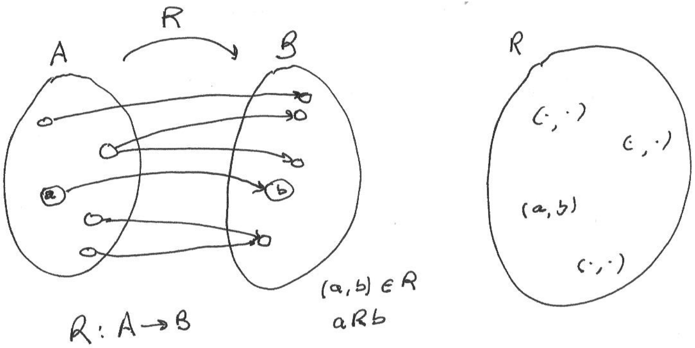
Conceptos preliminares, Lenguajes y Gramáticas

18 de agosto de 2021

Relaciones: algunas definiciones

Definición

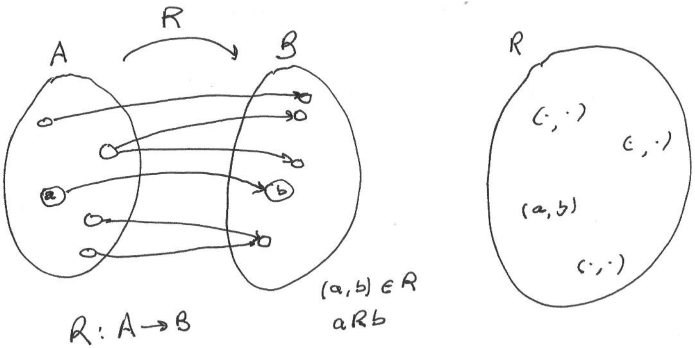
Dados los conjuntos A y B, se llama **relación** de A **en** B a todo subconjunto de $A \times B$.



Relaciones: algunas definiciones

Notación

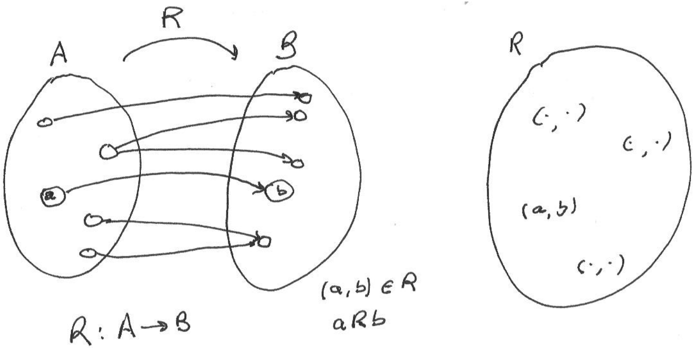
Si R es una relación de A en B (o sea, $R \subset A \times B$) esto se denota como $R : A \to B$.



Relaciones: algunas definiciones

Notación

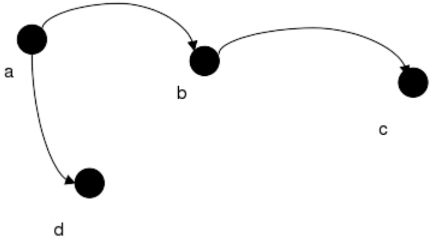
$a\,R\,b$ denota el hecho de que el par (a, b) pertenece a la relación R, esto es: $(a, b) \in R$.



Relaciones: algunas definiciones

Definición

En lo anterior, si $B = A$ se dice que R es una relación sobre A.



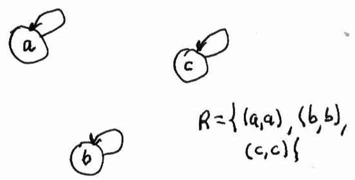
relación $R = \{(a,b), (a,d), (b,c)\}$ (Aho & Ullman. Vol. I)

Relaciones: propiedades de una relación R sobre A

Reflexividad

Una relación $R : A \to A$ es reflexiva cuando todo elemento de A está relacionado consigo mismo, o sea, cuando

$$\forall a \in A,\ (a, a) \in R.$$



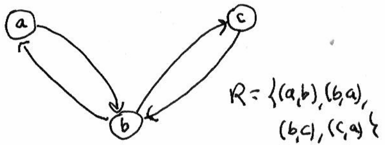
relación reflexiva

Relaciones: propiedades de una relación R sobre A

Simetría

Una relación $R : A \to A$ es simétrica cuando el hecho de que el par (a, b) pertenece a la relación R implica que el par (b, a) también pertenece a dicha relación, o sea, cuando

$$\forall a, b \in A,\ (a, b) \in R \Rightarrow (b, a) \in R.$$



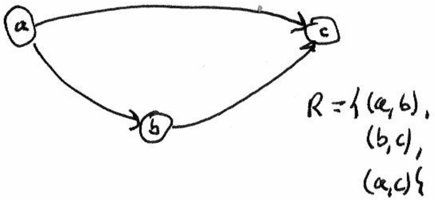
relación simétrica

Relaciones: propiedades de una relación R sobre A

Transitividad

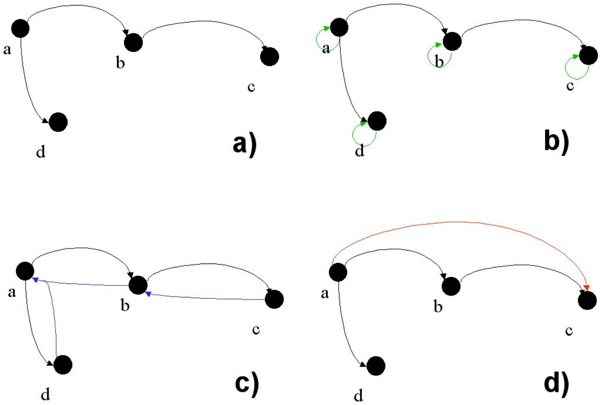
Una relación $R : A \to A$ es transitiva cuando el hecho de que los pares (a, b) y (b, c) pertenecen a la relación R implica que el par (a, c) también pertenece a dicha relación, o sea, cuando

$$\forall a, b, c \in A,\ (a, b) \in R \wedge (b, c) \in R \Rightarrow (a, c) \in R.$$



relación transitiva

Relaciones: propiedades de una relación R sobre A



a) relación $R = \{(a, b), (a, d), (b, c)\}$, b) clausura reflexiva, c) clausura simétrica, d) clausura transitiva.

Relaciones: definiciones

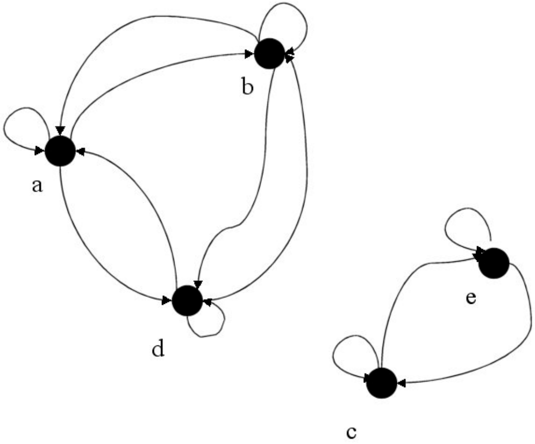
Relación de equivalencia

Una relación es de equivalencia, cuando es reflexiva, simétrica y transitiva.

Propiedad:

Una relación de equivalencia sobre un conjunto A particiona al mismo en subconjuntos disjuntos a los cuales se los llama **clases de equivalencia**.

Ejemplo de relación de equivalencia

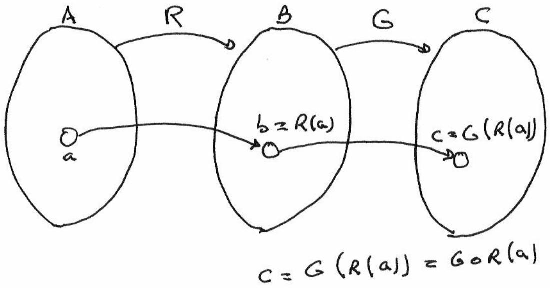


Relación de equivalencia en $\{a, b, c, d, e\}$. (Aho & Ullman. Vol. I)

Composición de relaciones:

Sean A, B y C tres conjuntos, y sean R y G dos relaciones tales que $R : A \to B$ y $G : B \to C$, se define la relación de composición $G \circ R$ como

$$G \circ R = \{(a, c), a \in A, c \in C : \exists b \in B \text{ tal que } aRb \wedge bGc\}.$$



Composición de relaciones

Relación de identidad:

Una relación R definida sobre A es de identidad (id_A) si se cumple que

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

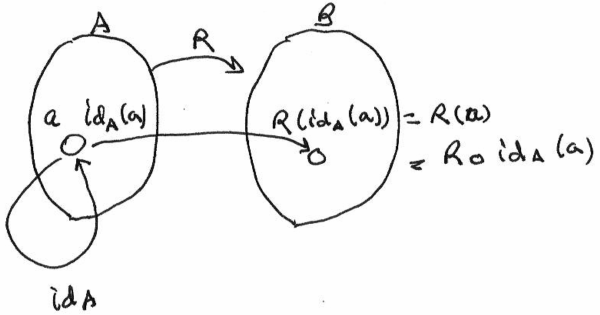
Propiedad:

la relación de identidad es el elemento neutro de la composición.

Relación de identidad

Dada una relación $R : A \to B$ es cierto que

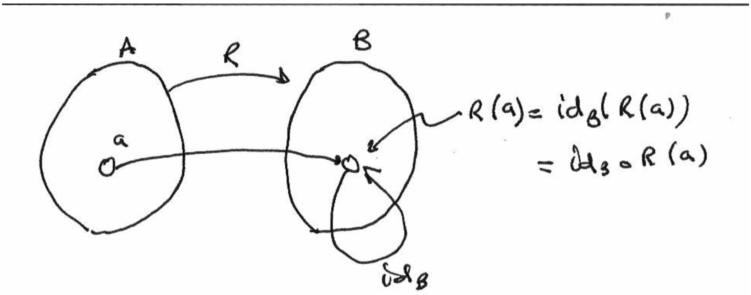
$$R \circ id_A = R$$



Relación de identidad

Dada una relación $R : A \to B$ es cierto que

$$id_B \circ R = R$$



Relación potencia:

Dada una relación R sobre A, se define R^n como:

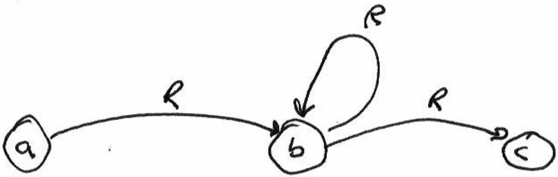
$$R^n = \begin{cases} id_A & \text{si } n = 0 \\ R \circ R^{n-1} & \text{si } n > 0 \end{cases}$$

con $R = R^1$.

Entonces

$$\begin{array}{lclcl} R^0 & = & id_A & & \\ R^1 & = & R \circ R^0 & = & R \circ id_A = R \\ R^n & = & R \circ R^{n-1} & = & \underbrace{R \circ R \ldots \circ R}_{n} \end{array}$$

Ejamplo de relación potencia



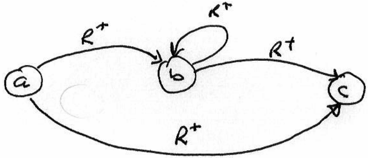
$(a,b) \in R^n$ con $n \geq 1$

$(a,c) \in R^n$ con $n \geq 2$

$(b,b) \in R^n$ con $n \geq 0$

$(b,c) \in R^n$ con $n \geq 1$

Clausura transitiva



Dada una relación R sobre A, se define clausura transitiva R^+ como:

$$
\begin{aligned}
R^+ &= \bigcup_{i=1}^{\infty} R^i, \text{ o sea} \\
R^+ &= R \cup R^2 \cup R^3 \ldots
\end{aligned}
$$

Clausura Positiva

Propiedades:

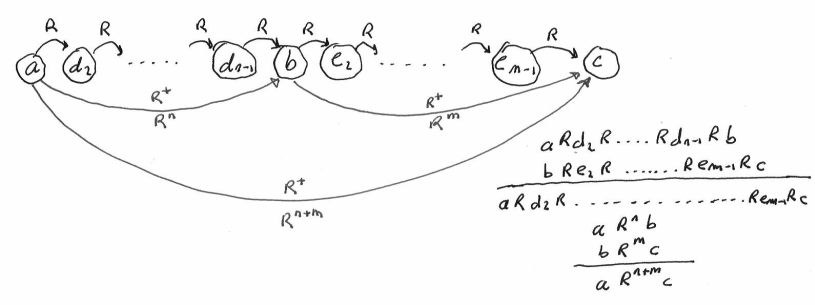
La clausura transitiva de una relación R posee las siguientes propiedades:

1. $R \subseteq R^+$
2. R^+ es transitiva
3. para toda relación G sobre A

$$R \subseteq G \quad \wedge \quad G \; transitiva \quad \Rightarrow \quad R^+ \subseteq G.$$

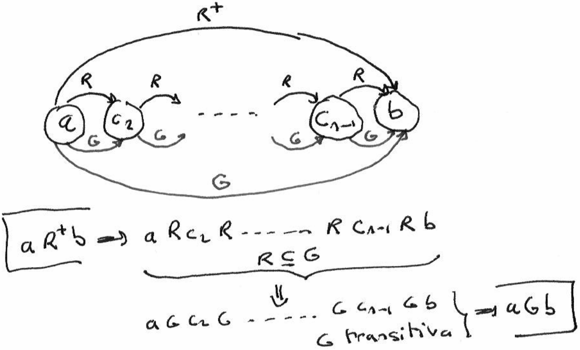
Demostración: R^+ es transitiva

Queremos probar que si aR^+b y bR^+c enonces aR^+c.
Si aR^+b, entonces existe una secuencia de elementos $d_1, \ldots, d_n$ tal que $d_1Rd_2, \ldots, d_{n-1}Rd_n$, donde $d_1 = a$ y $d_n = b$.
Análogamente, como bR^+c entonces existe una secuencia de elementos $e_1, \ldots, e_m$ tal que $e_1Re_2, \ldots, e_{m-1}Re_m$, donde $e_1 = b$ y $e_m = c$.
Por lo tanto, $aR^{n+m}c$, lo que a su vez implica que aR^+c.



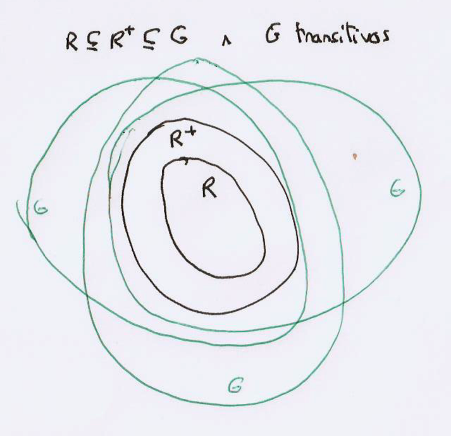
Demostración: $R \subseteq G \ \wedge \ G\ transitiva \Rightarrow R^+ \subseteq G$

si aR^+b, entonces existe una secuencia de elementos $c_1, \ldots, c_n$ tal que $c_1Rc_2, \ldots, c_{n-1}Rc_n$, donde $c_1 = a$ y $c_n = b$.
Como $R \subseteq G$ tenemos que $c_1Gc_2, \ldots, c_{n-1}Gc_n$, y como G es transitiva entonces, la aplicación repetida de la transitividad nos lleva a que c_1Gc_n, o sea aGb.



Demostración: $R \subseteq G \;\wedge\; G\ transitiva \Rightarrow R^+ \subseteq G$

De lo anterior surge que R^+ es la menor relación transitiva que incluye a la relación R.

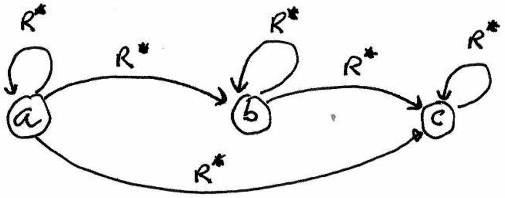


Clausura transitiva reflexiva: R^*

Definición

$$R^* = R^0 \cup R^+ = R^0 \cup R^1 \cup R^2 \ldots = \bigcup_{i=0}^{\infty} R^i.$$

donde R^0 es la relación identidad.



Algunas preguntas

Pregunta:

Dada una relación $R : A \to A$, si el conjunto A es finito, la relación R puede ser infinita?

Respuesta:

No, porque $R \subseteq A \times A$ y porque como A es finito, $A \times A$ también lo es.

Pregunta:

Puede darse que $R^* = R^+$?

Respuesta:

Sí, si la relación R es reflexiva.

Alfabeto

Es un conjunto finito de elementos o caracteres.

Cadena

Es un conjunto ordenado de elementos de un alfabeto.

Ejemplo

Dado el alfabeto: $\Sigma = \{a, b, c\}$, tenemos como cadenas posibles: $aaabbbccc$, $aaabbc$, $cbbbbbb$, etc.

Notación

Los símbolos son notados respetando el orden. Ejemplo: la cadena (a, b, c) (que es un conjunto ordenado) es notada abc.

Concatenación $\circ$

Es una operación entre un símbolo del alfabeto Σ y una cadena sobre dicho alfabeto

$$\circ : \Sigma \times \{\text{cadenas sobre } \Sigma\} \to \{\text{cadenas sobre } \Sigma\}\,.$$

Ejemplo

Si el alfabeto es $\Sigma = \{a, b, c\}$, $\alpha = ab$ es una cadena, y entonces $a \circ ab = aab$ es también una cadena.

Cadena nula λ

Es el neutro de la concatenación:

$$\forall a \in \Sigma,\ a \circ \lambda = a$$

Clausura de Kleene de Σ: Σ^*

- $\lambda \in \Sigma^*$
- $\alpha \in \Sigma^* \Rightarrow a \circ \alpha \in \Sigma^*,\ a \in \Sigma$

Clausura positiva de Σ: Σ^+

Si $\alpha \in \Sigma^*$entonces $a \circ \alpha \in \Sigma^+,\ a \in \Sigma$.

Ejemplo

Sea $\Sigma = \{a, b, c\}$, entonces $ccba \in \Sigma^*$ porque $\lambda \in \Sigma^*$, $a \circ \lambda = a \in \Sigma^*$, $b \circ a = ba \in \Sigma^*$, $c \circ ba = cba \in \Sigma^*$, y $c \circ cba = ccba \in \Sigma^*$.

Lenguaje sobre un alfabeto Σ: Conjunto de cadenas sobre un alfabeto Σ.

Ejemplo:

- ϕ es un lenguaje
- $\{\lambda\}$ es un lenguaje ($\neq$ de ϕ)
- Dado $\Sigma = \{0, 1\}$, $\{0, 01, 011, 0111, 01111, \ldots\ldots\}$, es un lenguaje sobre Σ.

Concatenación de lenguajes

Sea L_1 un lenguaje definido sobre el alfabeto Σ_1, y sea L_2 un lenguaje definido sobre el alfabeto Σ_2, se define la concatenacion de L_1 y L_2 como el lenguaje

$$L_1L_2 = \{xy : x \in L_1 \wedge y \in L_2\},$$

definido sobre el alfabeto $\Sigma_1 \cup \Sigma_2$.

Clausura de Kleene L^*

Se define por:

- $L^0 = \{\lambda\}$
- $L^n = LL^{n-1}$ para $n \geq 1$
- $L^* = \bigcup_{n \geq 0} L^n$.

Clausura positiva L^+

Se define por:

$$L^+ = \bigcup_{n \geq 1} L^n.$$

De lo anterior se ve que $L^+ = LL^* = L^*L$, y que $L^* = L^+ \cup \{\lambda\}$. También se ve que, si L es un lenguaje definido sobre Σ, entonces, $L \subseteq \Sigma^*$.

Gramáticas

Definición

Una gramática es una 4-upla $G = < V_N, V_T, P, S >$ donde

- V_N es un conjunto de símbolos llamados no-terminales (o también, variables o categorías sintácticas)
- V_T es un conjunto de símbolos terminales (tal como lo era Σ en los ejemplos anteriores)
- P es el conjunto de "producciones", el cual es un subconjunto finito de

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

  estas producciones son entonces pares ordenados (α, β), que usualmente son notados como $\alpha \to \beta$.
- $S \in V_N$ es el símbolo distinguido de V_N.

Forma sentencial de una gramática $G = \langle V_N, V_T, P, S \rangle$

- S es una forma sentencial de G.
- Si $\alpha\beta\gamma$ es una forma sentencial de G, y $(\beta \to \delta) \in P$, entonces $\alpha\delta\gamma$ es también una forma sentencial de G.

Derivación directa en G

Si $\alpha\beta\gamma \in (V_N \cup V_T)^*$ y $(\beta \to \delta) \in P$, se dice que $\alpha\delta\gamma$ se deriva directamente en G de $\alpha\beta\gamma$ y se denota como

$$\alpha\beta\gamma \underset{G}{\to} \alpha\delta\gamma$$

Definición

Denotaremos con $\underset{G}{\overset{+}{\rightarrow}}$ y con $\underset{G}{\overset{*}{\rightarrow}}$ a las clausura transitiva y a la clausura transitiva y reflexiva de $\underset{G}{\rightarrow}$ respectivamente.

Definición

Denotaremos con $\underset{G}{\overset{k}{\rightarrow}}$ a la potencia k de la relación $\underset{G}{\rightarrow}$.

Definición

Lenguaje generado por una gramática $G = \langle V_N, V_T, P, S \rangle$, el cual se denotará como $\mathcal{L}(G)$,

$$\mathcal{L}(G) = \left\{ \alpha \in V_T^* : S \underset{G}{\overset{+}{\rightarrow}} \alpha \right\}$$

Clasificación de gramáticas (Chomsky)

Gramáticas regulares (tipo 3)

- Si todas las producciones son de la forma $A \to xB$ o $A \to x$, donde $A, B \in V_N$ y $x \in V_T^*$, entonces la gramática es llamada "lineal a derecha".
- Si todas las producciones son de la forma $A \to Bx$ o $A \to x$, donde $A, B \in V_N$ y $x \in V_T^*$, entonces la gramática es llamada "lineal a izquierda".

ambos tipos de gramática son llamados regulares.

Clasificación de gramáticas (Chomsky)

Gramáticas regulares (tipo 3)

Una forma alternativa de escribir las gramáticas regulares es el siguiente:

- Todas las producciones son de la forma $A \to aB$ o $A \to a$ o $A \to \lambda$, donde $A, B \in V_N$ y $a \in V_T$, para el caso de gramáticas "lineales a derecha".
- Todas las producciones son de la forma $A \to Ba$ o $A \to a$ o $A \to \lambda$, donde $A, B \in V_N$ y $a \in V_T$, para el caso de gramáticas "lineales a izquierda".

ambos tipos de gramática son llamados regulares.

Ejemplo de gramática tipo 3 (regular):

Gramática

Gramática para generar $\left\{a^n b^m c^k : m, n, k \geq 1\right\}$,
$G = \langle \{S, A, B, C\}, \{a, b, c\}, S, P \rangle$, donde P está dado por

$$
\begin{array}{lll}
S \rightarrow aA & B \rightarrow bB & C \rightarrow cC \\
A \rightarrow aA & B \rightarrow bC & C \rightarrow c \\
A \rightarrow bB & &
\end{array}
$$

Ejemplo: derivación de la cadena $aabbbbccc$

$$
\begin{array}{llll}
S & \rightarrow aA & \rightarrow aaA & \rightarrow aabB \\
 & \rightarrow aabbB & \rightarrow aabbbB & \rightarrow aabbbbC \\
 & \rightarrow aabbbbcC & \rightarrow aabbbbccC & \rightarrow aabbbbccc
\end{array}
$$

Clasificación de gramáticas (Chomsky)

Gramáticas independientes del contexto (libres de contexto, tipo 2)

Cada producción es de la forma $A \to \alpha$, donde $A \in V_N$ y $\alpha \in (V_N \cup V_T)^*$.

De la definición anterior puede inferirse que toda gramática regular es independiente de (o libre del) contexto

Ejemplo de gramática tipo 2 (independiente del contexto)

Gramática

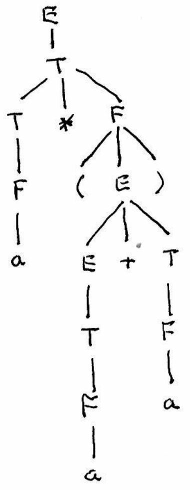
$G = \langle \{E, T, F\}, \{a, +, *, (,)\}, E, P \rangle$, donde P está dado por

$$
\begin{aligned}
E &\to E + T \\
E &\to T \\
T &\to T * F \\
T &\to F \\
F &\to (E) \\
F &\to a
\end{aligned}
$$

derivación de $a * (a + a)$

$$
\begin{aligned}
E &\to T \to T * F \\
&\to T * (E) \to F * (E) \\
&\to a * (E) \to a * (E + T) \\
&\to a * (T + T) \to a * (F + T) \\
&\to a * (a + T) \to a * (a + F) \\
&\to a * (a + a)
\end{aligned}
$$

Arbol de derivación



Dependientes del contexto (Sensitivas al contexto, tipo 1)

Cada producción es de la forma $\alpha \to \beta$, donde $\alpha, \beta \in (V_N \cup V_T)^*$ y $|\alpha| \leq |\beta|$. (Notar que esto impide la generación de la cadena nula λ)

De la definición anterior puede inferirse que:toda gramática independiente del (o libre del) contexto, que no posea reglas borradoras (o sea, reglas del tipo $A \to \lambda$, es también una gramática dependiente del (o sensitiva al) contexto

Ejemplo de gramática tipo 1 (dependiente del contexto):

Gramática

$G = \langle \{S, B, C\}, \{a, b, c\}, S, P \rangle$, donde P está dado por

$$\begin{array}{llll} S \to aSBC & CB \to BC & bB \to bb & cC \to cc \\ S \to abC & & bC \to bc & \end{array}$$

derivación de $aaabbbccc$:

$$\begin{array}{lll} S \to & & \\ \to aSBC & \to aaSBCBC & \to aaabCBCBC \\ \to aaabBCCBC & \to aaabBCBCC & \to aaabBBCCC \\ \to aaabbBCCC & \to aaabbbCCC & \to aaabbbcCC \\ \to aaabbbccC & \to aaabbbccc & \end{array}$$

Sin restricciones (tipo 0)

No poseen ninguna restricción como las anteriores.

Un lenguaje generado por una gramática tipo t es llamado "lenguaje t", p.ej.: un lenguaje generado por una gramática independiente del contexto es llamado también: "lenguaje independiente del contexto".

Alternativamente estos lenguajes son llamados tipo 3, 2, 1 y 0, respectivamente, p.ej.: un lenguaje generado por una gramática regular es llamado también: "lenguaje tipo 3".

El conjunto de las gramáticas tipo 0 incluye a todas las gramáticas.

Ejemplo de gramática tipo 0 (sin restricciones):

Gramática

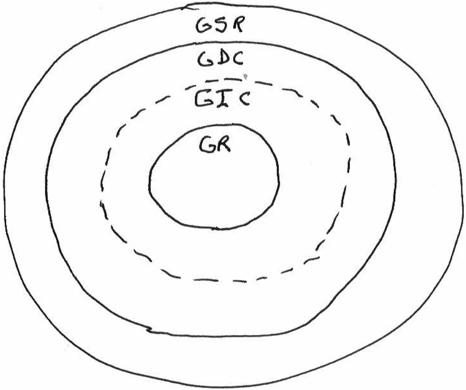
Gramática para generar $\{ww : w \in \{a,b\}^*\}$
$G = (\{S, A, B, C, D\}, \{a, b\}, S, P)$, donde P está dado por

$$\begin{array}{lllll} S \to CD & AD \to aD & Aa \to aA & Ba \to aB & C \to \lambda \\ C \to aCA & BD \to bD & Ab \to bA & Bb \to bB & D \to \lambda \\ C \to bCB & & & & \end{array}$$

derivación de $abaaabaa$

$$\begin{array}{llll} S & \to CD & \to aCAD & \to abCBAD \\ & \to abaCABAD & \to abaaCAABAD & \to abaaAABAD \\ & \to abaaAABaD & \to abaaAAaBD & \to abaaAaABD \\ & \to abaaaAABD & \to abaaaAAbD & \to abaaaAbAD \\ & \to abaaabAAD & \to abaaabAaD & \to abaaabaAD \\ & \to abaaabaaD & \to abaaabaa & \end{array}$$

Jerarquía de Chomsky



GSR = Gramáticas sin Restricciones – Tipo 0

GDC = Gramáticas Dependientes del Contexto – Tipo 1

GIC = Gramáticas Independientes del Contexto – Tipo 2

GR = Gramáticas Regulares – Tipo 3