

Instituto Politécnico Nacional

Escuela Superior de Cómputo



Teoría computacional

Clase 12: Clasificación de gramáticas

Solicitado: Ejercicios 10: Clasificación de gramáticas

M. en C. Edgardo Adrián Franco Martínez
<http://computacion.cs.cinvestav.mx/~efranco>

[@efranco_escom](https://twitter.com/efranco_escom)
edfrancom@ipn.mx



Contenido

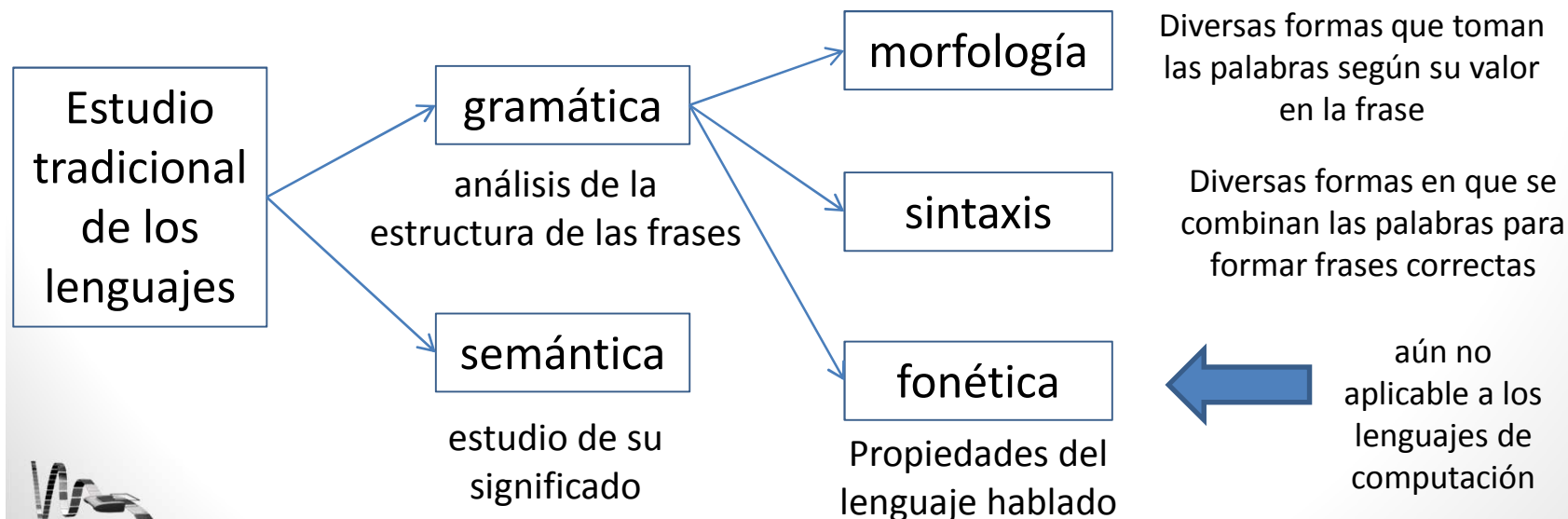
- Avram Noam Chomsky
- Jerarquía de Chomsky
 - Gramáticas tipo 3
 - Gramáticas tipo 2
 - Gramáticas tipo 1
 - Gramáticas tipo 0
- Clasificación de una gramática
- Descripción de las gramáticas
- Ejercicios 10: Clasificación de gramáticas



Avram Noam Chomsky

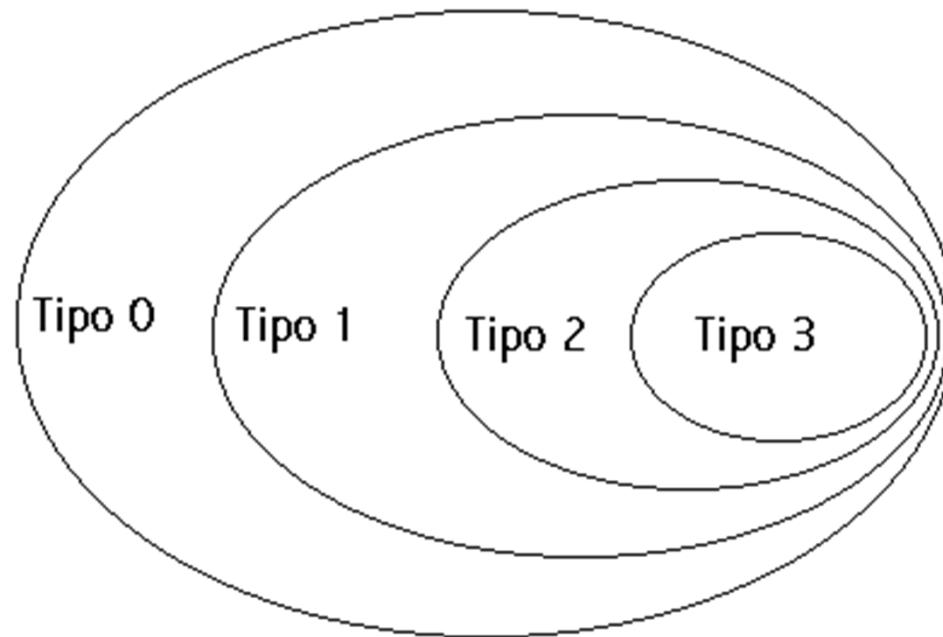


➤ En 1950 el lingüista norteamericano Avram Noam Chomsky introdujo la *teoría de las gramáticas transformacionales* o *teoría de lenguajes formales*, que convirtió la lingüística en una ciencia y proporcionó una herramienta que no sólo podía aplicarse a los lenguajes naturales, sino que facilitaba el estudio y la **formalización de los lenguajes para la programación** de computadoras (1960).





- En función de la forma de sus producciones, se puede caracterizar qué tan compleja es una gramática formal.
- Noam Chomsky mostró que esta caracterización clasifica jerárquicamente a las gramáticas formales: Gramáticas en un nivel están incluidas en los siguientes niveles y la inclusión entre niveles es propia.



Jerarquía de Chomsky



Gramáticas Tipo 3 (*gramáticas regulares*)

- Generan los **lenguajes regulares**. Las reglas (producciones) se restringen a un único no terminal en la parte izquierda y una parte derecha compuesta por un único terminal que puede estar seguido o no de un único no terminal. Es decir, normas del tipo:

$$A \rightarrow aB$$

$$A \rightarrow a$$

- Estos lenguajes son los que pueden ser decididos por un **autómata finito (*regular*)**. Los lenguajes regulares se utilizan para definir estructura léxica de los lenguajes de programación. **Definen la sintaxis** de los **identificadores**, **números**, **cadenas** y otros **elementos básicos** del lenguaje.



Gramáticas Tipo 2 (*independientes o libres de contexto*)



- Generan los lenguajes libres de contexto. Están definidas por reglas de la forma:

$$A \rightarrow \gamma$$

- A es un no terminal
- γ es una cadena de terminales y no terminales.
- Se denominan independientes de contexto porque A puede sustituirse por γ independientemente de las cadenas por las que esté acompañada.
- Estos lenguajes son todos los lenguajes que pueden ser reconocidos por los autómatas de pila.
 - Los lenguajes independientes de contexto **constituyen la base teórica para la sintaxis de la mayoría de los lenguajes de programación**. Definen la sintaxis de las declaraciones, las **proposiciones**, las **expresiones**, etc. (*i.e. la estructura de un programa*).



Gramáticas Tipo 1 (*dependientes de contexto*)



- Generan los lenguajes dependientes de contexto. Contienen reglas de producción de la forma:
$$\alpha A \beta \rightarrow \alpha \gamma \beta$$
 - **A** es un no terminal
 - α , β y γ **son** cadenas de terminales y no terminales.
 - α y β pueden ser vacíos, pero γ ha de ser distinto del vacío.
- Se denominan gramáticas dependientes del contexto, porque, como se observa, A puede ser sustituido por γ si está acompañada de α por la izquierda y de β por la derecha.
- Estos lenguajes son todos los lenguajes que pueden ser reconocidos por autómatas lineales acotados (Maquina de Turing Determinista).



Gramáticas Tipo 0 (*sin restricciones*)

- Incluyen todas las gramáticas formales.
- El más general, al que pertenece la semántica de los lenguajes naturales y artificiales.
- A estos lenguajes no se les impone restricción alguna.
- Estos lenguajes son todos los lenguajes que pueden ser reconocidos por una *máquina de Turing*.





Clasificación de una gramática

- Se dice que un lenguaje es de tipo k [$k = 0, k = 1, k = 2, k = 3$] cuando existe una gramática de tipo k que genera ese lenguaje.
- La clasificación de la gramática será la correspondiente al tipo de la producción de menor clasificación.





Gramática	Lenguaje	Reglas de Producción	Si $\mu \rightarrow \varphi$, relación entre $ \mu $ y $ \varphi $	Solución
Tipo-0	Rekursivas	Sin restricciones		Máquinas de Turing
Tipo-1	Dependiente de contexto	$\alpha A \beta \rightarrow \alpha \gamma \beta$	$ \mu \leq \varphi $	Autómatas lineales acotados
Tipo-2	Independiente de contexto	$A \rightarrow \gamma$	$ \mu = 1$	Autómatas de pila
Tipo-3	Regular	$A \rightarrow aB$ $A \rightarrow a$	$ \mu = 1$	Autómatas finitos, regulares

$$G_3 \subset G_2 \subset G_1 \subset G_0$$



Descripción de las gramáticas

Gramáticas Regulares (tipo 3 o G_3)

- **Gramáticas Regulares (tipo 3 o G_3)**
 - El lado izquierdo consiste sólo de una variable.
 - El lado derecho consiste de
 - Un símbolo terminal seguido de una variable ó
 - Sólo un símbolo terminal ó
 - La cadena vacía.

P.g.: $A \rightarrow aB \mid a \mid \lambda$



Descripción de las gramáticas

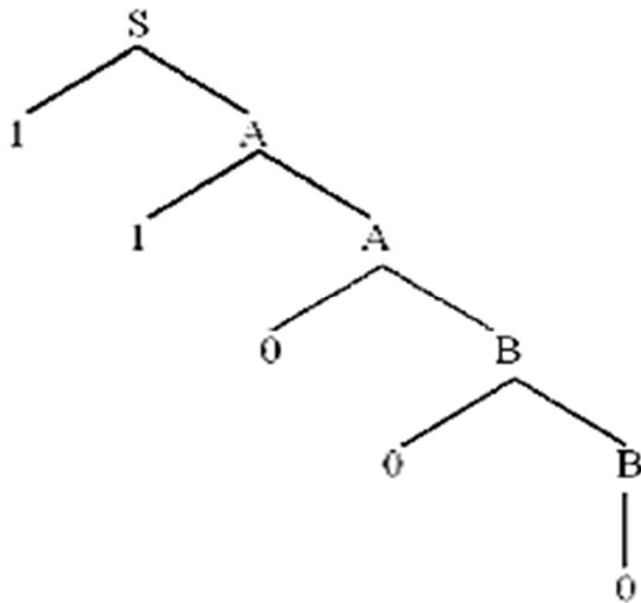
Gramáticas Regulares (tipo 3 o G_3)



Ejemplo: Las dos gramáticas G_r y G_l generan el lenguaje regular 11^*00^*

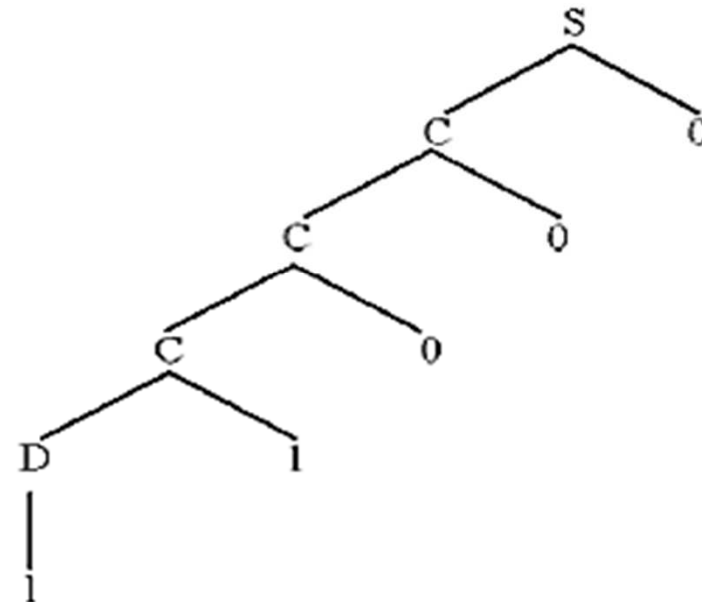
Regular por la derecha

$$G_r = \begin{cases} S \rightarrow 1A \\ A \rightarrow 1A \mid 0B \mid 0 \\ B \rightarrow 0B \mid 0 \end{cases}$$



Regular por la izquierda

$$G_l = \begin{cases} S \rightarrow C0 \\ C \rightarrow C0 \mid D1 \mid 1 \\ D \rightarrow D1 \mid 1 \end{cases}$$



Descripción de las gramáticas

Gramáticas Libres de Contexto, *GLC*, (tipo 2 o G_2)

- **Gramáticas Libres de Contexto, *GLC*, (tipo 2 o G_2)**
 - El lado **izquierdo** consiste **sólo de una variable**.
 - No hay restricciones para el lado derecho.

P.g.: $S \rightarrow aSb \mid ab \mid \lambda$



Descripción de las gramáticas

Gramáticas Libres de Contexto, GLC, (tipo 2 o G2)



Ejemplo: Las dos gramáticas G_1 y G_2 generan el lenguaje independiente del contexto $0^n 1^n 2^m$ con $n, m \geq 0$.

Lenguaje

GIC en formato no estricto

$$G_1 = \begin{cases} S \rightarrow AB \\ A \rightarrow 0A1 \mid \lambda \\ B \rightarrow 2B \mid \lambda \end{cases}$$

GIC en formato estricto

$$G_2 = \begin{cases} S \rightarrow AB \mid A \mid B \mid \lambda \\ A \rightarrow 0A1 \mid 01 \\ B \rightarrow 2B \mid 2 \end{cases}$$



Descripción de las gramáticas



Gramáticas Sensitivas al Contexto (tipo 1 o G_1)

- **Gramáticas Sensitivas al Contexto (tipo 1 o G_1)** A es un símbolo no terminal. Además, **las reglas son no-contractivas**, i.e. la longitud del lado izquierdo es menor o igual a la longitud del lado derecho. Esta propiedad de no-contracción garantiza que un lenguaje sensitivo al contexto no contiene λ .

P.g.:

$$S \rightarrow abc \mid aAbc$$

$$Ab \rightarrow Acb$$

$$Ac \rightarrow Bbcc$$

$$aB \rightarrow aa \mid aaA$$



Descripción de las gramáticas

Gramáticas Sensitivas al Contexto (tipo 1 o G_1)



Ejemplo: Sea la GDC $G = (\{a, b, c\}, \{S, M\}, S, P)$ donde

$$P = \begin{cases} S \rightarrow aMc \mid aSMc \\ aM \rightarrow ab \\ bM \rightarrow bb \\ cM \rightarrow Mc \end{cases}$$

La gramática G genera el lenguaje dependiente del contexto $a^n b^n c^n$ con $n > 0$. Un ejemplo de derivación sería:

$$S \Rightarrow aSMc \Rightarrow aaMcMc \Rightarrow aabcMc \Rightarrow aabMcc \Rightarrow aabbcc$$



Descripción de las gramáticas



Gramáticas sin restricción (tipo 0 o G_0)

- **Gramáticas sin restricción (tipo 0 o G_0)**, no hay restricciones para las reglas, excepto que el lado izquierdo no es λ .

P.g.:

$$\begin{aligned} S &\rightarrow aSBC \mid aBC \\ bB &\rightarrow bb \end{aligned}$$

$$\begin{aligned} CB &\rightarrow BC \\ bC &\rightarrow bc \end{aligned}$$

$$\begin{aligned} aB &\rightarrow ab \\ cC &\rightarrow cc \end{aligned}$$

$$A \rightarrow bc$$



Descripción de las gramáticas

Gramáticas sin restricción (tipo 0 o G_0)



Las gramáticas GEFs se caracterizan frente al resto (dejando aparte el caso $S \rightarrow \lambda$) en que admite reglas **compresoras**. Una regla compresora es aquella regla que cumple que el tamaño de su lado derecho es menor que el tamaño de su lado izquierdo.

Ejemplo: Sea la GEF $G = (\{a, b, c\}, \{S, M\}, S, P)$ donde

$$P = \begin{cases} S \rightarrow abMSc \mid \lambda \\ bMa \rightarrow abM \\ bMc \rightarrow bc \\ bMb \rightarrow bbM \end{cases}$$

La gramática G genera el lenguaje $a^n b^n c^n$ con $n \geq 0$. Un ejemplo de derivación sería:

$S \Rightarrow abMSc \Rightarrow abMabMSc \Rightarrow abMabMcc \Rightarrow abMabcc \Rightarrow aabMbcc \Rightarrow aabbMcc \Rightarrow aabbcc$





Ejercicios 10 “Clasificación de gramáticas”

- Clasificar las siguientes gramáticas dadas sus reglas de producción.

Ga	Gb	Gc	Gd	Ge	Gf	Gg
$Z \rightarrow yX$ $X \rightarrow y$ $X \rightarrow \lambda$ $yX \rightarrow yx$	$X \rightarrow xZyW$ $yW \rightarrow yx$ $Z \rightarrow vy$	$E \rightarrow E+T$ $E \rightarrow E-T$ $E \rightarrow T$ $T \rightarrow T * F$ $T \rightarrow T / F$ $T \rightarrow F$ $F \rightarrow (E)$ $F \rightarrow id$	$S \rightarrow aAbc$ $Ab \rightarrow bA$ $Ac \rightarrow Bbcc$ $bB \rightarrow bbaA$ $A \rightarrow aa$ $B \rightarrow bb$	$A \rightarrow bC$ $A \rightarrow bBC$ $bB \rightarrow bCa$ $C \rightarrow b$ $C \rightarrow \lambda$ $yCc \rightarrow yCc$	$S \rightarrow aAB A$ $A \rightarrow cBd$ $B \rightarrow e fS$ $C \rightarrow gD hDt$ $D \rightarrow x y z$	$S \rightarrow aS$ $S \rightarrow aN$ $N \rightarrow bN$ $N \rightarrow bM$ $N \rightarrow b$ $M \rightarrow c$

Se entregarán antes del día **Lunes 14 de Octubre de 2013
(23:59:59 hora limite)*

**Incluir la redacción de cada ejercicio*

**Portada y encabezados de pagina*

