Sintaxis y Semántica de los lenguajes

Introducción

(2021)

Facultad Regional Delta,
Universidad Tecnológica Nacional

Prof. Juan Miguel Santos

Introducción

- **Relación**.
  - Propiedades de las relaciones.
  - Reflexividad. Simetría. Transitividad.
  - Relaciones de equivalencia.
  - Relación de composición.
  - Relación de identidad.
  - Relación de potencia.
- **Clausura transitiva**.
- **Clausura transitiva reflexiva**.
- **Representación de relaciones**.
  - Matriz de unión.
  - Matriz de intersección.
  - Matriz de composición.
  - Algoritmo de Wharshall.
- **Inducción estructural**.

Preliminares.

- Definición de **alfabeto**.
- Definición de **cadena**.
- **Concatenación** de cadenas.
- **Asociatividad**.
- Operaciones sobre cadenas.
  - Agregar a la cola.
  - Cabeza.
  - Cuerpo.
  - **Reversa**.
- Abusos de notación.

Lenguajes y Gramáticas.

- **Lenguaje**.
- **Gramática**.
- Lenguaje generado por una gramática.
- **Clasificación de Chomsky**.
  - Gramáticas del tipo 0.
  - Gramáticas del tipo 1.
  - Gramáticas del tipo 2.
  - Gramáticas del tipo 3.
- Inclusión de los lenguajes dados por las gramáticas.
- Lenguaje estrictamente del tipo i.
- **Operaciones sobre lenguajes**.
  - Producto de lenguajes o concatenación.
  - Potencia de Lenguajes.
  - Clausura de un Lenguaje.
  - Clausura de Kleene de un lenguaje.
  - Relación entre A* y A+.

Relación:

Dados dos conjuntos A y B, se define una relación como el conjunto de pares ordenados (a,b) donde $a \in A$ y $b \in B$.

La **notación** es **aRb** ó **R:A→B**.

Toda relación R:A →B está incluída en el producto cartesiano de

A x B, esto es

$$R: A \rightarrow B \Rightarrow R \subseteq A \times B.$$

En particular, A puede ser igual a B, en cuyo caso la relación se define como R: A→A ó R: B→B.

Algunas propiedades de las relaciones

Reflexividad: Sea R una relación en A (R: $A \rightarrow A$), R es reflexiva si $\forall a \in A$ se cumple que aRa.

Ejemplo: "$\leq$": $\aleph \rightarrow \aleph$

Simetría: Sea R una relación en A (R: $A \rightarrow A$), R es simétrica si $\forall a,b \in A$, si aRb $\Rightarrow$ bRa.

Ejemplo: "$\neq$": $\aleph \rightarrow \aleph$ es simétrica y

"$<$": $\aleph \rightarrow \aleph$ no es simétrica.

Transitividad: Sea R una relación en A (R: $A \rightarrow A$), R es transitiva si $\forall a,b,c \in A$, si aRb $\wedge$ bRc $\Rightarrow$ aRc.

Ejemplo: "ser paralela": $X \rightarrow X$, donde X es el conjunto de rectas en el plano es transitiva, y

"X es hijo de Y" no es transitiva.

Relación de equivalencia: una relación es de equivalencia si es transitiva, simétrica y reflexiva.

Una característica de estas relaciones es que **inducen clases** de equivalencias (una partición del conjunto).

Ejemplo: Sea R sobre los números enteros y sean a,b dos números enteros tal que “**a** tiene el mismo resto que **b** al ser dividido por 3”, entonces **a** y **b** pertenecen a la misma clase.

Esta relación que es de equivalencia induce 3 clases: la del 0, la del 1 y la del 2.

(…-5,-2,1,4,7,10,13,16,…) pertenecen a la clase del 1

Relación de composición:

Sean A, B y C tres conjuntos y R, G dos relaciones tal que

$R: A \rightarrow B$ y $G: B \rightarrow C$,

luego la relación de composición RoG se define como

$RoG = \{ (a,c), a \in A, c \in C \,/\, \exists b \in B \wedge aRb \wedge bGc\}$.

Relación de identidad:

Sea R una relación sobre A es de identidad (id_A) si se cumple que

$\forall a,b \in A,\ a\ id_A\ b \Rightarrow a=b$.

Relación potencia:

Sea R una relación sobre A se cumple que

$$R^n = \begin{cases} id_A & \text{si} \quad n = 0 \\ RoR^{n-1} & \text{si} \quad n > 0 \end{cases}$$

Notas: - la relación potencia es también un conjunto de pares,

- este signo es el de composición.

Clausura transitiva:

Sea R definida sobre A, una clausura transitiva está definida por

$$R^{+} = \bigcup_{i \in \aleph, i \geq 1} R^{i} = \bigcup_{i=1}^{\infty} R^{i}$$

Clausura transitiva reflexiva:

Sea R definida sobre A, una clausura transitiva está definida por

$$R^{*} = R^{+} \cup id = \bigcup_{i=0}^{\infty} R^{i}$$

Representación de relaciones:

Sea A un conjunto finito definido por $A = \{a_1, a_2, .., a_n\}$ y

M_R la matriz asociada a una relación R sobre A,

M_R está definida por

$$M_R[i,j] = \begin{cases} 1 & \text{si} \quad a_i R a_j \\ 0 & \text{sino} \end{cases}$$

Ejemplo:

A = {1,3,7}

aRb es "**a** *menor a* **b**"

$$M_R = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Matriz de unión, intersección y composición:

Sean dos relaciones R y G sobre A,

$$R: A \to A \text{ y } G: A \to A,$$

con las matrices que las representan

$$M_R=(r_{ij}),\ M_G=(g_{ij}),$$

la matriz unión es

$$M_{RUG}=(m_{ij}),\ \text{donde } m_{ij} = r_{ij} \text{ OR } g_{ij},$$

la matriz intersección es

$$M_{RnG}=(m_{ij}),\ \text{donde } m_{ij} = r_{ij} \text{ AND } g_{ij},$$

la matriz composición es

$$M_{RoG}=(c_{ij}),\ \text{donde} \quad c_{ij} = \bigvee_{k=1}^{n} \left(r_{ik} \quad \text{AND} \quad g_{kj} \right).$$

Nota: para los operadores lógicos considere que los operandos 0 y 1 corresponden a falso y verdadero respectivamente.

Algoritmo de Warshall:

Sea la relación R sobre A, MR la matriz de la relación y n el cardinal de A, la clausura de R estará dada por

```
Comienzo.
T←R
Para k : 1 .. n
    Para i : 1 .. n
        Para j : 1 .. n
            T(i, j) = T(i, j)  ∨  (T(i, k) ∧ T(k, j))
Fin.
```

Inducción estructural

La idea es realizar inducción pero no sobre los naturales sino sobre una estructura (la cual tiene una definición inductiva).

Ejemplo: Consideremos listas de elementos definidas de la siguiente manera

i) [ ] es la lista vacía,

ii) Si *a* pertenece al conjunto de posibles elementos de la lista y *l* es una lista, *a*::*l* es una lista.

Sea ahora el conjunto de listas definidas sobre los naturales (esto es, sus elementos son números naturales) con las siguientes operaciones:

Max(l) devuelve el elemento mayor de l

Sum(l) devuelve la suma de los números en l (sum[ ] = 0)

Long(l) devuelve la cantidad de elementos de l

Queremos demostrar que se cumple $\forall$ *l* la siguiente propiedad P(*l*)

$$\mathrm{Sum}(l) \leq \mathrm{Max}(l) * \mathrm{Long}(l).$$

Hacemos inducción en la longitud de la lista.

Caso base ($l = [\ ]$)

$P([\ ])$ es verdadero pues $Sum\ [\ ] = 0$

Paso inductivo ($P(n::l)$ asumiendo $P(l)$ es verdadero)

$Sum(n::l) = n + Sum(l)$

Por Hipótesis Inductiva $P(l)$ es verdadero y entonces

$$Sum(l) \leq Max(l) * Long(l).$$

Así, $Sum(n::l) \leq n + Max(l) * Long(l)$.

Además sabemos que $n \leq Max(n::l)$ y $Max(l) \leq Max(n::l)$.

Podemos escribir entonces

$$Sum(n::l) \leq Max(n::l) + Max(n::l)*Long(l) =$$

$$Max(n::l)(1+Long(l))=Max(n::l)*Long(n::l).$$

Alfabeto:

Es un conjunto finito de símbolos o caracteres.

Cadena:

Es una colección o conjunto ordenado de elementos de un alfabeto.

Ejemplo: sea el alfabeto $\Sigma = \{ I, V, X, L, C, M \}$,

los siguientes conjuntos ordenados son cadenas

I, II, IV, XCIII.

pero

IIIIII, XXXXXVVVIIII, XIIIC

también son cadenas…

Σ* (definición provisoria) :

es el conjunto de todas las cadenas que se pueden construir a partir de Σ.

Nota: $\lambda \in \Sigma^*$

Ejemplo: Sea $\Sigma = \{a,b\}$, aa, ab, a, b, aaa, ababa, etc pertenecen a Σ^*.

Concatenación: (entre un símbolo de Σ y una cadena)

$\circ$: Σ x {cadenas sobre Σ} $\rightarrow$ {cadenas sobre Σ} ó

o: $\Sigma \times \Sigma^* \rightarrow \Sigma^*$.

Ejemplo: Sea $\Sigma = \{a,b,c\}$ y α = ab es una cadena, luego $a \circ ab = aab$.

Σ* (clausura de Σ):

i) $\lambda \in \Sigma^*$

ii) Si $\alpha \in \Sigma^* \Rightarrow \forall a \in \Sigma, (a \circ \alpha \in \Sigma^*)$

donde $\circ$ (punto) es la concatenación u operador de concatenación definido por "$\circ$" : $\Sigma \times \Sigma^* \rightarrow \Sigma^*$

Concatenación de cadenas (+):

$$+ : \Sigma^* x\ \Sigma^* \rightarrow \Sigma^*$$

Propiedades de la concatenación:

Sea a $\in \Sigma$ y α, $\beta \in \Sigma^*$

i) $\lambda + \beta = \beta$

ii) a. $\alpha + \beta$ = a.($\alpha + \beta$)

Asociatividad de la concatenación de cadenas:

Sean $\alpha, \beta, \gamma \in \Sigma^*$ entonces $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$

Demostración por inducción en la longitud de α

<u>Caso base</u> ($\alpha = \lambda$)

$\alpha + (\beta + \gamma) = \lambda + (\beta + \gamma) = (\beta + \gamma)$ por ser λ el elemento neutro (prop. i).

Luego $(\beta + \gamma) = \beta + \gamma = (\lambda + \beta) + \gamma = (\alpha + \beta) + \gamma$.

<u>Paso inductivo</u> (la HI es verdadera para cadenas de longitud n)

Sea $a \in \Sigma$ y $\alpha = a\, \alpha'$, $|\alpha'| = n$,

$(a\, \alpha') + (\beta + \gamma) = a\, (\alpha' + (\beta + \gamma))$ por prop. ii de concatenación de cadenas.

Por HI y nuevamente por def. ii de concatenación de cadenas

$a\, (\alpha' + (\beta + \gamma)) = a\, ((\alpha' + \beta) + \gamma) =$

$= a\, (\alpha' + \beta) + \gamma = (a\, \alpha' + \beta) + \gamma = (\alpha + \beta) + \gamma$.

Proposición: $\alpha + \lambda = \alpha$

Demostración por inducción en la longitud de α

Caso base ($\alpha = \lambda$)

$\alpha + \lambda = \lambda + \lambda = \lambda = \alpha$ por ser λ elemento neutro.

Paso inductivo (la HI es verdadera para cadenas de longitud n)

Sea $a \in \Sigma$ y $\alpha = a\,\alpha'$, $|\alpha'| = n$,

$\alpha + \lambda = a\alpha' + \lambda = a(\alpha' + \lambda)$

Aplicando la HI, tenemos que $(\alpha' + \lambda) = \alpha'$ y así

$\alpha + \lambda = a(\alpha' + \lambda) = a\alpha' = \alpha$.

Operaciones sobre Cadenas

Agregar a la cola: Σ x $\Sigma^* \rightarrow \Sigma^*$

α = Agregar_a_la_cola(a, α') $\Leftrightarrow \alpha = \alpha$'a

Cabeza: $\Sigma^+ \rightarrow \Sigma$

a = Cabeza(α) $\Leftrightarrow \alpha = a\ \alpha'$

Cuerpo: $\Sigma^+ \rightarrow \Sigma^*$

α' = Cuerpo(α) $\Leftrightarrow \alpha = a\alpha'$

Reversa: $\Sigma^* \rightarrow \Sigma^*$

definida por

i) $\lambda^{R} = \lambda$

ii) $(a\ \alpha)^R = \alpha^R\ a$

Nota: Abusos de notación

i) $a \circ \alpha = a\alpha$

ii) $\alpha \circ a = \alpha a$

iii) $\alpha + \beta = \alpha\ \beta$

Lenguaje (sobre Σ):

Es un conjunto de cadenas L talque $L \subseteq \Sigma^*$

Gramática

Según la definición de Chomsky (1959) una gramática G es una upla

$$G = < V_N, V_T, P, S >$$

donde **V_N** es un conjunto de símbolos llamados **no terminales**

V_T es un conjunto de símbolos llamados **terminales**

P es un conjunto de **producciones** (o reglas gramaticales)

y **S** es el **símbolo distinguido**, $S \in V_N$.

Las reglas de producción son pares ordenados en

$$(V_T \cup V_N)^+ \times (V_T \cup V_N)^*$$

Notación

$\alpha \rightarrow \beta$

(α,β)

α se reescribe como β

β se deriva de α

α produce β

Toda cadena de un lenguaje se forma utilizando las reglas de producción a partir del símbolo distinguido.

Ejemplo:

VN

(conjunto de símbolos no terminales)

P

(conjunto de producciones)

Sea G = < {S,A}, {a,b}, {S→aA, A→aA, A→b}, S >.

VT

(conjunto de símbolos terminales)

S

(símbolo distinguido)

Para formar la cadena aaab partimos del símbolo distinguido S y aplicamos primero la producción S→aA reemplazando el símbolo S por la cadena aA.

A partir de la cadena obtenida (aA) uso alguna otra producción para reemplazar el símbolo no terminal A. En este caso uso A→aA y así obtenemos la cadena aaA. Nuevamente A→aA y finalmente A→b.

S →aA →aaA →aaaA →aaab

Notación: S →+ aaab quiere decir que se han usado 1 o más producciones para obtener de S la cadena aaab.

Lenguaje generado por un gramática

$$L(G) = \{\alpha \in V_T^* \ / \ S \rightarrow^+ \alpha\}$$

Clasificación de Chomsky

Chomsky agrupó a las gramáticas en 4 clases.

<u>Gramáticas del tipo 0</u> (sin restricciones)

$$\alpha \rightarrow \beta, \quad \alpha \in (V_N \cup V_T)+,\ \beta \in (V_N \cup V_T)^*$$

<u>Gramáticas del tipo 1</u> (sensibles al contexto)

$$\alpha A \beta \rightarrow \alpha\, \gamma \beta, \quad A \in V_N,\ \alpha,\ \beta \in (V_N \cup V_T)^*,\ \gamma \in (V_N \cup V_T)+$$

La parte derecha de una producción debe ser, en longitud, mayor o igual que a la longitud de la parte izquierda.

<u>Gramáticas del tipo 2</u> (independientes del contexto)

$$A \rightarrow \alpha, \quad A \in V_N,\ \alpha \in (V_N \cup V_T)^*$$

<u>Gramáticas del tipo 3</u> (regulares)

$A \rightarrow aB,$

$A \rightarrow a, \qquad a \in \Sigma,\ A,B \in V_N$

ó

$A \rightarrow Ba,$

$A \rightarrow a, \qquad a \in \Sigma,\ A,B \in V_N$

Ejemplo: de una gramática tipo 1 :

G = < {S,B,C}, {a,b,c}, P, S >

donde P =	{S→aSBC,	bB →bb,
	S → aBC,	CB → BC,
	C → c,	aB → ab}

S →aBC →abC →abc

S →aSBC →aaBCBC →aabCBC →aabBCC →aabbCC →aabbcC →
→aabbcc

S →aSBC →aaSBCBC →aaaBCBCBC →aaabCBCBC →aaabBCCBC →
→aaabbCCBC →aaabbCBCC →aaabbBCCC →aaabbbCCC →
→aaabbbcCC →aaabbbccC →aaabbbccc

Ejemplo: de una gramática tipo 2 o libre de contexto:

G = < {S}, {a,b}, P, S >

donde P = { $S \to aSb$,

$S \to ab$}

$S \to ab$

$S \to aSb \to aabb$

$S \to aSb \to aaSbb \to aaabbb$

Ejemplo: de una gramática tipo 3 o regular:

G = < {S}, {a}, P, S >

donde P = {$S \to aS$,

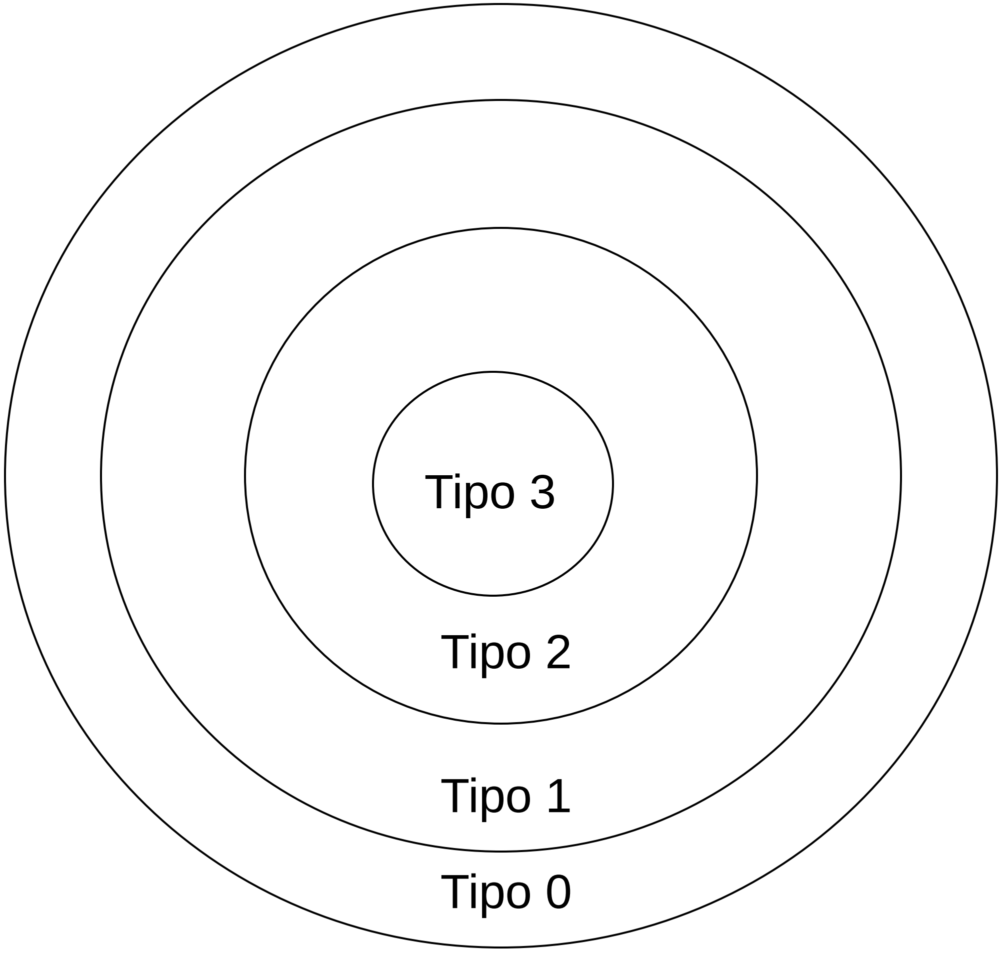
$S \to a$}

$S \to a$

$S \to aS \to aa$

$S \to aS \to aaS \to aaa$

Inclusión de los lenguajes generados por las distintas clases de gramáticas



Lenguaje del tipo i:

L es de tipo i si existe una gramática G de tipo i talque L = L(G).

Lenguaje estrictamente del tipo i:

L es estrictamente de tipo i si existe una gramática G de tipo i tal que L = L(G) y (i es 3 ó no existe G' del tipo i + 1 talque L = L(G')).

Operaciones sobre lenguajes

Sea Σ un alfabeto y A,B lenguajes sobre Σ

Producto de lenguajes o concatenación:

$$AoB = AB = \{\alpha\beta \ / \ \alpha \in A \wedge \beta \in B\}.$$

Nota: id = $\{\lambda\}$ es el elemento neutro.

Potencia de lenguajes:

$$A^n = \begin{cases} id & \text{si} \quad n = 0 \\ AoA^{n-1} & \text{si} \quad n > 0 \end{cases}$$

Clausura positiva de un lenguaje:

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

Clausura de Kleene de un lenguaje:

$$A^* = A^+ + \{\lambda\}$$

Nota: en general no es cierto que $A^+ = A^* - \{\lambda\}$.

(Pensar que λ puede pertenecer a A.)