\ \alpha.\alpha^{n-1} & , \text{ si } n > 0 \end{cases}$$



Ejemplos

Ejemplos:

$$\alpha^0 = \lambda, \alpha^1 = ab, \alpha^2 = abab, \alpha^3 = ababab, \dots$$



Poniendo reversa

Definición (inductiva sobre la estructura de la cadena):

$$\begin{aligned}\lambda^r &= \lambda \\ (a.\alpha)^r &= \alpha^r.a\end{aligned}$$

Ejemplos:

$$\alpha^r = ba, \beta^r = gebab, \gamma^r = g, \lambda^r = \lambda$$



Para tener en *cuenta*

Longitud

Definición:

$$\begin{aligned} |\lambda| &= 0 \\ |a.\alpha| &= 1 + |\alpha| \end{aligned}$$

Cantidad de apariciones

$|\alpha|_b$ es la cantidad de apariciones del símbolo b en la cadena α . Se puede definir de la siguiente manera:

$$\begin{aligned} |\lambda|_b &= 0 \\ |a.\alpha|_b &= \begin{cases} 1 + |\alpha|_b & , \text{ si } a = b \\ |\alpha|_b & , \text{ si } a \neq b \end{cases} \end{aligned}$$

Ejercicio

Demostrar que la longitud de la concatenación de dos cadenas es la suma de sus longitudes respectivas, es decir: $|\alpha\beta| = |\alpha| + |\beta|$

Acá vamos a utilizar inducción estructural, que ya la vimos en Algoritmos II. La idea es pensar que todas las cadenas son de la forma:

- $\omega = \lambda$
- $\omega = a.\omega'$, con $\omega' \in \Sigma^*$, $a \in \Sigma$

Luego, probamos el caso base y el paso inductivo sobre la estructura de las cadenas.

En este caso tenemos dos cadenas. Hacemos inducción sobre una sola de ellas.

1. Si $\alpha = \lambda$:

$$|\lambda\beta| = |\beta| = 0 + |\beta| = |\lambda| + |\beta|$$

2. Si $\alpha = a\gamma$ y para γ vale la hipótesis:

$$\begin{aligned} |(a\gamma)\beta| &= |a(\gamma\beta)| \text{ (por asociatividad concat)} = 1 + |\gamma\beta| \text{ (por def long)} = \\ &= 1 + |\gamma| + |\beta| \text{ (por HI)} = |a\gamma| + |\beta| \text{ (por def long)} \end{aligned}$$



Ejercicio

Demostrar que $|\alpha^n| = n|\alpha|$

En este caso usamos inducción sobre naturales, vista en Álgebra I.

1. Si $n = 0$

$$|\alpha^0| = |\lambda| = 0 = 0|\alpha|$$

2. Si $n > 0$ y para $n - 1$ vale HI:

$$\begin{aligned} |\alpha^n| &= |\alpha\alpha^{n-1}| \text{ (por def potencia)} = |\alpha| + |\alpha^{n-1}| \text{ (por prop demostrada)} = \\ &= |\alpha| + (n-1)|\alpha| \text{ (por HI)} = n|\alpha| \end{aligned}$$



Ejercicios

- Para toda α y toda β , $(\alpha\beta)^r = \beta^r\alpha^r$
- Para toda cadena α , $(\alpha^r)^r = \alpha$



Lenguaje

- Dado un alfabeto, un *lenguaje* es un conjunto cualquiera de cadenas de ese alfabeto.
- Es decir, L es un lenguaje sobre el alfabeto Σ si $L \subseteq \Sigma^*$.
- Los denotamos con letras romanas mayúsculas.
- Idea: pensar en cadenas consideradas *sintácticamente válidas*

Ejemplos

- Sobre $\Sigma = \{a, b, c\}$:

$$L_1 = \{a, aa, baba, ab, cab\}$$

$$L_2 = \{\lambda, b, a\}$$

$$L_3 = \emptyset$$

$$L_4 = \{\lambda\} = \Lambda$$

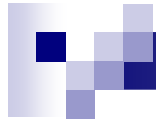
(Λ es la forma de denotar al lenguaje cuya única cadena es la vacía)

- Sobre $\Gamma = \{0, 1\}$:

$$M_1 = \{0, 01, 1010\}$$

$$M_2 = \{1, 01, 001, 0001, 00001, \dots\} = \{0^n.1 | n \geq 0\}$$

(M_2 fue definido primero por extensión y luego por comprensión)



Operaciones con lenguajes

No sólo las clásicas: \cup \cap \setminus

Veremos otras, como
concatenación y clausuras.



Unión

Sean L_1, L_2 lenguajes sobre Σ . La unión de lenguajes es otro lenguaje.
Usando los conjuntos definidos anteriormente

$$L_1 \cup L_2 = \{ \lambda, a, b, aa, baba, ab, cab \}$$

$$L_1 \cup \emptyset = L_1$$

$$L_1 \cup \Lambda = \{ \lambda, a, aa, baba, ab, cab \}$$

$$L_1 \cup \Sigma^* = \Sigma^*$$

\emptyset es el elemento neutro de la unión de lenguajes.

Para la intersección, ¿cuál es el elemento neutro?



Complemento

El complemento de un lenguaje es el conjunto de todas las cadenas del mismo alfabeto que no están en él. Es decir:

$$L^c = \bar{L} = \Sigma^* \setminus L = \{\alpha \mid \alpha \in \Sigma^* \wedge \alpha \notin L\}$$

Ejemplo: Para $\Sigma = \{a\}$, $\{a^n \mid n \geq 3\}^c = \{\lambda, a, a^2\}$

Para $\Sigma = \{a,b\}$, $\{a^n \mid n \geq 3\}^c$ es más complicado...



Concatenación entre lenguajes

La concatenación de dos lenguajes es un lenguaje formado por todas las posibles combinaciones de concatenar una cadena del primero con una del segundo:

$$L.M = \{\alpha.\beta \mid \alpha \in L \wedge \beta \in M\}$$

- La concatenación de dos lenguajes es un lenguaje formado por todas las posibles combinaciones de concatenar una cadena del primero con una del segundo.
- A semeja a la idea del producto cartesiano.
- También es asociativa y no conmutativa.



Ejemplos

$$L_1.L_2 = \{a, ab, aa, aab, aaa, baba, babab, babaa, abb, aba, cab, cabb, caba\}$$

$$\{a^n / n \geq 0\} \{a\} = \{a^n / n \geq 1\} = \{a\} \{a\}^* = \{a\} \{a\}^*$$

$$\{a^n / n \geq r\} \{a^n / n \geq s\} = \{a^n / n \geq r+s\} = \{a^n / n \geq r\} \{a^s\}$$



Preguntas

- ¿Hay un elemento absorbente para la concatenación? i.e., $X / \forall L \quad L X = X L = \emptyset$
- ¿Hay un elemento neutro para la concatenación? i.e., $X / \forall L \quad L X = X L = L$
- Pensar.

- Respuestas:

$$L . \Lambda = L$$

$$L . \emptyset = \emptyset$$



Poniendo potencia

Definición (similar a la de potencia de cadenas):

$$L^n = \begin{cases} \Lambda & , \text{ si } n = 0 \\ L.L^{n-1} & , \text{ si } n > 0 \end{cases}$$

Ejemplo:

$$L_2^0 = \{\lambda\}$$

$$L_2^1 = \{\lambda, b, a\}$$

$$L_2^2 = \{\lambda, b, a, bb, ba, ab, aa\}$$

$$L_2^3 = \{\lambda, b, a, bb, ba, ab, aa, bbb, bba, bab, baa, abb, aba, aab, aaa\}$$

...



Yendo hacia arriba - pregunta

¿Para todo L , para todo n , $L^n \subseteq L^{n+1}$?

Por ejemplo, ¿valdrá $L \subseteq LL$ para todo L ?



Observación

Notar que como λ está en L_2 , para todo n vale: $L_2^n \subseteq L_2^{n+1}$. En cambio, lo mismo no se cumple para L_1 :

$$L_1^0 = \{\lambda\}$$

$$L_1^1 = \{a, aa, baba, ab, cab\}$$

$$L_1^2 = \{aa, aaa, ababa, aab, acab, aaaa, aababa, aaab, aacab, babaa, babaaa, babababa, babaab, babacab, aba, abaa, abbaba, abab, abcab, caba, cabaa, cabbaba, cabab, cabcab\}$$



Clausura

Ya vimos la clausura de Kleene de un alfabeto (Σ^* es la clausura de Kleene de Σ). Para un lenguaje se define análogamente (coloquialmente la llamamos “L estrella”):

$$L^* = \bigcup_{i \geq 0} L^i = L^0 \cup L^1 \cup L^2 \cup L^3 \dots \quad (1)$$

Llamamos clausura positiva (coloquialmente “L más”) a:

$$L^+ = \bigcup_{i > 0} L^i = L^1 \cup L^2 \cup L^3 \dots \quad (2)$$

Ejemplo:

$$\{a, b\}^+ = \{a, b, aa, ab, ba, bb, aaa, aba, baa, bba, aab, abb, bab, bbb, \dots\}$$

$$\begin{aligned} L^* &= L^+ \cup \Lambda \\ L^+ &= L.L^* \end{aligned}$$



Ejercicios

$$L \subseteq L' \Rightarrow L^* \subseteq L'^* \wedge L^+ \subseteq L'^+$$

$$(L^*)^* = L^*$$

$$(L^+)^+ = L^+$$

$$(L^*)^+ = L^*$$

$$(L^+)^* = L^*$$



Ejercicios (más bien preguntas)

- ¿Vale que $(A \cup B)^* = A^* \cup B^*$?
- ¿Vale que $(A \cap B)^* = A^* \cap B^*$?
- ¿Vale que $(A^2)^* = A^*$?
- ¿Vale que $(A \cup A^2)^* = A^*$?
- Hallar L tal que $L^2 \subset L$ pero $L^3 \not\subset L$.

Sub cadenas, iniciales y finales

- Las *subcadenas* de L , $\text{Sub}(L)$ o $S(L)$, son todas las cadenas que se pueden formar tomando una cadena cualquiera de L y quitándole 0 o más símbolos del comienzo y 0 o más símbolos del final.
- Las *cadenas iniciales* de L , $\text{Init}(L)$ o $I(L)$, son todas las cadenas que se pueden formar tomando una cadena cualquiera de L y quitándole 0 o más símbolos del final.
- Las *cadenas finales* de L , $\text{Fin}(L)$ o $F(L)$, son todas las cadenas que se pueden formar tomando una cadena cualquiera de L y quitándole 0 o más símbolos del comienzo.

Formalmente:

$\text{Sub}(L)$	$=$	$S(L)$	$=$	$\{\alpha \mid \beta\alpha\gamma \in L\}$
$\text{Init}(L)$	$=$	$I(L)$	$=$	$\{\alpha \mid \alpha\gamma \in L\}$
$\text{Fin}(L)$	$=$	$F(L)$	$=$	$\{\alpha \mid \beta\alpha \in L\}$



Observaciones conjuntas

$$L \subseteq I(L) \subseteq S(L)$$

$$L \subseteq F(L) \subseteq S(L)$$

$$\lambda \in I(L), F(L), S(L) \text{ (¡siempre que } L \neq \emptyset\text{!)}$$

$$\lambda \in L \subseteq I(L) \cap F(L) \subseteq I(L) \cup F(L) \subseteq S(L)$$



Ejercicio

Demostrar que $S(L) = F(I(L))$ Para esto hay que ver que:

$$\forall \alpha (\alpha \in S(L) \iff \alpha \in F(I(L)))$$

Lo hacemos en forma directa a partir de las definiciones:

$$\begin{aligned} \alpha \in S(L) &\iff \exists \beta \exists \gamma (\beta \alpha \gamma \in L) && \text{(por definición de Sub)} \\ &\iff \exists \beta (\beta \alpha \in I(L)) && \text{(por definición de Init)} \\ &\iff \alpha \in F(I(L)) && \text{(por definición de Fin)} \end{aligned}$$



Ejercicios

$$I(I(L)) = I(L)$$

$$F(F(L)) = F(L)$$

$$S(S(L)) = S(L)$$

$$S(I(L)) = I(S(L)) = S(F(L)) = F(S(L)) = S(L)$$

$$I \circ I = I$$

$$F \circ F = F$$

$$S \circ I = I \circ S = S \circ F = F \circ S = \\ = F \circ I = I \circ F = S \circ S = S$$



Ejercicios

$$(L^n)^m = (L^m)^n = L^{m+n}$$

$$(L^n)^* = (L^*)^n$$

$$(L^n)^+ = (L^+)^n$$

$$(L^n)^r = (L^r)^n$$

$$(L^*)^r = (L^r)^*$$

$$(L^+)^r = (L^r)^+$$



Ejercicios: ifs...fis

$$I(L^r) = F(L)^r$$

$$F(L^r) = I(L)^r$$

$$S(L^r) = S(L)^r$$

¡Símil leyes de de Morgan!

Seguimos la próxima...